Thesis for the degree of doctor of philosophy

Observability and identifiability of nonlinear systems with applications in biology

Milena Anguelova

Department of Mathematical Sciences
Division of Mathematics
Chalmers University of Technology and Göteborg University
Göteborg, Sweden 2007

Observability and identifiability of nonlinear systems with applications in biology
Milena Anguelova
ISBN 978-91-7385-035-3

Doktorsavhandlingar vid Chalmers Tekniska högskola
Ny serie nr 2716
ISSN 0346-718X

Department of Mathematical Sciences
Division of Mathematics
Chalmers University of Technology and Göteborg University
SE-412 96 Göteborg
Sweden
Telephone: +46 (0)31-772 1000

Printed at the Department of Mathematical Sciences
Göteborg, Sweden, 2007

Observability and identifiability of nonlinear systems with applications in biology

Milena Anguelova
Department of Mathematical Sciences
Chalmers University of Technology and Göteborg University

Abstract

This thesis concerns the properties of observability and identifiability of nonlinear systems. It consists of two parts, the first dealing with systems of ordinary differential equations and the second with delay-differential equations with discrete time delays.

The first part presents a review of two different approaches to study the observability of nonlinear ODE-systems found in literature. The differential-geometric and algebraic approaches both lead to the so-called rank test where the observability of a control system is determined by calculating the dimension of the space spanned by gradients of the time-derivatives of its output functions. We show that for analytic systems affine in the input variables, the number of time-derivatives of the output that have to be considered in the rank test is limited by the number of state variables.

Parameter identifiability is a special case of the observability problem. A case study is presented in which the parameter identifiability of a previously published kinetic model for the metabolism of *S. cerevisiae* (baker's yeast) has been analysed. The results show that some of the model parameters cannot be identified from any set of experimental data.

The general features of kinetic models of metabolism are examined and shown to allow a simplified identifiability analysis, where all sources of structural unidentifiability are to be found in single reaction rate expressions. We show how the assumption of an algebraic relation between concentrations in metabolic models can cause parameters to be unidentifiable.

The second part concerning delay systems begins by an introduction to the algebraic framework of modules over noncommutative rings. We then present both previously published and new results on the problem of observability. New results are shown on the problems of state elimination and characterisation of the identifiability of time-lag parameters. Their identifiability is determined by the form of the system's input-output representation. Linear-algebraic criteria are formulated to decide the identifiability of the delay parameters which eliminate the need for explicit computation of the input-output equations. The criteria are applied in the analysis of biological

models from the literature.

Keywords: Observability, identifiability, nonlinear systems, time delay, delay systems, state elimination, metabolism, conservation laws, signalling pathways.

Contents

This thesis consists of two parts, Part I and Part II, and includes Papers I, II, III and IV as follows:

Part I. Observability and identifiability of nonlinear ODE systems: General theory and a case study of metabolic models

- *Paper I.* Anguelova, M., Cedersund, G., Johansson, M., Franzén, C.J. and Wennberg, B. Conservation laws and unidentifiability of rate expressions in biochemical models, *IET Systems Biology, 2007, 1(4), pp. 230-237.*

Part II. Observability and identifiability for nonlinear systems of delay-differential equations with discrete time-delays

- *Paper II* Anguelova, M. and Wennberg, B. State elimination and identifiability of the delay parameter for nonlinear time-delay systems, *Automatica, 2007, to appear.*

- *Paper III* Anguelova, M. and Wennberg, B. Identifiability of the time-lag parameter in delay systems with applications to systems biology, *in Proc. of FOSBE 2007 (Foundation of Systems Biology in Engineering), Stuttgart, Germany, September 9-13, 2007.*

- *Paper IV* Anguelova, M. and Wennberg, B. State elimination and identifiability of delay parameters for nonlinear systems with multiple time-delays, *in Proc. of IFAC Workshop on Time Delay Systems, TDS'07 Nantes, France, September 17-19, 2007.*

Acknowledgements

First and foremost, I would like to thank my advisor Bernt Wennberg - for suggesting that I work on identifiability in the first place, which turned out to be such an interesting topic; for always matching my enthusiasm at every small progress; for managing to help me every time I got stuck; for all the good ideas that he came up with; for reading my manuscripts repeatedly at all times of the day and night.

I would also like to thank my co-advisor Carl Johan Franzén for his help with all the biochemistry- and biology-related questions and for spending a huge amount of time improving our manuscripts. For this, I would also like to thank my pair-project collaborator, Mikael Johansson.

The National Research School in Genomics and Bioinformatics and its coordinator Anders Blomberg are gratefully acknowledged for both financial and moral support of this project.

Thanks to my sister Iana Anguelova (who just happens to be a *real* mathematician) for help with the algebra, valuable discussions and comments to the manuscript.

Big thanks to Borys Stoew who sacrificed his time to read and comment the manuscript. This led to a momentous improvement in its appearance.

Thanks to Karin Kraft, Anna Nyström and Yulia Yurgens for the pleasant coffee-breaks.

I would like to thank my parents for their invaluable help with the logistics of life with children and for their support at all times.

Finally, I have Vessen, Maria and Emanuella to thank for my perfect life.

Preface

The work described in this thesis has been financed by the National Research School in Genomics and Bioinformatics under the project title *Large Scale Metabolic Modelling*. The work is also related to the more experimentally-oriented project *A metabolome and metabolic modeling approach to functional genomics*, also sponsored by the research school. The aim of the latter has involved the construction of metabolic models for the understanding of regulation and signal transduction within cells. The analysis of structural properties of metabolic models is the aim of this work and the properties that we have investigated are observability and identifiability. The biological models that we have come across have motivated the study of both ODE and delay systems and the division of the theoretical results of this thesis into two parts. Here follows a brief description of the latter.

Part I. Observability and identifiability of nonlinear ODE systems: General theory and a case study of metabolic models

This part of the thesis concerns the observability and identifiability problem for nonlinear systems of ordinary differential equations with applications in the kinetic modelling of metabolism in yeast. It consists of a monograph and one paper, Paper I, see below. The monograph reviews already published work before describing some new results and introduces a case study of a kinetic model of glycolysis from the literature. Of a particular interest for the study are enzymatic rate equations and how they are parameterised. This is discussed further in Paper I, which is briefly introduced in the monograph.

Part II. Observability and identifiability for nonlinear systems of delay-differential equations with discrete time-delays

This part of the thesis concerns some control problems for nonlinear time-delay systems, such as observability, identifiability and state elimination and application of the results to biological models from the literature. It consists of a monograph and three papers, Paper II, III and IV. The monograph introduces a mathematical framework for control based on modules over noncommutative rings before describing both previously published and new results on the observability problem. A new result on state elimination is shown, which leads to a characterisation of the identifiability of the delay parameters, the main result of this part of the thesis. The monograph concludes by an application of the result to a model of genetic regulation from the literature. Application to other models from systems biology can be found in Paper III. The theoretical results from Papers II and IV are described in the monograph part, omitting detailed derivations, for which the reader is referred to the papers themselves.

Part I

Observability and identifiability of nonlinear ODE systems

General theory and a case study of metabolic models

Milena Anguelova

Abstract

Observability is a structural property of a control system defined as the possibility to deduce the state of the system from observing its input-output behaviour.

We present a review of two different methods to test the observability of nonlinear control systems found in literature. The differential geometric and algebraic approaches have been applied to different classes of control systems. Both methods lead to the so-called rank test where the observability of a control system is determined by calculating the dimension of the space spanned by gradients of the time-derivatives of its output functions. It has been shown previously that for rational systems with n state-variables, only the first $n-1$ time-derivatives have to be considered in the rank test. In this work, we show that this result applies for a broader class of analytic systems.

The rank test can be used to determine parameter identifiability which is a special case of the observability problem. A case study is presented in which the parameter identifiability of a previously published kinetic model for the metabolism of *S. cerevisiae* (baker's yeast) has been analysed. The results show that some of the model parameters cannot be identified from any set of experimental data.

The general features of kinetic models of metabolism are examined and shown to allow a simplified identifiability analysis, where all sources of structural unidentifiability are to be found in single reaction rate expressions. We show how the assumption of an algebraic relation between concentrations in metabolic models can cause parameters to be unidentifiable. A general method is presented to determine whether a conserved moiety renders a given rate expression unidentifiable and to reparameterise it into identifiable parameters.

Keywords: Nonlinear systems, observability, identifiability, observability rank condition, metabolic model, kinetic model, metabolism, glycolysis, *Saccharomyces cerevisiae*

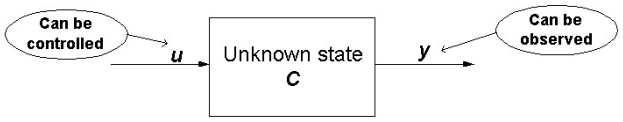
Contents

1 Introduction

1.1 Motivation for studying observability and identifiability

Consider a culture of yeast cells grown in a reactor and the chemical reactions that take place in their metabolism. Thus far, the possibility to observe what occurs inside a single cell as far as metabolic fluxes are concerned is very limited. It is therefore not unnatural to consider the cell as a box where we see what goes in (nutrients) and what comes out (secreted products), but not what happens inside.

There is, however, extensive knowledge of the chemistry and biology that takes place within the cell, and based on that, models are made for the transformation that occurs inside the box. In preparation for a mathematical description of the situation, we transform the above picture as follows:



We now label the part that can be controlled - for example, the amount of food given to the cells - u and call it "input", or "control variable". The input often varies with time and is thus a function $u(t)$. The part that can be observed over some interval of time - e.g. the different secreted products the fluxes of which can be measured - is denoted by $y(t)$ and called "output". What happens inside the cell is accounted for in terms of changes in the concentrations of the different chemical species present with respect to time; these concentrations are referred to as "state-variables" and denoted by $c(t)$. We also have a number of parameters that come with the model used for cellular metabolism, denoted by p. In this first part of the thesis, we assume that the future concentrations of the chemical species c depend only on their present concentrations and those of the inputs. Thus, the history of the cell does not matter and the changes in the concentrations with respect to time

can be described by ordinary differential equations. The following continuous "state-space" model can be formulated:

$$\begin{cases} \dot{p}(t) &= 0 \\ \dot{c}(t) &= f(c(t), p, u(t)) \\ y(t) &= g(c(t), p(t)) \quad , \end{cases} \tag{1.1}$$

where $\dot{c}(t)$ denotes the time-derivative of the state-variables at time t. A hypothetical setting is assumed where we start feeding an input $u(t)$ to the cell at time zero when the system is at an unknown state $c(0)$ and we observe the cell's behaviour in terms of the outputs produced. It is assumed that u is a function of time that we can choose, and that the values of y and all its time-derivatives at the starting point (time zero) can be measured. The variation with time, (t), will not be explicitly written when it is clear from the context.

It is often the case that metabolic models contain numerous parameters with unknown *in vivo* values. Sometimes, for the purpose of simulation, the latter are approximated by their *in vitro* values, see for example (Teusink et al., 2000). Often, however, one is interested in obtaining the values that fit a given set of experimental data. Thus, the parameters are estimated based on observing the input-output behaviour of the system. The property of identifiability is the possibility to define the values of the model parameters uniquely in terms of known quantities, that is, inputs, outputs and their time-derivatives.

1.2 Problem statement

A generalisation of identifiability is the property of observability. Consider the following "control system" which generalises the example above:

$$\Sigma \doteq \begin{cases} \dot{x}(t) &= f(x, u) \\ y &= h(x, u) \quad . \end{cases} \tag{1.2}$$

In this system, x are the state-variables, the inputs are denoted by u and the outputs by y; all their components are functions of time. Note that parameters can be considered state-variables with time-derivative zero. We have no knowledge of the initial conditions for the state-variables (or, respectively, of the parameter values). It is assumed that we have a perfect measurement of the outputs so that they are known as functions of time in some interval and all their time-derivatives at time zero can be calculated. The observability problem consists of investigating whether there exist relations binding the state-variables to the inputs, outputs and their time-derivatives and thus

locally defining them uniquely in terms of controllable/measurable quantities without the need for knowing the initial conditions. If no such relations exist, the initial state of the system cannot be deduced from observing its input-output behaviour. In the biological setting above, for instance, this can mean that there are infinitely many parameter sets that produce exactly the same output for every input and thus the model parameters cannot be estimated from any experimental measurements.

Before we define the problem of observability, consider the following example of a control system taken from Sedoglavic (2002):

$$\begin{array}{rcl} \dot{x}_1 & = & \frac{x_2}{x_1} \\ \dot{x}_2 & = & \frac{x_3}{x_2} \\ \dot{x}_3 & = & x_1\theta - u \\ y & = & x_1 \quad . \end{array} \tag{1.3}$$

In this system, x_1, x_2 and x_3 are state-variables, θ is a parameter, there is a single input u and a single output y. In the following we use capital letters to denote initial values of a function and its derivatives, i.e. $u^{(r)}(0) = U^{(r)}$, $y^{(r)}(0) = Y^{(r)}$ for $r \geq 0$. By computing time-derivatives of the output at time zero, we obtain the equations:

$$Y^{(1)} = \dot{x}_1 = \frac{x_2}{x_1} \tag{1.4}$$

$$Y^{(2)} = \ddot{x}_1 = \frac{\dot{x}_2 x_1 - \dot{x}_1 x_2}{x_1^2} = \frac{\frac{x_3}{x_2}x_1 - \frac{x_2}{x_1}x_2}{x_1^2} = \frac{x_3}{x_1 x_2} - \frac{x_2^2}{x_1^3} \tag{1.5}$$

$$\begin{array}{rcl} Y^{(3)} & = & x_1^{(3)} = \frac{\dot{x}_3 x_1 x_2 - x_3(\dot{x}_1 x_2 + x_1 \dot{x}_2)}{x_1^2 x_2^2} - \frac{2x_2\dot{x}_2 x_1^3 - x_2^2 3x_1^2 \dot{x}_1}{x_1^6} = \\ & = & \frac{(x_1\theta - U^{(0)})x_1 x_2 - x_3(\frac{x_2}{x_1}x_2 + x_1\frac{x_3}{x_2})}{x_1^2 x_2^2} - \frac{2x_2\frac{x_3}{x_2}x_1^3 - x_2^2 3x_1^2\frac{x_2}{x_1}}{x_1^6} = \\ & = & \frac{\theta}{x_2} - \frac{U^{(0)}}{x_1 x_2} - \frac{x_3^2}{x_1 x_2^3} - \frac{3x_3}{x_1^3} + \frac{3x_2^3}{x_1^5} \quad . \end{array} \tag{1.6}$$

For this simple example, it is actually possible to explicitly calculate the initial values of the state-variables and the parameters in terms of the inputs and outputs and their time-derivatives at time zero as shown in Sedoglavic (2002):

$$x_1 = Y^{(0)} \tag{1.7}$$

$$x_2 = Y^{(0)}Y^{(1)} \tag{1.8}$$

$$x_3 = Y^{(0)}Y^{(1)}\left((Y^{(1)})^2 + Y^{(0)}Y^{(2)}\right) \tag{1.9}$$

$$\theta = \frac{\left((Y^{(1)})^2 + Y^{(0)}Y^{(2)}\right)^2 + Y^{(0)}Y^{(1)}(3Y^{(1)}Y^{(2)} + Y^{(0)}Y^{(3)}) - U^{(0)}}{Y^{(0)}} \tag{1.10}$$

A given input-output behaviour thus corresponds to a unique state of the system. In general, we are not going to demand a globally unique state. It is enough that the equations have a finite number of solutions each defining a locally unique state. The observability problem concerns the existence of such relations and not the explicit calculation of the state variables from the equations. Depending on the theoretical approach, different definitions of observability can be given, as shown in this report.

1.3 Organisation of the report

In this work, a method for investigating the observability of certain classes of nonlinear control systems is described by using different theoretical points of view, each of which adds to our understanding of the problem.

Sections 2 and 3 present a survey of the theory on nonlinear observability available in the literature. Observability has been dealt with in both a differential geometrical interpretation, and an algebraical one. The two approaches are introduced and the results in terms of obtaining an observability test are described.

Section 4 attempts to answer the following question that arises during the literature surveys. If the derived observability test is to be applied in practice, a bound must be introduced for the number of time-derivatives of the output that have to be considered in obtaining equations for the variables. Such an upper bound is given for rational systems in Section 3. In Section 4 this bound is shown to apply for analytic systems.

Section 5 describes the identifiability problem as a special case of observability.

In Section 6 we apply the theory discussed in the preceding sections to a case study of a kinetic model for the metabolism of *Saccharomyces cerevisiae*, also known as bakers yeast. We use an algorithm by Sedoglavic (2002) and its implementation in Maple which performs an observability/identifiability test of rational models. We obtain results for the identifiability of the kinetic model and find the non-identifiable parameters. The results are interpreted in terms of the biological structure of the model.

The case study in Section 6 leads us to consider whether the special structure of metabolic models allows for a simplified identifiability test, in which only individual reaction rate expressions need to be analysed. Assumptions of conserved moieties of chemical species, often used in kinetic modelling of metabolism, are shown to lead to unidentifiable rate expressions, and in turn, to unidentifiable parameters in the models. This is discussed in detail in Paper I, where we also show how the models can be reparameterised into identifiable rate expressions.

2 The differential-geometric approach to nonlinear observability

In this section we present the basics of the theory of nonlinear observability in a differential-geometric approach that we have gathered from the works of Hermann and Krener (1977); Krener (1985); Isidori (1995); Sontag (1991) and Sussmann (1979).

2.1 Definitions

Throughout this section we will consider control systems affine in the input variables which is a valid description of many real-world systems. They have the form:

$$\Sigma \begin{cases} \dot{x}(t) &= f(x(t),u(t)) = g^0(x(t)) + g(x(t))u(t) \\ y(t) &= h(x(t)) \quad , \end{cases} \tag{2.1}$$

where $u(t)$ denotes the input, $x(t)$ the state variables and $y(t)$ the outputs (measurements). Throughout the text, the time-dependence (t) will not be written explicitly where it is understood from the context. We assume that $x \in M$ where M is an open subset of $\mathbb{R}^n$, $u \in \mathbb{R}^m$, $y \in \mathbb{R}^p$ and g^0, and the m columns of g, denoted by g^i for $i = 1, \ldots, m$, are analytic vector fields defined on M. We also have to assume that the system is complete, that is, for every bounded measurable input $u(t)$ and every $x^0 \in M$ there exists a solution to $\dot{x}(t) = f(x(t),u(t))$ such that $x(0) = x^0$ and $x(t) \in M$ for all $t \in \mathbb{R}$.

Here follow several definitions. Let W denote an open subset of M.

Definition 2.1 *A pair of points x^0 and x^1 in M are **W-distinguishable** if there exists a measurable bounded input $u(t)$ defined on the interval [0,T] that generates solutions $x^0(t)$ and $x^1(t)$ of $\dot{x} = f(x,u)$ satisfying $x^i(0) = x^i$ such that $x^i(t) \in W$ for all $t \in [0,T]$ and $h(x^0(t)) \neq h(x^1(t))$ for some $t \in [0,T]$. We denote by $I(x^0,W)$ all points $x^1 \in W$ that are not **W-distinguishable** from x^0.*

Definition 2.2 *The system Σ is **observable** at $x^0 \in M$ if $I(x^0,M) = x^0$.*

If a system is observable according to the above definition, it is still possible that there is an arbitrarily large interval of time in which two points of M cannot be distinguished form each other. Therefore a local concept is introduced which guarantees that to distinguish between the points of an open subset W of M, we do not have to go outside of it, which necessarily sets a limit to the time interval as well.

Definition 2.3 *The system Σ is **locally observable** at $x^0 \in M$ if for every open neighborhood W of x^0, $I(x^0, W) = x^0$.*

Clearly, local observability implies observability as we can set W in Definition 2.3 equal to M. On the other hand, since W can be chosen arbitrarily small, local observability implies that we can distinguish between neighbouring points instantaneously (since the trajectory is bound to be within W, setting a limit to the time interval).

Remark: In this section "local observability" is a stronger property than "observability" because it implies that only *local* information is needed.

Both the definitions above ensure that a point $x^0 \in M$ can be distinguished from every other point in M. For practical purposes though, it is often enough to be able to distinguish between neighbours in M, which leads us to the following two concepts:

Definition 2.4 *The system Σ has the **distinguishability property** at $x^0 \in M$ if x^0 has an open neighborhood V such that $I(x^0, M) \cap V = x^0$.*

In a system having this property, any point x^0 can be distinguished from neighbouring points but there could be arbitrarily large intervals of time $[0, T]$ in which the points cannot be distinguished. In order to set a limit on the time interval, a stronger concept is introduced:

Definition 2.5 *The system Σ has the **local distinguishability property** at $x^0 \in M$ if x^0 has an open neighbourhood V such that for every open neighbourhood W of x^0, $I(x^0, W) \cap V = x^0$.*

Clearly, local observability implies local distinguishability as we can set V equal to M. Thus, if a system does not have the local distinguishability property at some x^0, it is not locally observable at that point either.

It is the final property of local distinguishability that lends itself to a test.

2.2 The observability rank condition (ORC)

This subsection describes how to determine if a system possesses the local distinguishability property by the so-called "observability rank condition" as introduced by Hermann and Krener (1977).

Throughout this subsection, we will use the following simple example of a control system:

$$\begin{cases} \dot{x}_1 &= 0 \\ \dot{x}_2 &= u - x_1 x_2 \\ y &= x_1 x_2 \quad . \end{cases} \tag{2.2}$$

For this system, $g^0(x_1,x_2) = \begin{pmatrix} 0 \\ -x_1x_2 \end{pmatrix}$, $g(x_1,x_2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $h(x_1,x_2) = x_1x_2$ (according to the notation introduced in the previous subsection) with $p=1$, $m=1$, and $n=2$.

We now introduce differentiation with respect to time along the system dynamics. Formally, this is done by so-called Lie-differentiation. The Lie derivative of a C^∞ function ϕ on M by a vector field v on M is

$$L_v(\phi)(x) :=< d\phi, v > \quad . \tag{2.3}$$

Here $<>$ denotes the scalar product and $d\phi$ the gradient of ϕ.

Applying this to our example system, note that $g^0(x_1,x_2)$ and $g(x_1,x_2)$ are vector fields on M and we can calculate the Lie derivative of $h(x_1,x_2)$ along them:

$$L_{g^0}(h)(x_1,x_2) =< dh, g^0 >= (x_2 \quad x_1) \begin{pmatrix} 0 \\ -x_1x_2 \end{pmatrix} = -x_1^2x_2 \tag{2.4}$$

and

$$L_g(h)(x_1,x_2) =< dh, g >= (x_2 \quad x_1) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = x_1 \quad . \tag{2.5}$$

The flow $\Phi(t,x)$ of a vector field v on M is by definition the solution of:

$$\begin{cases} \frac{\partial}{\partial t}\Phi(t,x) &= v(\Phi(t,x)) \\ \Phi(0,x) &= x \quad . \end{cases} \tag{2.6}$$

Observe that we have the following equality:

$$L_v(\phi)(x) = \left.\frac{d}{dt}\right|_{t=0} \big(\phi(\Phi(t,x)\big) \quad . \tag{2.7}$$

The Taylor series of $\phi(\Phi(t,x))$ with respect to t are called Lie series:

$$\phi\big(\Phi(t,x)\big) = \sum_{l=0}^{\infty} \frac{t^l}{l!} L_v^l(\phi)(x) \quad . \tag{2.8}$$

Let us now link the local distinguishability property to these new concepts. First of all, as observed in (Sussmann, 1979), if two points x^0 and x^1 in M are W-distinguishable by a bounded measurable input, then they must be W-distinguishable by a piecewise constant input. This is due to uniform convergence since the outputs depend continuously on the inputs. For a constant input u, $f(x,u)$ defines a vector field on M and we can define the flow $\Phi(t,x)$ and the Lie series expansion of $h_i(\Phi(t,x))$ for $i=1,\ldots,p$. To see

how this generalises to piecewise-constant inputs, we follow (Isidori, 1995) and consider the input such that for $i = 1, \ldots, m$,

$$\begin{cases} u_i(t) = u_i^1, & t \in [0, t_1) \\ u_i(t) = u_i^l, & t \in [t_1 + \cdots + t_{l-1}, t_1 + \cdots + t_l), \quad l \geq 2 \end{cases} , \tag{2.9}$$

where $u_i^l \in \mathbb{R}$. With no loss of generality, we can assume that $t_1 = \cdots = t_l$. Define the vector fields

$$\theta_l = g^0 + gu^l \tag{2.10}$$

and denote their corresponding flows by Φ_t^l. Under this input, the state reached at time $t_1 + \cdots + t_l$ starting from x^0 at $t = 0$ can be expressed as

$$x(t_l) = \Phi_{t_l}^l \circ \cdots \circ \Phi_{t_1}^1(x^0) \quad . \tag{2.11}$$

The corresponding output becomes

$$y_i(t_1 + \cdots + t_l) = h_i\big(\Phi_{t_l}^l \circ \cdots \circ \Phi_{t_1}^1(x^0)\big) \quad . \tag{2.12}$$

This output can be regarded as the value of the mapping

$$\begin{aligned} F_i^{x^0} : \quad (-\epsilon, \epsilon)^l \quad &\to \mathbb{R} \\ (t_1, \ldots, t_l) \quad &\mapsto h_i \circ \Phi_{t_l}^l \circ \cdots \circ \Phi_{t_1}^1(x^0) \quad . \end{aligned} \tag{2.13}$$

If two initial states x^0 and x^1 are such that they produce the same output for all possible piecewise-constant inputs, then

$$F_i^{x^0}(t_1, \ldots, t_l) = F_i^{x^1}(t_1, \ldots, t_l) \tag{2.14}$$

for all possible $(t_1, \ldots, t_l)$ with $0 \leq t_j < \epsilon$ and all $i = 1, \ldots, p$. Thus,

$$\Big(\frac{\partial^l F_i^{x^0}}{\partial t_1 \cdots \partial t_l}\Big)_{t_1 = \cdots = t_l = 0} = \Big(\frac{\partial^l F_i^{x^1}}{\partial t_1 \cdots \partial t_l}\Big)_{t_1 = \cdots = t_l = 0} \quad . \tag{2.15}$$

Since

$$\Big(\frac{\partial^l F_i^{x^0}}{\partial t_1 \cdots \partial t_l}\Big)_{t_1 = \cdots = t_l = 0} = L_{\theta_1} \cdots L_{\theta_l} h_i(x^0) \quad , \tag{2.16}$$

we must have,

$$L_{\theta_1} \cdots L_{\theta_l} h_i(x^0) = L_{\theta_1} \cdots L_{\theta_l} h_i(x^1) \quad . \tag{2.17}$$

Suppose now that there exists an open neighbourhood V of x^0 such that all points in V are distinguishable from x^0 instantaneously (which is the requirement for local distinguishability). Then, there exists a piecewise-constant input u such that the map from V to the space spanned by the Lie derivatives $L_{\theta_1} \cdots L_{\theta_l} h_i$ is $1:1$. Let us formally describe the "observation" space

spanned by the $L_{\theta_1} \cdots L_{\theta_l} h_i$ which will be denoted by $\mathcal{G}$. It can be shown, (Isidori, 1995; Sontag, 1991), that

$$\mathcal{G} = \text{span}_{\mathbb{R}}\{L_{g^{i_1}} L_{g^{i_2}} \cdots L_{g^{i_r}}(h_i) : \quad r \geq 0, \quad i_j = 0, \ldots, m, \quad i = 1, \ldots, p\} \quad . \tag{2.18}$$

Since we are interested in the Jacobian of the $1:1$ map mentioned above, the space spanned by the gradients of the elements of $\mathcal{G}$ is introduced and denoted by $d\mathcal{G}$:

$$d\mathcal{G} = \text{span}_{\mathbb{R}_x}\{d\phi : \quad \phi \in \mathcal{G}\} \quad , \tag{2.19}$$

where $\mathbb{R}_x$ denotes the field of meromorphic functions on M.

It is the dimension of $d\mathcal{G}$ which determines the local distinguishability property. For each $x \in M$, let $d\mathcal{G}(x)$ be the subspace of the cotangent space at x obtained by evaluating the elements of $d\mathcal{G}$ at x. The rank of $d\mathcal{G}(x)$ is constant in M except at certain singular points, where the rank is smaller (this property is due to the system being analytic, see for example (Krener, 1985) or Chapter 3 in (Isidori, 1995). Then $\dim_{\mathbb{R}_x} d\mathcal{G}$ is defined as the generic or maximal rank of $d\mathcal{G}(x)$, that is, $\dim_{\mathbb{R}_x} d\mathcal{G} = max_{x \in M}(\dim_{\mathbb{R}} d\mathcal{G}(x))$.

We can now formulate the so-called "observability rank condition" introduced by Hermann and Krener (1977):

Theorem 2.1 *The system Σ has the **local distinguishability property** for all x in an open dense set of M if and only if $\dim_{\mathbb{R}_x} d\mathcal{G} = n$.*

Let us apply this test to the example system. We observe by inspection that the space $\mathcal{G}$ for this system is spanned by functions of the forms x_1^k and $x_1^k x_2$ (the first two Lie derivatives were calculated above). Thus, the space $d\mathcal{G}$ is spanned by one-forms of the type $(kx_1^{k-1} \qquad 0)$ and $(kx_1^{k-1}x_2 \qquad x_1^k)$. Therefore we conclude that this example system has the local distinguishability property almost everywhere except on the line $x_1 = 0$.

Consider another example:

$$\begin{cases} \dot{x}_1 &= u - x_1 \\ \dot{x}_2 &= u - x_2 \\ y &= x_1 + x_2 \quad . \end{cases} \tag{2.20}$$

For this system, $g^0(x_1, x_2) = \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix}$, $h(x_1, x_2) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $h(x_1, x_2) = x_1 + x_2$ (according to the previously used notation). The first two Lie derivatives are

$$L_{g^0}(h)(x_1, x_2) =< dh, g^0 >= (1 \quad 1) \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix} = -x_1 - x_2 \tag{2.21}$$

and

$$L_g(h)(x_1, x_2) =< dh, g >= (1 \quad 1) \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2 \quad . \tag{2.22}$$

The space $\mathcal{G}$ for this example is spanned by constant functions and the function $x_1 + x_2$. Thus, the space $d\mathcal{G}$ is spanned by one-forms of the type $(1 \quad 1)$ and $(0 \quad 0)$. Clearly, this space is of dimension 1, which means that the system does not have the local distinguishability property anywhere.

2.3 From piecewise-constant to differentiable inputs - a different definition of observation space

2.3.1 Observation space for analytic inputs

In the previous section, the observation space was defined in terms of piecewise-constant inputs to be:

$$\mathcal{G} = \text{span}_{\mathbb{R}}\{L_{g^{i_1}} L_{g^{i_2}} \cdots L_{g^{i_r}}(h_i) : \quad r \geq 0, \quad i_j = 0, \ldots, m, \quad i = 1, \ldots, p\} \quad . \tag{2.23}$$

In this subsection it is shown that the observation space can be defined equally well in terms of analytic inputs. We follow the works of Sontag (1991) and Krener (1985).

A time-dependent vector field $v(t, x)$ defines a time-dependent flow in a similar way as in the previous section:

$$\begin{cases} \frac{\partial}{\partial t}\Phi(t, x) &= v(t, \Phi(t, x)) \\ \Phi(0, x) &= x \quad . \end{cases} \tag{2.24}$$

Let $\Phi_u(t, x)$ denote the time-dependent flow corresponding to the time-dependent vector field $f(x(t), u(t))$, where we now assume that we have a single input u which is an analytic function of time (the results in this section can be generalised to apply for vector-valued inputs). Let the initial values of u and its derivatives be $u^{(r)}(0) = U^{(r)}$ for $r \geq 0$ with $U^{(r)} \in \mathbb{R}$. For any non-negative integer l and any $U = (U^{(0)}, \ldots, U^{(l-1)}) \in \mathbb{R}^l$, define the functions

$$\psi_{rm+i}(x, U) = \left.\frac{d^r}{dt^r}\right|_{t=0} g_i(\Phi_u(t, x)) \quad , \tag{2.25}$$

for $1 \leq i \leq p$, $0 \leq r \leq l - 1$. (Observe that the result of this formula is actually the Lie derivation defined earlier, where extra terms appear due to the time dependence of the input. In fact, $\psi_{rp+i}(x, U) = L_f^r h_i$ where we define $L_f = \sum_{j=1}^n f_j \frac{\partial}{\partial x_j} + \sum_{l=0} U^{(l+1)} \frac{\partial}{\partial u^{(l)}}$.) Applying repeatedly the chain rule, we

see that the functions ψ_i can be expressed as polynomials in $U^{(0)}, \ldots, U^{(l-1)}$ with coefficients that are functions of x, (Sontag, 1991).

As in Subsection 2.2, we can again define the Taylor series of $g(\Phi_u(t,x))$ with respect to t:

$$h_i(\Phi_u(t,x)) = \sum_{r=0}^{\infty} \psi_{rp+i}(x,U)\frac{t^r}{r!} \quad . \tag{2.26}$$

Similarly to Subsection 2.2, where we considered the space spanned by the coefficients of the Lie series for $h_i(\Phi_u(t,x))$, we now construct the space spanned by the ψ_j:

$$\hat{\mathcal{G}} = \mathrm{span}_{\mathbb{R}}\{\psi_{lp+i}(x,U): \quad U \in \mathbb{R}^l, l \geq 0, \quad i = 1, \ldots, p\} \quad . \tag{2.27}$$

Wang and Sontag (1989) proved that $\mathcal{G} = \hat{\mathcal{G}}$. We can illustrate this with the observable example from Subsection 2.2:

$$\left\{ \begin{array}{lcl} \dot{x}_1 & = & 0 \\ \dot{x}_2 & = & u - x_1 x_2 \\ y & = & x_1 x_2 \quad . \end{array} \right. \tag{2.28}$$

The time-dependent flow for the time-dependent vector field $f(x,u) = \begin{pmatrix} 0 \\ u - x_1 x_2 \end{pmatrix}$ becomes $\Phi_u(t,x) = \begin{pmatrix} \Phi_{u,1}(t,x) \\ \Phi_{u,2}(t,x) \end{pmatrix}$, where

$$\left\{ \begin{array}{lcl} \frac{\partial}{\partial t}\Phi_{u,1}(t,x) & = & 0 \\ \frac{\partial}{\partial t}\Phi_{u,2}(t,x) & = & u(t) - \Phi_{u,1}(t,x)\Phi_{u,2}(t,x) \\ \Phi_{u,1}(0,x) & = & x_1 \\ \Phi_{u,2}(0,x) & = & x_2 \quad . \end{array} \right. \tag{2.29}$$

The first few ψ_i:s can be calculated as follows:

$$\begin{array}{rcl} \psi_1(x,U) & = & h(\Phi_u(t,x))_{|t=0} = \big(\Phi_{u,1}(t,x)\Phi_{u,2}(t,x)\big)_{|t=0} = \\ & = & \Phi_{u,1}(0,x)\Phi_{u,2}(0,x) = x_1 x_2 \\ \psi_2(x,U) & = & \left.\frac{dh(\Phi_u(t,x))}{dt}\right|_{t=0} = \left.\frac{\partial\big(\Phi_{u,1}(t,x)\Phi_{u,2}(t,x)\big)}{\partial t}\right|_{t=0} = \\ & = & \big(\Phi_{u,2}(t,x)\frac{\partial\Phi_{u,1}(t,x)}{\partial t} + \Phi_{u,1}(t,x)\frac{\partial\Phi_{u,2}(t,x)}{\partial t}\big)_{|t=0} = \\ & = & \big(\Phi_{u,2}(t,x)\cdot 0 + \Phi_{u,1}(t,x)(u(t) - \Phi_{u,1}(t,x)\Phi_{u,2}(t,x))\big)_{|t=0} = \\ & = & \Phi_{u,1}(0,x)(U^{(0)} - \Phi_{u,1}(0,x)\Phi_{u,2}(0,x)) = x_1(U^{(0)} - x_1 x_2) \\ \psi_3(x,U) & = & \left.\frac{d^2h(\Phi_u(t,x))}{dt^2}\right|_{t=0} = \left.\frac{\partial^2\big(\Phi_{u,1}(t,x)\Phi_{u,2}(t,x)\big)}{\partial t^2}\right|_{t=0} = \\ & = & x_1\big(U^{(1)} - x_1(U^{(0)} - x_1 x_2)\big) \quad . \end{array} \tag{2.30}$$

We see now that if the $U^{(i)}$:s are free to vary over $\mathbb{R}$, then the space $\hat{\mathcal{G}}$ for this example is spanned by the functions x_1^k and $x_1^k x_2$ for $k \geq 1$, exactly as the space $\mathcal{G}$ that we calculated in Subsection 2.2.

2.3.2 Symbolic notation for the inputs

Following (Sontag, 1991), let us now consider the ψ_j:s as formal polynomials in $U^{(0)}, U^{(1)}, \ldots$ with coefficients that are functions of x. Denote by $\mathcal{K} = \mathbb{R}(U^{(0)}, U^{(1)}, \ldots)$ the field obtained by adjoining the indeterminates $U^{(0)}, U^{(1)}, \ldots$ to $\mathbb{R}$. Recall that $\mathbb{R}_x$ is the field of meromorphic functions on M. Define $\mathcal{K}_x = \mathbb{R}_x(U^{(0)}, U^{(1)}, \ldots)$ as the field obtained by adjoining (a finite number of) the indeterminates $U^{(0)}, U^{(1)}, \ldots$ to $\mathbb{R}_x$, where $\mathcal{K}_x$ is seen as a vector space over $\mathcal{K}$. Let $\mathcal{F}^{\mathcal{K}}$ be the subspace of $\mathcal{K}_x$ spanned by the functions ψ_j over $\mathcal{K}$, that is,

$$\mathcal{F}^{\mathcal{K}} = \text{span}_{\mathcal{K}}\{\psi_j : \quad j \geq 1\} \quad . \tag{2.31}$$

This is now a different definition of the "observation space". As before, we are also interested in the space spanned by the differentials of the elements of $\mathcal{F}^{\mathcal{K}}$. The latter can be seen as polynomial functions of $U^{(0)}, U^{(1)}, \ldots$ with coefficients that are covector fields on M. For the example in Subsection 2.3.1, the differentials of the ψ_j:s can be written:

$$\begin{aligned}
d\psi_1 &= (x_2 \qquad x_1) \\
d\psi_2 &= (U^{(0)} - 2x_1x_2 \qquad - x_1^2) = (1 \qquad 0)U^{(0)} + (-2x_1x_2 \qquad - x_1^2) \\
d\psi_3 &= (U^{(1)} - 2U^{(0)}x_1 + 3x_1^2x_2 \qquad x_1^3) = \\
&= (1 \qquad 0)U^{(1)} + (-2x_1 \qquad 0)U^{(0)} + (3x_1^2x_2 \qquad x_1^3) \quad .
\end{aligned}$$

Recall from the previous section that the space $d\mathcal{G}$ for this example is spanned by one-forms of the type $(kx_1^{k-1} \qquad 0)$ and $(kx_1^{k-1}x_2 \qquad x_1^k)$. The covector fields calculated above are clearly of the same form.

Now let

$$\mathcal{O}^{\mathcal{K}} = \text{span}_{\mathcal{K}_x}\{d\psi_i : \quad \psi_i \in \mathcal{F}^{\mathcal{K}}\} \quad . \tag{2.32}$$

Sontag (1991) proved the following result:

Theorem 2.2 *For the analytic system* (2.1)

$$\dim_{\mathbb{R}_x} d\mathcal{G} = \dim_{\mathcal{K}_x} \mathcal{O}^{\mathcal{K}} \quad . \tag{2.33}$$

Thus, the property of local distinguishability can be determined from the dimension of the space $\mathcal{O}^{\mathcal{K}}$. The significance of this result is that u can now be treated symbolically in calculating the rank. This observation is used in Section 4 to derive an upper bound for the number of $d\psi_j$ that have to be considered in the rank test.

3 The algebraic point of view: observability of rational models

This section introduces the algebraic point of view in the treatment of the observability problem according to the works of Diop and Fliess (1991a,b); Diop and Wang (1993) and Sedoglavic (2002).

3.1 Example

Before we describe the algebraic setting for our general control problem, consider the following simple example:

$$\left\{ \begin{array}{lcl} \dot{x}_1 & = & x_1 x_2^2 + u \\ \dot{x}_2 & = & x_1 \\ y & = & x_1 \quad . \end{array} \right. \tag{3.1}$$

We obtain two equations for the state-variables x_1 and x_2 from the output function and its first Lie derivative where we use the notations $u^{(r)}(0) = U^{(r)}$ and $y^{(r)}(0) = Y^{(r)}$, $r \geq 0$ for the time derivatives at zero of the input and output, respectively:

$$Y^{(0)} = x_1 \tag{3.2}$$

$$Y^{(1)} = L_f x_1 = x_1 x_2^2 + U^{(0)} \quad . \tag{3.3}$$

By simple algebraic manipulation of these equations, we can obtain the following polynomial equations for each of the variables with coefficients in $U = (U^{(0)}, U^{(1)}, \ldots)$ and $Y = (Y^{(0)}, Y^{(1)}, \ldots)$:

$$x_1 = Y^{(0)} \tag{3.4}$$

$$Y^{(0)} x_2^2 + U^{(0)} - Y^{(1)} = 0 \quad . \tag{3.5}$$

There are finitely many (two) solutions of these equations for a given set of inputs and outputs (except on the lines $x_1 = 0$ and $x_2 = 0$). Each one is locally unique and determines the state of the system completely from information on the input and output values. (In the terminology of Section 2, this example system has the local distinguishability property for all x except for those on the lines $x_1 = 0$ and $x_2 = 0$).

This was a very simple example where we could derive (and solve) these polynomial equations for the variables explicitly. In general, however, the observability problem concerns the existence of such equations rather than their explicit calculation. We now review an algebraic formulation of observability for control systems consisting of polynomial or rational expressions.

3.2 Algebraic setting

3.2.1 Algebraic observability

Consider now "polynomial" control systems of the form:

$$\Sigma \begin{cases} \dot{x} &= f(x,u) \\ y &= h(x,u) \end{cases}, \tag{3.6}$$

where u stands for the m input variables, f and h are for now vectors of n and p polynomial functions, respectively (we will make the transition to rational functions later).

The equations obtained by differentiating the output functions will now contain polynomial expressions only. This allows us to make a new definition of observability based on the following rather intuitive idea - the state-variable x_i, $i = 1, \ldots, n$ is observable if there exists an algebraic relation that binds x_i to the inputs, outputs and a finite number of their time-derivatives. If each x_i is the solution of a polynomial equation in U and Y, then we know that a given input-output map corresponds to a locally unique state of the system. We will now prepare for a formal definition of algebraic observability.

Let $\mathbb{R}\langle U, Y\rangle$ denote the field obtained by adjoining the indeterminates $U_i^{(0)}$, $U_i^{(1)}, \ldots$, $i = 1, \ldots, m$ and $Y_j^{(0)}, Y_j^{(1)}, \ldots$, $j = 1, \ldots, p$ to $\mathbb{R}$ (or any other field of characteristic zero). Then we can make the following definition of algebraic observability:

Definition 3.1 *x_i, $i \in \{1, \ldots, n\}$ is algebraically observable if x_i is algebraic over the field $\mathbb{R}\langle U, Y\rangle$. The system Σ is algebraically observable if the field extension $\mathbb{R}\langle U, Y\rangle \hookrightarrow \mathbb{R}\langle U, Y\rangle(x)$ is purely algebraic.*

3.2.2 Derivations and transcendence degree

The transcendence degree of the field extension $\mathbb{R}\langle U, Y\rangle \hookrightarrow \mathbb{R}\langle U, Y\rangle(x)$ is now equal to the number of non-observable state-variables which should be assumed known (i.e. should have known initial conditions) in order to obtain an observable system. Our purpose is now to find a way to calculate this transcendence degree. For this, the theory of derivations over subfields as described in (Jacobsson, 1980) and (Lang, 1993) is used.

Definition 3.2 *A derivation D of a ring R is a linear map $D : R \to R$ such that*

$$D(a+b) = D(a) + D(b) \tag{3.7}$$

$$D(ab) = aD(b) + D(a)b \quad , \tag{3.8}$$

for $a, b \in R$.

For example, the partial derivative $\frac{\partial}{\partial X_i}$, $i = 1, \ldots, n$, is a derivation of the polynomial ring $k[X_1, \ldots, X_n]$ over a field k.

Consider now a field F of characteristic 0 and a finitely-generated field extension $E = F(x) = F(x_1, \ldots, x_k)$. Can a derivation D of F be extended to a derivation D^* of E which coincides with D on F? Consider the ideal determined by (x) in $F[X]$ and denoted by I, that is, the set of polynomials in $F[X]$ vanishing on (x). If such a derivation D^* exists and $p(X) \in I$, then the following must hold:

$$0 = D(0) = D^*0 = D^*p(x) = p^D(x) + \sum_{i=1}^{n} \frac{\partial p}{\partial x_i} D^*x_i \quad , \tag{3.9}$$

where p^D denotes the polynomial obtained by applying D to all the coefficients of p (which are elements of F) and $\frac{\partial p}{\partial x_i}$ denotes the polynomial $\frac{\partial p}{\partial X_i}$ evaluated at (x). If the above is true for a set of generators of the ideal I, then it is satisfied by all polynomials in I. This is now a necessary condition for extending the derivation D to $E = F(x)$. It is also a sufficient condition as shown in (Jacobsson, 1980) and (Lang, 1993):

Theorem 3.1 *Let D be a derivation of a field F. Let $(x) = (x_1, \ldots, x_n)$ be a finite family of elements in an extension of F. Let $p_\alpha(X)$ be a set of generators for the ideal determined by (x) in $F[X]$. Then, if (w) is any set of elements of $F(x)$ satisfying the equations*

$$0 = p^D(x) + \sum_{i=1}^{n} \frac{\partial p_\alpha}{\partial x_i} w_i \quad , \tag{3.10}$$

there is one and only one derivation D^* of $F(x)$ coinciding with D on F and such that $D^*x_i = w_i$.

Suppose now that the derivation D on F is the trivial derivation, that is, $Dx = 0$ for all $x \in F$. Then, $p^D(x) = 0$ in the equation above and thus, $0 = \sum_{i=1}^{n} \frac{\partial p_\alpha}{\partial x_i} w_i$. The w_i:s are thus solutions of a homogeneous linear equation system and there exists a non-trivial derivation D^* of $E = F(x)$ only if the matrix formed by the $\frac{\partial p_\alpha}{\partial x_i}$:s is not full-ranked.

Let $Der_F E$ denote the set of derivations of $E = F(x)$ that are trivial on F. $Der_F E$ forms a vector space over E if we define $(bD)(x) = b(D(x))$ for $b \in E$. The dimension of this vector space can be calculated as follows, see (Jacobsson, 1980):

Theorem 3.2 *Let $E = F(x_1, \ldots, x_n)$ and let $X = \{p_1, \ldots, p_q\}$ be a finite set of generators for the ideal of polynomials p in $F[X_1, \ldots, X_n]$ such that*

$p(x_1, \ldots, x_n) = 0$ *(this set exists due to Hilbert's basis theorem). Then:*

$$[Der_F E : E] = n - \text{rank}(J(p_1, \ldots, p_q)) \quad , \tag{3.11}$$

where $J(p_1, \ldots, p_q)$ *is the Jacobian matrix*

$$\begin{bmatrix} \frac{\partial p_1}{\partial x_1} & \cdots & \frac{\partial p_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial p_q}{\partial x_1} & \cdots & \frac{\partial p_q}{\partial x_n} \end{bmatrix} \quad . \tag{3.12}$$

To see how the space $Der_F E$ is related to the transcendence degree of the field extension $F \hookrightarrow E$ suppose that $E = F(x)$ and x is algebraic over F with minimal polynomial p. If D is a derivation of E which is trivial on F, then $0 = p'(x)Dx$ and thus $Dx = 0$ since $p'(x)$ cannot be zero (the field F has characteristic zero). Therefore D is trivial on E. We have the following general result Jacobsson (1980):

Theorem 3.3 *If* $E = F(x_1, \ldots, x_n)$*, then* $Der_F E = 0$ *if and only if* E *is algebraic over* F*. Moreover,* $[Der_F E : E]$ *is equal to the transcendence degree of* E *over* F*.*

3.2.3 Rank calculation

We now have a way of calculating the transcendence degree of $E = F(x)$ over F by a rank calculation. Suppose that the transcendence degree is equal to $r > 0$ and thus some of the x_i:s are not algebraic over F. We wish to know if element x_j is algebraic over F. Consider the field extensions $F \hookrightarrow F(x_j) \hookrightarrow E$. We can calculate the transcendence degree of the field extension $F(x_j) \hookrightarrow E$ by the method described above. Since $E = F(x_j)(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$, this will involve a calculation of the rank of the following matrix:

$$\begin{bmatrix} \frac{\partial p_1}{\partial p_1} & \cdots & \frac{\partial p_1}{\partial x_{j-1}} & \frac{\partial p_1}{\partial x_{j+1}} & \cdots & \frac{\partial p_1}{\partial x_n} \\ \vdots & & \ddots & & & \vdots \\ \frac{\partial p_q}{\partial x_1} & \cdots & \frac{\partial p_q}{\partial x_{j-1}} & \frac{\partial p_q}{\partial x_{j+1}} & \cdots & \frac{\partial p_q}{\partial x_n} \end{bmatrix} \quad . \tag{3.13}$$

If the transcendence degree of the field extension $F(x_j) \hookrightarrow E$ is equal to r (i.e. the above matrix has rank $(n-1)-r$), then the variable x_j is algebraic over F. This is due to the fact that if we have the field extensions $F \hookrightarrow F' \hookrightarrow E$, then (Lang, 1993):

$$\text{tr.deg.}(E/F) = \text{tr.deg.}(E/F') + \text{tr.deg.}(F'/F) \quad . \tag{3.14}$$

We thus have a way of classifying all x_i as either algebraic over F or not by eliminating the i:th column in the Jacobian and observing if there is a change of its rank.

3.3 The observability rank condition (ORC) for rational systems

3.3.1 The ORC for polynomial systems

Setting $F = \mathbb{R}\langle U, Y\rangle$ and $E = F(x_1, \ldots, x_n)$, we can apply the theory from Subsection 3.2 to our control problem. We have obtained a method for testing the observability of polynomial control systems by calculating the transcendence degree of the field extension $\mathbb{R}\langle U, Y\rangle \hookrightarrow \mathbb{R}\langle U, Y\rangle(x)$. In order to perform the calculations described above, we need to describe the ideal I of polynomials p in $k\langle U, Y\rangle[X]$ such that $p(x_1, \ldots, x_n) = 0$. Clearly, $Y_j^{(0)} - g_j \in I$ for all $j = 1, \ldots, p$. Differentiating the j:th output variable with respect to time at zero we obtain (by Lie-differentiation where the time-dependence of the inputs is taken into account, as in Section 2.3 of the previous section):

$$Y_j^{(1)} = L_f g_j = \sum_{k=0} \sum_{i=1}^{l} \frac{\partial g_j}{\partial u_i^{(k)}} U_i^{(k+1)} \tag{3.15}$$

$$Y_j^{(2)} = L_f^2 g_j = \sum_{k=0} \sum_{i=1}^{l} \frac{\partial (L_f g_j)}{\partial u_i^{(k)}} U_i^{(k+1)} \quad , \tag{3.16}$$

etc. Clearly $Y_j^{(1)} - L_f g_j$ and $Y_j^{(2)} - L_f^2 g_j$ are elements of $\mathbb{R}\langle U, Y\rangle(x)$ and polynomials in I. In fact, all such polynomials obtained by Lie-derivation belong to I. It can be shown that I is generated by the polynomials $Y_j^{(i)} - L_f^i g_j$ for $j = 1, \ldots, p$, $i = 0, \ldots, n-1$ by the following argument of Sedoglavic's (Sedoglavic, 2002).

We have

$$\mathbb{R}\langle U\rangle \subset \mathbb{R}\langle U, Y\rangle \subset \mathbb{R}\langle U\rangle(x) \quad , \tag{3.17}$$

since each $Y_j^{(i)}$ is a polynomial function of x with coefficients in $\mathbb{R}\langle U\rangle$. Thus, as in 3.2.3,

$$\begin{aligned} &\text{tr.deg.}(\mathbb{R}\langle U\rangle(x)/\mathbb{R}\langle U\rangle) = \\ &\quad = \text{tr.deg.}(\mathbb{R}\langle U\rangle(x)/\mathbb{R}\langle U, Y\rangle) + \text{tr.deg.}(\mathbb{R}\langle U, Y\rangle/\mathbb{R}\langle U\rangle) \quad , \end{aligned} \tag{3.18}$$

and the transcendence degree of the field extension $\mathbb{R}\langle U\rangle \hookrightarrow \mathbb{R}\langle U, Y\rangle$ is therefore at most n. Thus, for every $j = 1, \ldots, p$, there exists an algebraic relation $q_j(Y_j^{(0)}, \ldots, Y_j^{(n)}) = 0$ with coefficients in $\mathbb{R}\langle U\rangle$. Thus the polynomial $Y_j^{(n)} - L_f^n g_j$ belongs to the ideal generated by the polynomials $Y_j^{(i)} - L_f^i g_j$ for $i = 1, \ldots, n-1$. We therefore conclude that we need only consider the equations obtained by the first $n-1$ Lie-derivatives of the output functions.

Hence, according to Theorem 3.2, in order to calculate the transcendence degree of the field extension $\mathbb{R}\langle U, Y\rangle \hookrightarrow \mathbb{R}\langle U, Y\rangle(x)$ we have to find the rank of the following matrix:

$$\begin{bmatrix} \frac{\partial L_f^0 g_1}{\partial x_1} & \cdots & \frac{\partial L_f^0 g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial L_f^0 g_p}{\partial x_1} & \cdots & \frac{\partial L_f^0 g_p}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial L_f^{n-1} g_1}{\partial x_1} & \cdots & \frac{\partial L_f^{n-1} g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial L_f^{n-1} g_p}{\partial x_1} & \cdots & \frac{\partial L_f^{n-1} g_p}{\partial x_n} \end{bmatrix} . \tag{3.19}$$

If this Jacobian matrix is full-ranked, then the transcendence degree is zero by Theorems 3.2 and 3.3 and we have an algebraically observable system. We have arrived at the observability rank condition that was derived for differentiable inputs in the differential geometric approach in Subsection 2.3.2, but this time we have a finite number of Lie derivatives to consider.

If the system is not algebraically observable, we can find the non-observable variables by removing columns in this matrix and calculating the rank of the reduced matrices, as described in Subsection 3.2.3.

3.3.2 The ORC for rational systems

We will now generalise this theory to apply for rational systems of type:

$$\Sigma \begin{cases} \dot{x} &= f(x,u) \\ y &= g(x,u) \end{cases} , \tag{3.20}$$

where now $f_i = p_i(u,x)/q_i(x)$ for $i = 1, \ldots, n$ and $g_j = r_j(x,u)/s_j(x)$ for $j = 1, \ldots, p$ with p_i, q_i, r_j and s_j polynomial functions.

We observe that just as before, $Y_j^{(i)} - L_f^i g_j \in \mathbb{R}\langle U, Y\rangle(x)$ for all $i = 0, \ldots, n-1$, $j = 1, \ldots, p$, but they are no longer polynomials. However, as shown by Diop et al. (1993) and Sedoglavic (2002), these rational expressions can be used in the rank test instead of the polynomials that generate the ideal I, and therefore we may use the same Jacobian in this case, as for polynomial systems.

Remark: Observe that the algebraic interpretation has lead us to the observability rank condition derived for analytic inputs in Subsection 2.3 of the previous section, showing the equivalence of algebraic observability and local distinguishability, see (Diop et al., 1993). In fact, the ideal I of

polynomials p in $\mathbb{R}\langle U, Y\rangle[X]$ such that $p(x_1, \ldots, x_n) = 0$ is generated by the same functions that span the space $\mathcal{F}^{\mathcal{K}}$ defined in Section 2. The rank of the Jacobian

$$\begin{bmatrix} \frac{\partial L_f^0 g_1}{\partial x_1} & \cdots & \frac{\partial L_f^0 g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial L_f^0 g_p}{\partial x_1} & \cdots & \frac{\partial L_f^0 g_p}{\partial x_n} \\ \cdots & & \\ \frac{\partial L_f^{n-1} g_1}{\partial x_1} & \cdots & \frac{\partial L_f^{n-1} g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial L_f^{n-1} g_m}{\partial x_1} & \cdots & \frac{\partial L_f^{n-1} g_m}{\partial x_n} \end{bmatrix} \tag{3.21}$$

is exactly the dimension of the space $\mathcal{O}^{\mathcal{K}}$ which, as we recall, determines the local distingushability property according to Theorem 2.2. The result of the algebraic approach of this section is that we have been able to show that for rational systems the space $\mathcal{O}^{\mathcal{K}}$ is generated by a finite number of functions. In Section 4, we take a different approach to show that this is in fact true for all analytical systems of the form (2.1).

3.4 Symmetry

Suppose now that by applying the rank test above, we find that our control system is not algebraically observable and that the transcendence degree is r. This means that $Der_{\mathbb{R}\langle U,Y\rangle}\mathbb{R}\langle U, Y\rangle(x)$ is not empty and has dimension r. The differential-geometric concept that corresponds to derivations is that of tangent vectors. We can therefore interpret the existence of derivations on $\mathbb{R}\langle U, Y\rangle(x)$ that are trivial on $\mathbb{R}\langle U, Y\rangle$ as the existence of tangent vectors to the space of solutions to our control system, such that if we move in their direction, the output remains the same and we cannot observe that the system is in a different state. In other words, there are infinitely many trajectories for the control system that cannot be distinguished from each other by observing the input-output map.

A derivation therefore generates a family of symmetries for the control system - symmetries in the variables leaving the inputs and outputs invariable. In this section we will show how these can be calculated.

First of all, observe that the partial derivatives $\frac{\partial}{\partial x_i}$ form a basis for the derivations on $\mathbb{R}\langle U\rangle(x)$ that are trivial on $\mathbb{R}\langle U\rangle$ (see Appendix 8.2 for explanation). Among these, we wish to find the ones that are trivial also on $\mathbb{R}\langle U, Y\rangle$. If v is one of them, recall from Theorems 3.1 and 3.2 that we must

have:

$$\begin{bmatrix} \frac{\partial L_f^0 g_1}{\partial x_1} & \cdots & \frac{\partial L_f^0 g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial L_f^0 g_p}{\partial x_1} & \cdots & \frac{\partial L_f^0 g_p}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial L_f^{n-1} g_1}{\partial x_1} & \cdots & \frac{\partial L_f^{n-1} g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial L_f^{n-1} g_p}{\partial x_1} & \cdots & \frac{\partial L_f^{n-1} g_p}{\partial x_n} \end{bmatrix} \cdot v = 0 \quad . \tag{3.22}$$

Thus, v belongs to the kernel of the above Jacobian matrix. Suppose that $v = (v_1, \ldots, v_n)$, where $v_i \in \mathbb{R}\langle U, Y\rangle(x)$. Then v is the Lie-derivation $v = \sum_{i=1}^{n} v_i \frac{\partial}{\partial x_i}$ which corresponds to a vector field v and a flow $\Phi(\rho, x)$ of v given by (see Section 2):

$$\begin{cases} \frac{\partial}{\partial \rho}\Phi(\rho, x) &= v(\Phi(\rho, x)) \\ \Phi(0, x) &= x \quad . \end{cases} \tag{3.23}$$

The solution of this system of differential equations evaluated at any $\rho > 0$ corresponds to a new initial state of the system which cannot be distinguished from the original one, $(x_1, \ldots, x_n)$, by observing the input-output it produces.

We now have a strategy for finding the families of symmetries for our control system. First, we have to define a basis for the kernel of the Jacobian matrix. In order to obtain the associated families of symmetries, we have to solve the system of differential equations that correspond to each element of the chosen basis. To make the calculations simpler, we can use the observations from Subsection 3.2.3 to find the non-observable variables. Instead of calculating the kernel of the Jacobian matrix, we can calculate the kernel of its maximal singular minor which is obtained when the columns and rows corresponding to the observable variables are removed. Then, the system of differential equations to be solved will only involve the non-observable variables.

We will now apply this to a non-observable example:

$$\begin{cases} \dot{x}_1 &= x_2 x_4 + u \\ \dot{x}_2 &= x_2 x_3 \\ \dot{x}_3 &= 0 \\ \dot{x}_4 &= 0 \\ y &= x_1 \quad . \end{cases} \tag{3.24}$$

We need to calculate the first three Lie-derivatives of the output function:

$$\begin{aligned} Y^{(1)} &= L_f x_1 = x_2 x_4 + U^{(0)} && (3.25)\\ Y^{(2)} &= L_f^2 x_1 = L_f(x_2 x_4 + u) = x_4 x_2 x_3 + U^{(1)} && (3.26)\\ Y^{(3)} &= L_f^3 x_1 = L_f(x_4 x_2 x_3 + \dot{u}) = x_4 x_3 x_2 x_3 + U^{(2)} \quad . && (3.27) \end{aligned}$$

Thus the Jacobian matrix becomes:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & x_4 & 0 & x_2 \\ 0 & x_3 x_4 & x_2 x_4 & x_2 x_3 \\ 0 & x_3^2 x_4 & 2x_2 x_3 x_4 & x_2 x_3^2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & x_4 & 0 & x_2 \\ 0 & 0 & x_2 x_4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad . \tag{3.28}$$

Clearly, this matrix has rank 3 and the non-observable variables are x_2 and x_4 - removing the second or fourth column does not change the rank of the matrix. We can now eliminate the first and third rows and columns and consider the kernel of the remaining minor, which is the matrix

$$\begin{bmatrix} x_4 & x_2 \\ x_3^2 x_4 & x_2 x_3^2 \end{bmatrix} \quad . \tag{3.29}$$

This kernel is generated by the vector $(x_2, -x_4)$. The derivation $x_2 \frac{\partial}{\partial x_2} - x_4 \frac{\partial}{\partial x_4}$ thus corresponds to the system of differential equations:

$$\begin{cases} \dot{\Phi}_2(\rho, x) &= \Phi_2(\rho, x) \\ \dot{\Phi}_4(\rho, x) &= -\Phi_4(\rho, x) \\ \Phi_2(0, x) &= x_2 \\ \Phi_4(0, x) &= x_4 \quad . \end{cases} \tag{3.30}$$

The solution is:

$$\begin{cases} \Phi_2(\rho, x) &= x_2 e^{\rho} \\ \Phi_4(\rho, x) &= x_4 e^{-\rho} \quad . \end{cases} \tag{3.31}$$

If we set $e^{\rho} = \lambda$, we find that multiplying x_2 by λ and dividing x_4 by it defines a new state $\bar{x}$ that is indistinguishable from the original one for any λ. Indeed, we see that performing this procedure does not change the output

and its Lie-derivatives:

$$\begin{cases} \dot{\bar{x}}_1 &= \bar{x}_2\bar{x}_4 + u = \lambda x_2 x_4/\lambda + u = x_2x_4 + u \\ \dot{\bar{x}}_2 &= \bar{x}_2\bar{x}_3 = \bar{x}_2\bar{x}_3 = \lambda x_2 x_3 \\ \dot{\bar{x}}_3 &= 0 \\ \dot{\bar{x}}_4 &= 0 \\ \bar{Y}^{(0)} &= \bar{x}_1 = x_1 = Y^{(0)} \\ \bar{Y}^{(1)} &= L_{\bar{f}}\bar{x}_1 = \bar{x}_2\bar{x}_4 + U^{(0)} = x_2x_4 + U^{(0)} = Y^{(1)} \\ \bar{Y}^{(2)} &= L^2_{\bar{f}}\bar{x}_1 = L_{\bar{f}}(\bar{x}_2\bar{x}_4 + u) = \bar{x}_4\dot{\bar{x}}_2 + U^{(1)} = \bar{x}_4\bar{x}_2\bar{x}_3 + U^{(1)} = \\ &= \tfrac{1}{\lambda}x_4\lambda x_2x_3 + U^{(1)} = Y^{(2)} \\ \bar{Y}^{(3)} &= L^3_{\bar{f}}\bar{x}_1 = L_{\bar{f}}(\bar{x}_4\bar{x}_2\bar{x}_3 + \dot{u}) = \bar{x}_3\bar{x}_4\dot{\bar{x}}_2 + U^{(2)} = \\ &= x_3\tfrac{1}{\lambda}x_4\lambda x_2x_3 + U^{(2)} = Y^{(3)} \quad . \end{cases} \tag{3.32}$$

We know from Subsection 3.3.1 that we need not consider any further Lie derivatives since they depend on the previous ones.

We have now defined a family of symmetries

$$\sigma_\lambda : \{x_1, x_2, x_3, x_4\} \to \{x_1, \lambda x_2, x_3, x_4/\lambda\} \tag{3.33}$$

of the control system which leaves the input and output invariant.

3.5 Sedoglavic's algorithm

There is a published algorithm by Sedoglavic (2002) with a Maple implementation which performs an observability test of rational systems and for non-identifiable systems, predicts the non-identifiable variables with high probability. This is done in polynomial time with respect to system complexity.

The algorithm is mainly based on generic rank computation, for details, see (Sedoglavic, 2002). The symbolic computation of the Jacobian matrix defined in Subsection 3.3 can be cumbersome for systems with many variables and parameters and it cannot be done in polynomial time. Instead, the parameters are specialised on some random integer values, and the inputs are specialised on a power series of t with integer coefficients. To limit the growth of these integers in the process of rank computation, the calculations are done on a finite field $\mathbb{F}_p$ (p refers to a prime number). The probabilistic aspects of the algorithm concern the choice of specialisation of parameters and inputs and also the fact that cancelation of the determinant of the Jacobian modulo p has to be avoided. The calculation of the rank is deterministic for observable systems, that is, when the process states that the system is observable, the answer is correct. For non-observable systems, the probability of a correct answer depends on the complexity of the system

and on the prime number p. The predicted non-observable variables can be further analysed to find a family of symmetries which then can confirm the test result.

The Maple implementation takes as an input a rational system of differential equations where parameters, state-variables and inputs have to be stated as such, and also a set of outputs has to be defined. The transcendence degree of the field extension associated to the system is calculated and the non-observable parameters and state-variables are predicted.

We have used this implementation for our case study in Section 6.

4 The first $n-1$ derivatives of the output function determine the observability of analytic systems with n state variables

This section deals with several questions that arise from Sections 2 and 3. The differential-geometric approach from Section 2 results in the observability rank test for observability of analytic systems. In this test, the rank of the linear space containing the gradients of all Lie derivatives of the output functions must be calculated. Since no bound is given for the number of Lie derivatives necessary for the calculation, the practical application of the test to other than the simplest examples is difficult. Such an upper bound is derived for the case of rational systems in Section 3 using the algebraical approach. The following questions now arise. Can an upper bound be given only for rational systems? How do such requirements for the class of the system arise? In this section, we attempt to extend the upper bound for the number of time-derivatives of the output function to apply for the class of analytical systems affine in the input variable that are addressed by the differential-geometric approach in Section 2. We are going to use the results by Sontag (1991) described in Subsection 2.3 where the observability rank condition was defined in terms of differentiable inputs.

Consider once again the example from the introduction, taken from Sedoglavic (2002):

$$\left\{\begin{array}{lcl} \dot{x}_1 &=& \frac{x_2}{x_1} \\ \dot{x}_2 &=& \frac{x_3}{x_2} \\ \dot{x}_3 &=& x_1\theta - u \\ y &=& x_1 \quad . \end{array}\right. \tag{4.1}$$

Recall that we obtained the following equations for the state-variables and the parameter from calculating the first three time-derivatives at zero of the output (see Subsection 1.2):

$$\begin{array}{lcl} r_1(x_1, Y^{(0)}) = Y^{(0)} - x_1 &=& 0 \\ r_2(x_1, x_2, Y^{(1)}) = Y^{(1)} - \frac{x_2}{x_1} &=& 0 \\ r_3(x_1, x_2, x_3, Y^{(2)}) = Y^{(2)} - (\frac{x_3}{x_1 x_2} - \frac{x_2^2}{x_1^3}) &=& 0 \\ r_4(x_1, x_2, x_3, \theta, Y^{(3)}) = Y^{(3)} - (\frac{\theta}{x_2} - \frac{U^{(0)}}{x_1 x_2} - \frac{3x_3}{x_1^3} - \frac{x_3^2}{x_1 x_2^3} + \frac{3x_2^3}{x_1^5}) &=& 0 \quad . \end{array} \tag{4.2}$$

The problem now is to determine whether these equations are enough to ensure that a given input-output behaviour corresponds to a locally unique

state of the system. From the implicit function theorem it follows that the variables x_1, x_2, x_3 and the parameter θ can be expressed locally (in the neighbourhood of a given point in the space of solutions of the differential equations) as functions of $U^{(0)}$ and $Y^{(0)}, Y^{(1)}, Y^{(2)}, Y^{(3)}$ if the rank of the following Jacobian matrix evaluated at that point is equal to four:

$$\begin{bmatrix} \frac{\partial(r_1)}{\partial x_1} & \frac{\partial(r_1)}{\partial x_2} & \frac{\partial(r_1)}{\partial x_3} & \frac{\partial(r_1)}{\partial \theta} \\ \frac{\partial(r_2)}{\partial x_1} & \frac{\partial(r_2)}{\partial x_2} & \frac{\partial(r_2)}{\partial x_3} & \frac{\partial(r_2)}{\partial \theta} \\ \frac{\partial(r_3)}{\partial x_1} & \frac{\partial(r_3)}{\partial x_2} & \frac{\partial(r_3)}{\partial x_3} & \frac{\partial(r_3)}{\partial \theta} \\ \frac{\partial(r_4)}{\partial x_1} & \frac{\partial(r_4)}{\partial x_2} & \frac{\partial(r_4)}{\partial x_3} & \frac{\partial(r_4)}{\partial \theta} \end{bmatrix} = \tag{4.3}$$

$$= - \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{x_2}{x_1^2} & \frac{1}{x_1} & 0 & 0 \\ -\frac{x_3}{x_1^2 x_2} + \frac{3x_2^2}{x_1^4} & -\frac{x_3}{x_1 x_2^2} + \frac{2x_2}{x_1^3} & \frac{1}{x_1^2 x_2} & 0 \\ \frac{u}{x_2 x_1^2} + \frac{9x_3}{x_1^4} + \frac{x_3^2}{x_2^3 x_1^2} - \frac{15x_2^3}{x_1^6} & -\frac{\theta}{x_2^2} + \frac{u}{x_1 x_2^2} + \frac{3x_3^2}{x_1 x_2^4} + \frac{9x_2^2}{x_1^5} & -\frac{3}{x_1^3} - \frac{2x_3}{x_1 x_2^3} & \frac{1}{x_2} \end{bmatrix}.$$

Clearly, this matrix has full rank for all values of x_1, x_2, x_3 and θ and thus the system has a locally unique state for a given input-output behaviour.

Now the following question arises - if the rank of the above matrix is not full, can we then conclude that the system is not locally observable without considering further derivatives of the output function which would produce new equations? In other words, is the rank of the Jacobian determined by the first n equations, where n is the total number of state-variables and parameters? We will now show that this is true for the analytic systems affine in the input variable that were discussed in Section 2.

Consider again the analytic control system of the form (equation (2.1)):

$$\Sigma \begin{cases} \dot{x} &= f(x,u) = g^0(x) + g(x)u \\ y &= h(x) \quad . \end{cases} \tag{4.4}$$

As previously (Section 2), the elements of the n-dimensional vectors g^0 and g are analytic functions of x and we assume for the moment that we have a single analytic output $h(x)$ and also a single analytic input u. The n state-variables x are assumed to occupy an open subset M of $\mathbb{R}^n$.

The first two equations obtained by differentiating the output function

with respect to time at zero are:

$$\begin{aligned}
Y^{(1)} &= L_f h(x) = dh \cdot f_{|t=0} = dh \cdot (g^0 + gU^{(0)}) = \\
&= dh \cdot g^0 + U^{(0)}(dh \cdot g) \qquad (4.5)\\
Y^{(2)} &= L_f^2 h(x) = L_f(dh \cdot f) = (d(dh \cdot f) \cdot f)_{|t=0} + \frac{\partial (dh \cdot f)}{\partial u} U^{(1)} = \\
&= d\big(dh \cdot g^0 + U^{(0)}(dh \cdot g)\big) \cdot \big(g^0 + gU^{(0)}\big) + \\
&\quad + \frac{\partial \big(dh \cdot g^0 + u(dh \cdot g)\big)}{\partial u} U^{(1)} = \\
&= d\big(dh \cdot g^0 + U^{(0)}(dh \cdot g)\big) \cdot h^0 + U^{(0)} d\big(dh \cdot g^0 + U^{(0)}(dh \cdot g)\big) \cdot g + \\
&\quad + U^{(1)}(dh \cdot g) = \\
&= d(dh \cdot g^0) \cdot g^0 + U^{(0)}\big(d(dh \cdot g) \cdot g^0 + d(dh \cdot g^0) \cdot g\big) + \\
&\quad + (U^{(0)})^2 \big(d(dh \cdot g)\big) \cdot g) + U^{(1)}(dh \cdot g) \quad . \qquad (4.6)
\end{aligned}$$

These calculations confirm the result by Sontag (1991) that the first $n-1$ Lie derivatives of the output function $g(x)$ for the system (2.1) are polynomial functions of $U^{(0)}, U^{(1)}, \ldots, U^{(n-2)}$ with coefficients that are analytic functions on M.

Thus we have that $L_f^{(i)} h \in \mathcal{K}_x$ for $i = 0, \ldots, n-1$; recall from Subsection 2.3 that $\mathcal{K}_x = \mathbb{R}_x(U^{(0)}, U^{(1)}, \ldots)$ is the field of meromorphic functions on M to which we add the indeterminates $U^{(0)}, U^{(1)}, \ldots$ and obtain rational functions of $U^{(0)}, U^{(1)}, \ldots$ with coefficients that are meromorphic functions on M. See also (Sontag, 1991).

Following the notation from example (4.1) above, the first n equations for the state-variables can now be formulated:

$$\begin{aligned}
r_1(x, Y^{(0)}) &= Y^{(0)} - h = 0 \qquad (4.7)\\
r_2(x, u, Y^{(1)}) &= Y^{(1)} - L_f h = 0 \qquad (4.8)\\
&\vdots \\
r_n(x, u, \ldots, u^{(n-2)}, Y^{(n-1)}) &= Y^{(n-1)} - L_f^{n-1} h = 0 \quad . \qquad (4.9)
\end{aligned}$$

Therefore, the Jacobian that we are interested in is:

$$-\begin{bmatrix} \frac{\partial h}{\partial x_1} & \cdots & \frac{\partial h}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial L_f^{n-1} h}{\partial x_1} & \cdots & \frac{\partial L_f^{n-1} h}{\partial x_n} \end{bmatrix} \quad . \qquad (4.10)$$

Since $L_f^{(i)} h \in \mathcal{K}_x$ for $i = 0, \ldots, n-1$, the elements of this Jacobian also belong to $\mathcal{K}_x$. We will now show that if this Jacobian is not full-ranked,

that is, the first n gradients of the output function and its Lie derivatives are linearly dependent over the field $\mathcal{K}_x$, then any further Lie derivative produces a gradient which is linearly dependent of the first n and we can thus conclude that the system is not locally observable. Furthermore, if the first q gradients, where $q \leq n$, are linearly-dependent, then no further gradients are necessary for the calculation of the rank, which becomes $\leq q-1$. In fact, we can stop Lie differentiating the output function at the first instance of linear dependence.

Remark: To be able to discuss linear dependence, we have to know that the gradients of the Lie derivatives produce a linear space over a field (or a free module over a commutative ring). This was the case for the rational systems in Section 3 and this is also the case here for analytic systems of the above type, because the elements of the Jacobian belong to the field $\mathcal{K}_x$.

Theorem 4.1 *Let Σ be the system*

$$\begin{cases} \dot{x} &= f(x,u) = g^0(x) + g(x)u \\ y &= h(x) \quad , \end{cases} \tag{4.11}$$

where x is a vector of n state-variables occupying an open subset M of $\mathbb{R}^n$, g^0 and g are n-dimensional vectors of analytic functions on M, the output $h(x)$ is an analytic function on M and the control variable u is an analytic function of time.

If q is an integer such that $dL_f^{(i)}h$, $i = 0, \ldots, q$ are linearly dependent over the field $\mathcal{K}_x$, then the dimension of the space $\mathcal{O}^{\mathcal{K}} = \text{span}_{\mathcal{K}_x}\{dL_f^{(i)}h, \quad i \geq 0\}$ (see Subsection 2.3) is less than or equal to $q-1$. If $q < n$, the system Σ is not locally observable.

Proof: Suppose that the first q gradients are linearly dependent and q is the least such number (it certainly exits as the rank is $\leq n$ and a single non-zero vector is linearly independent of itself). Then, there exist coefficients $k_i \in \mathcal{K}_x, i = 0, \ldots, q-1$, not all of them zero, such that

$$\sum_{i=0}^{q-1} k_i dL_f^i h = \mathbf{0} \quad . \tag{4.12}$$

We can take the Lie derivative of both sides (which are co-vector fields) to obtain:

$$\mathbf{0} = L_f(\sum_{i=0}^{q-1} k_i dL_f^i h) = \sum_{i=0}^{q-1} L_f(k_i dL_f^i h) = \sum_{i=0}^{q-1} \left((L_f k_i) dL_f^i h + k_i L_f(dL_f^i h)\right) \quad . \tag{4.13}$$

We now observe the following fact (which is simply saying that the d and L_f operators commute even when f depends on a control variable $u(t)$, see Appendix 8.1 for derivation):

$$L_f(dL_f^i h) = dL_f^{i+1} h \quad , \tag{4.14}$$

for $i \geq 0$.

It follows that

$$\mathbf{0} = \sum_{i=0}^{q-1} \left((L_f k_i) dL_f^i h + k_i dL_f^{i+1} h \right) \quad . \tag{4.15}$$

Recalling the structure of the field $\mathcal{K}_x$, we know that $L_f k_i \in \mathcal{K}_x$ since

$$L_f k_i = dk_i \cdot f + \frac{\partial k_i}{\partial u} U^{(1)} = dk_i \cdot (g^0 + gU^{(0)}) + \frac{\partial k_i}{\partial u} U^{(1)} \quad , \tag{4.16}$$

which is clearly a rational function of $U^{(0)}, U^{(1)}, \ldots$ with coefficients that are meromorphic functions of x. Since we know that k_{q-1} is not zero (we assumed that q was the least number such that the first q gradients are linearly dependent), we conclude that $dL_f^q h$ is linearly dependent on the preceding gradients.

Using the same calculations we can prove by induction that any further gradient is linearly dependent on the previous ones which then means that $dL_f^0 h, \ldots, dL_f^{q-1} h$ form a basis for the space $\mathcal{O}^{\mathcal{K}}$ which determines local distinguishability (by Theorems 2.1 and 2.2) and thus local observability. If $q < n$, this space has rank less than n and thus the system is not locally observable. □

Thus it is enough to consider the first $n-1$ Lie derivatives of the output function in the rank test and also, we can stop calculating further derivatives of the output function at the first instance of linear dependence among their gradients.

Remark: We note that in the case of multiple output functions one needs to calculate $n-1$ time-derivatives of each.

5 Parameter identifiability

In this relatively short section we will present the problem of parameter identifiability of nonlinear control systems as a special case of the observability problem.

Identifiability is the possibility to identify the parameters of a control system from its input-output behaviour. By considering parameters as state-variables with time derivative zero, one can use the observability rank test to determine identifiability. The property of local observability is then interpreted as the existence of only finitely many parameter sets that fit the observed data, each of them locally unique. The use of the rank test for determining the identifiability of nonlinear systems dates back to at least 1978 when Pohjanpalo (1978) used the coefficients of the Taylor series expansion of the output to determine the parameter identifiability of a class of nonlinear systems applied in the analysis of saturation phenomena in pharmacokinetic studies. A more recent example is the work by Xia and Moog (2003) where different concepts of nonlinear identifiability are studied in an algebraic framework. They apply the theory to a four-dimensional HIV/AIDS model, and show that their theoretical results can be used to determine whether all the parameters in the model are determinable from the measurement of CD4+ T cells and virus load, and if not, what else has to be measured. The minimal number of measurements of the variables for the complete determination of all parameters and the best period of time to make such measurements are calculated. Another example with biological application is the work by Margaria et al. (2004) where the identifiability of some highly structured biological models of infectious disease dynamics is analysed both using the rank method and Sedoglavic's algorithm, (Sedoglavic, 2002) and also by the constructive method of characteristic set computation described by Ollivier (1990); Ljung and Glad (1990) and others. Due to the fact that its computational complexity is exponential in the number of parameters, the latter method can only be applied to relatively small control systems.

We will now describe how the observability rank test can be used to determine parameter identifiability. Consider a physical/chemical/biological model:

$$\Sigma \begin{cases} \dot{x} &= f(x,p,u) \\ y &= h(x,p) \end{cases}, \tag{5.1}$$

where as before, x denote the n state-variables, u the m inputs and y the p observed quantities. The l model parameters are denoted by p and $f(x,p,u)$ and $h(x,p)$ are vectors of analytical functions. We may or may not be given

a set of initial conditions for the state-variables:

$$x(0) = x^0 \quad . \tag{5.2}$$

In order to be able to use the theory from the previous sections, we observe that the above model can be represented by the following control system:

$$\Sigma \begin{cases} \dot{p} &= 0 \\ \dot{x} &= f(x,p,u) \\ y &= h(x,p) \quad , \end{cases} \tag{5.3}$$

where x and p can now be considered as the same type of variables. We can apply the rank test to this system in exactly the same way as discussed in the previous sections.

Without initial conditions for x, the non-observable variables can be both in x and in p. Suppose now that we are given a full set of initial conditions on x:

$$x(0) = x^0 \quad . \tag{5.4}$$

The problem of the observability of the x variables now disappears as the initial state is already uniquely defined. What is left, is exactly the problem of identifiability for the parameters - is the set of parameters that realises a given input-output map unique, at least locally?

This can be determined by the rank test described in the previous chapters. For analytical systems (see Section 4) the rank test amounts to calculating the rank of the following Jacobian matrix:

$$\begin{bmatrix} \frac{\partial L_f^0 g_1}{\partial p_1} & \cdots & \frac{\partial L_f^0 g_1}{\partial p_l} \\ \vdots & \ddots & \vdots \\ \frac{\partial L_f^0 g_m}{\partial p_1} & \cdots & \frac{\partial L_f^0 g_m}{\partial p_l} \\ \vdots & \ddots & \vdots \\ \frac{\partial L_f^{n-1} g_1}{\partial p_1} & \cdots & \frac{\partial L_f^{n-1} g_1}{\partial p_l} \\ \vdots & \ddots & \vdots \\ \frac{\partial L_f^{n-1} g_m}{\partial p_1} & \cdots & \frac{\partial L_f^{n-1} g_m}{\partial p_l} \end{bmatrix} \quad . \tag{5.5}$$

If the rank of this matrix is l, then the model is identifiable. If not, the non-identifiable parameters can be found using the same procedure we used earlier for finding non-observable variables, see Subsection 3.2.3.

6 Case study: Identifiability analysis of a kinetic model for *S. cerevisiae*

In this case study, we have investigated the identifiability of a published model of the metabolic dynamics in *S. cerevisiae* by Rizzi et al. (1997). We begin by a short description of the biochemistry of the central metabolic pathways.

6.1 The central metabolic pathways

Metabolism is the overall network of enzyme-catalysed reactions in a cell. Its degradative, or energy-releasing phase is called catabolism. The central catabolic pathways which are more or less universal among organisms consist of glycolysis, the pentose phosphate pathway and the citric acid cycle. In glycolysis sugars are degraded to a three-carbon compound called pyruvate. In the absence of oxygen pyruvate is then reduced to lactate, ethanol or other fermentation products. In aerobic conditions, it is instead oxidised via the citric acid cycle in the process of cellular respiration. A simplified scheme of some of the most important reactions in the central metabolic pathways is shown in the figure on the next page.

The different species in the boxes are called metabolites. The reactions marked by arrows are catalysed by enzymes which determine their "reaction rate" or "flux", that is, the speed with which the reaction occurs.

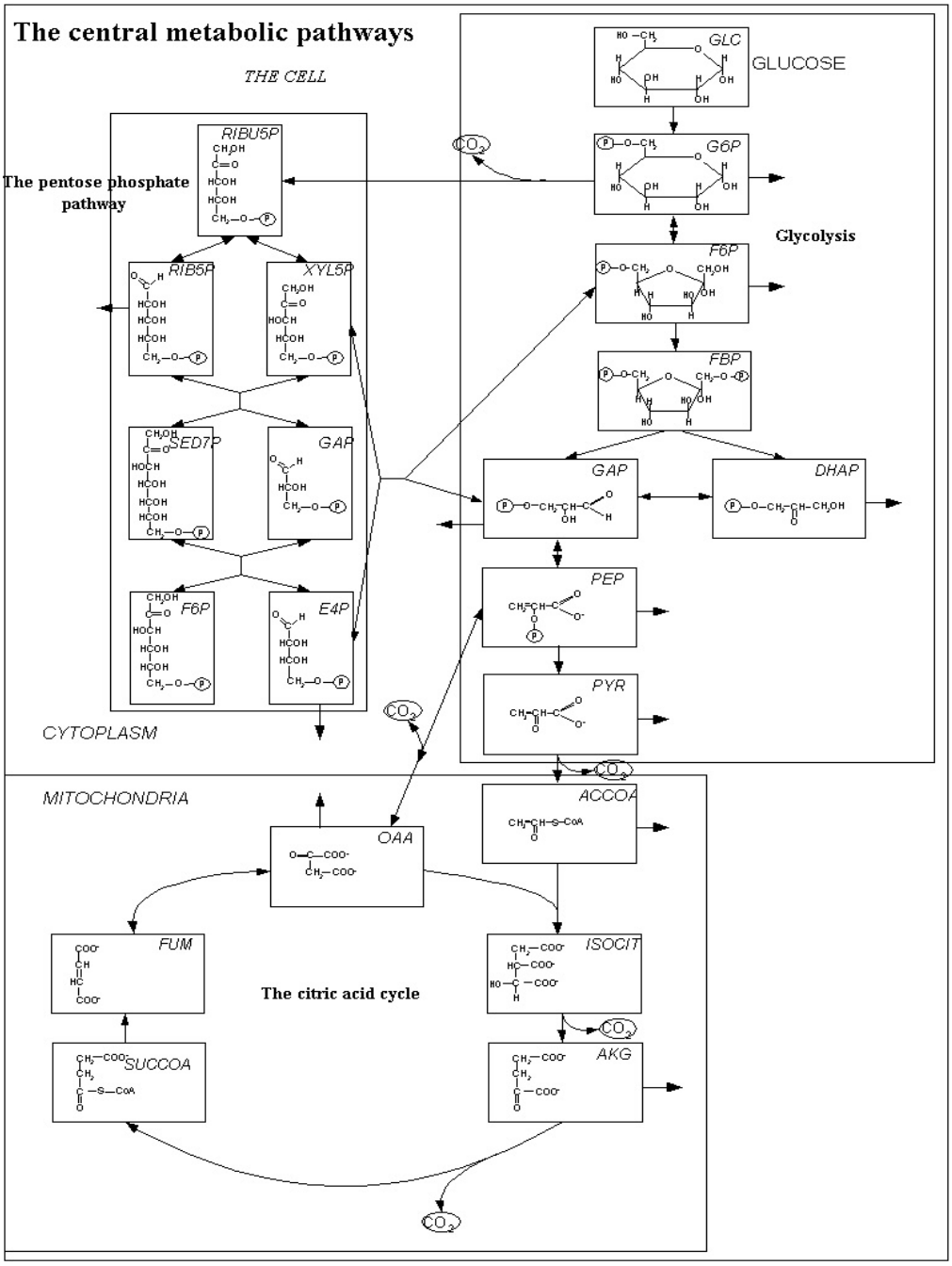
Reaction rates are often modelled by using so-called Michaelis-Menten or Hill kinetics where an equation is derived for the reaction rate based on a biochemical description of the general way in which enzymatic reactions occur. For example, a reaction in which a single substrate (reactant) A is transformed to a single product B under the catalysis of a single enzyme E has the following rate equation, (Lehninger, 2000):

$$r = \frac{r^{\max} A}{K_{\mathrm{m}} + A} \quad , \tag{6.1}$$

where the constants $r^{\max}$ and K_{m} are specific for this reaction. $r^{\max}$ is the maximal rate of the reaction and K_{m} is the substrate concentration at which the reaction rate is half $r^{\max}$.

An enzyme can have several binding sites for the substrate in which case a so-called Hill equation is used. For the above reaction where we allow n binding sites for the enzyme, the equation becomes (see for example Chapter 5 in (Stephanopoulos et al., 1998)):

$$r = \frac{r^{\max} A^n}{K_{\mathrm{m}} + A^n} \quad . \tag{6.2}$$

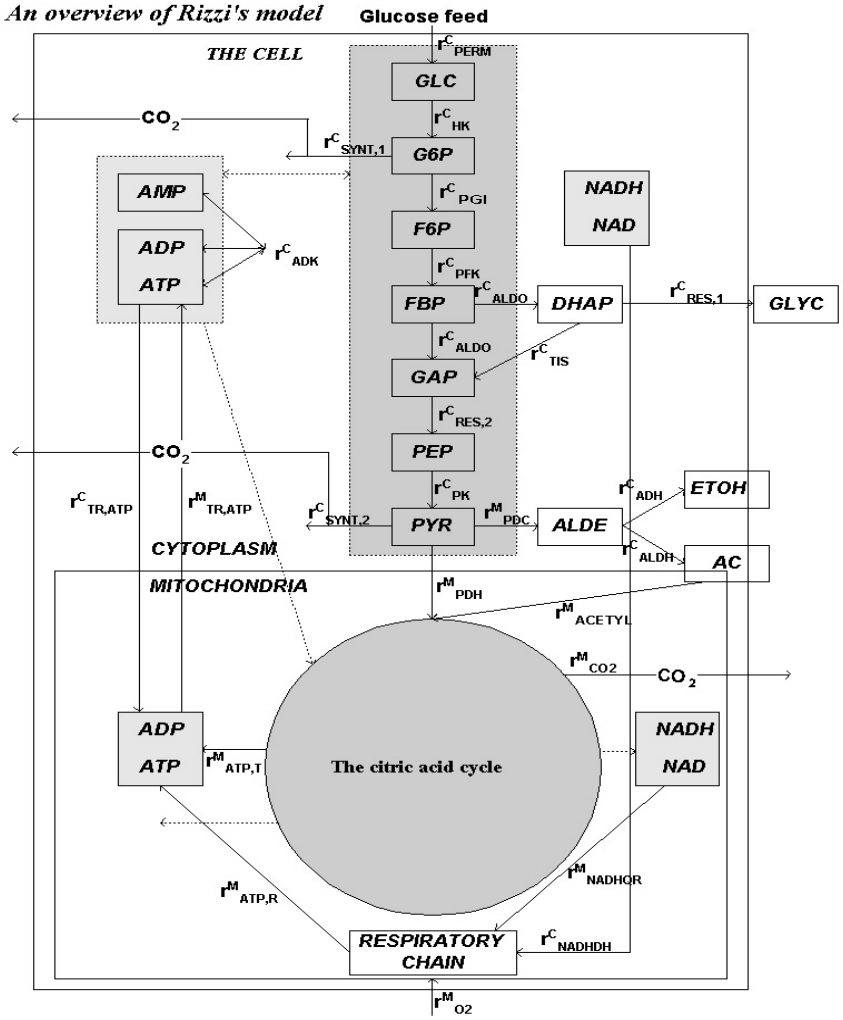


These equations can take much more complicated forms depending on the number of substrates and products and other factors such as reversibility of the reaction, inhibition and cooperation effects on the enzymes, etc., see (Lehninger, 2000; Segel, 1975) and (Rizzi et al., 1997) for details.

6.2 The model of metabolic dynamics by Rizzi et al.

We now proceed to describe a mathematical model of the dynamics of the chemical reactions in the central metabolic pathways formulated by Rizzi et al. (1997).

The authors propose a kinetic model for the reactions of glycolysis, the citric acid cycle, the glyoxylate cycle and the respiratory chain in growing cells of *S. cerevisiae*. The model aims to predict the short-term changes in the metabolic states of the cells under *in vivo* conditions after a change in the glucose feed rate. A schematic picture adapted from (Rizzi et al., 1997) describing the metabolites and fluxes included in the model is shown below.



For each metabolite in the scheme, a mass balance is written where the change of its concentration in time is expressed accounting for the incoming and outcoming fluxes as well as the effect of dilution. The following system of differential equations is obtained for the concentrations of the different species (see Appendix 8.3 for nomenclature and parameter description):

$$\frac{dc^e_{\mathrm{GLC}}}{dt} = D(c^0_{\mathrm{GLC}} - c^e_{\mathrm{GLC}}) - \frac{c_{\mathrm{X}}}{\rho} r^C_{\mathrm{PERM}} \tag{6.3}$$

$$\frac{dc^e_{\mathrm{GLYC}}}{dt} = \frac{c_{\mathrm{X}}}{\rho} r^C_{\mathrm{RES,1}} - D c^e_{\mathrm{GLYC}} \tag{6.4}$$

$$\frac{dc^e_{\mathrm{AC}}}{dt} = \frac{c_{\mathrm{X}}}{\rho} r^C_{\mathrm{ALDH}} - D c^e_{\mathrm{AC}} \tag{6.5}$$

$$\frac{dc^e_{\mathrm{ETOH}}}{dt} = \frac{c_{\mathrm{X}}}{\rho} r^C_{\mathrm{ADH}} - D c^e_{\mathrm{ETOH}} \tag{6.6}$$

$$\frac{dc^e_{\mathrm{CO_2}}}{dt} = \frac{c_{\mathrm{X}}}{\rho}\left(\frac{V_M}{V_C} r^M_{\mathrm{CO_2}} + r^C_{\mathrm{PDC}} + a_{\mathrm{CO_2},1} r^C_{\mathrm{SYNT,1}} + a_{\mathrm{CO_2},2} r^C_{\mathrm{SYNT,2}}\right) + + S_{\mathrm{CO_2}} \tag{6.7}$$

$$\frac{dc^e_{\mathrm{O_2}}}{dt} = -\frac{V_M}{V_C}\frac{C_{\mathrm{X}}}{\rho} r^M_{\mathrm{O_2}} + S_{\mathrm{O_2}} \tag{6.8}$$

$$\frac{dc^C_{\mathrm{GLC}}}{dt} = r^C_{\mathrm{PERM}} - r^C_{\mathrm{HK}} - \mu c^C_{\mathrm{GLC}} \tag{6.9}$$

$$\frac{dc^C_{\mathrm{G6P}}}{dt} = r^C_{\mathrm{HK}} - r^C_{\mathrm{PGI}} - r^C_{\mathrm{SYNT,1}} - \mu c^C_{\mathrm{G6P}} \tag{6.10}$$

$$\frac{dc^C_{\mathrm{F6P}}}{dt} = r^C_{\mathrm{PGI}} - r^C_{\mathrm{PFK}} - \mu c^C_{\mathrm{F6P}} \tag{6.11}$$

$$\frac{dc^C_{\mathrm{FBP}}}{dt} = r^C_{\mathrm{PFK}} - r^C_{\mathrm{ALDO}} - \mu c^C_{\mathrm{FBP}} \tag{6.12}$$

$$\frac{dc^C_{\mathrm{DHAP}}}{dt} = r^C_{\mathrm{ALDO}} - r^C_{\mathrm{TIS}} - r^C_{\mathrm{RES,1}} - \mu c^C_{\mathrm{DHAP}} \tag{6.13}$$

$$\frac{dc^C_{\mathrm{GAP}}}{dt} = r^C_{\mathrm{ALDO}} + r^C_{\mathrm{TIS}} - r^C_{\mathrm{RES,2}} - \mu c^C_{\mathrm{GAP}} \tag{6.14}$$

$$\frac{dc^C_{\mathrm{PEP}}}{dt} = r^C_{\mathrm{RES,2}} - r^C_{\mathrm{PK}} - \mu c^C_{\mathrm{PEP}} \tag{6.15}$$

$$\frac{dc^C_{\mathrm{PYR}}}{dt} = r^C_{\mathrm{PK}} - \frac{V_M}{V_C} r^M_{\mathrm{PDH}} - r^C_{\mathrm{PDC}} - r^C_{\mathrm{SYNT,2}} - \mu c^C_{\mathrm{PYR}} \tag{6.16}$$

$$\frac{dc^C_{\mathrm{ALDE}}}{dt} = r^C_{\mathrm{PDC}} - r^C_{\mathrm{ADH}} - r^C_{\mathrm{ALDH}} - \frac{V_M}{V_C} r^M_{\mathrm{ACETYL}} - \mu c^C_{\mathrm{ALDE}} \tag{6.17}$$

$$\begin{aligned}
\frac{dc^C_{\mathrm{ADP}}}{dt} &= r^C_{\mathrm{HK}} + r^C_{\mathrm{PFK}} + a_{\mathrm{ATP},1} r^C_{\mathrm{SYNT},1} + a_{\mathrm{ATP},2} r^C_{\mathrm{SYNT},2} + m_{\mathrm{ATP}} - \\
&\quad - 2r^C_{\mathrm{ADK}} - r^C_{\mathrm{RES},2} - r^C_{\mathrm{PK}} - r^C_{\mathrm{TR,ADP}} - \mu c^C_{\mathrm{ADP}} \qquad (6.18)\\
\frac{dc^C_{\mathrm{ATP}}}{dt} &= r^C_{\mathrm{RES},2} + r^C_{\mathrm{PK}} + \frac{V_M}{V_C} r^M_{\mathrm{TR,ATP}} + r^C_{\mathrm{ADK}} - r^C_{\mathrm{HK}} - r^C_{\mathrm{PFK}} - \\
&\quad - a_{\mathrm{ATP},1} r^C_{\mathrm{SYNT},1} - m_{\mathrm{ATP}} - a_{\mathrm{ATP},2} r^C_{\mathrm{SYNT},2} - \mu c^C_{\mathrm{ATP}} \qquad (6.19)\\
\frac{dc^C_{\mathrm{AMP}}}{dt} &= r^C_{\mathrm{ADK}} - \mu c^C_{\mathrm{AMP}} \qquad (6.20)\\
\frac{dc^C_{\mathrm{NADH}}}{dt} &= r^C_{\mathrm{RES},2} + a_{\mathrm{NADH},2} r^C_{\mathrm{SYNT},2} - r^C_{\mathrm{RES},1} - r^C_{\mathrm{ADH}} - r^C_{\mathrm{ALDH}} - \\
&\quad - r^C_{\mathrm{NADHDH}} - \mu c^C_{\mathrm{NADH}} \qquad (6.21)\\
\frac{dc^M_{\mathrm{ATP}}}{dt} &= r^M_{\mathrm{ATP,R}} + r^M_{\mathrm{ATP,T}} - r^M_{\mathrm{TR,ATP}} - \mu c^M_{\mathrm{ATP}} \qquad (6.22)\\
\frac{dc^M_{\mathrm{NADH}}}{dt} &= r^M_{\mathrm{NADH,T}} - r^M_{\mathrm{NADHQR}} - \mu c^M_{\mathrm{NADH}} \quad . \qquad (6.23)
\end{aligned}$$

Remark: After comparison with the original version of the model, see (Baltes, 1996), some minor modifications have been made to the model description in (Rizzi et al., 1997) due to what appears to be typing errors in the latter, see (Johansson, 2007).

The fluxes r have rate equations based on Michaelis-Menten, Hill or other types of enzyme kinetics gathered from the literature or proposed by the authors. For example, the triosephosphate isomerase reaction converting dihydroxyacetone phosphate to glyceraldehyde-3-phosphate has the rate expression

$$r^C_{\mathrm{TIS}} = r^{\max}_{\mathrm{TIS}} \frac{c^C_{\mathrm{DHAP}} - \frac{c^C_{\mathrm{GAP}}}{K_{eq,6}}}{K_{\mathrm{DHAP},6}(1 + \frac{c^C_{\mathrm{GAP}}}{K_{\mathrm{GAP},6}}) + c^C_{\mathrm{DHAP}}} \quad . \qquad (6.24)$$

Most of the fluxes in the model are rational expressions with the exception of those fluxes where the Hill coefficients are not integers. This fact is important for the identifiability analysis and will be discussed later.

The above model was evaluated in (Rizzi et al., 1997) on the basis of experimental observations previously described in (Theobald et al., 1997). The model predictions were compared to the experimental results and the parameters were estimated from the data (Rizzi et al., 1997). We are now going to use the theory of identifiability to find out whether the kinetic parameters of this model can be uniquely determined from a perfect set of experimental data. We must first formulate a control system for the model. For this, the appropriate set of inputs and outputs must be chosen from the description of the experimental setting in (Theobald et al., 1997). In the latter, a

methodology was developed where the changes in metabolite concentrations after a glucose feed pulse (a fast injection of a certain volume of glucose in the medium, (Theobald et al., 1997), were measured over time. The initial conditions were the priorly-known values of the metabolite concentrations under so-called "steady-state growth" - a condition when biomass concentration (and other factors) has stabilised to a constant value for the culture, see (Theobald et al., 1997) for details.

In order to translate the information in the above paragraph into mathematical language, we include a perfect measurement of all metabolite concentrations c (thus including the given initial conditions) in the outputs of the control system:

$$\begin{cases} \dot{p} & = & 0 \\ \dot{c} & = & f(c,p,u) \\ y & = & c \\ (c(0) & = & c^0) \quad , \end{cases} \tag{6.25}$$

where c is the vector of metabolite concentrations, f is the right-hand side of the equation array 6.3 and we denote all the model parameters by p. The initial conditions are in parenthesis as the information they provide is included in the output set. The input u is assumed to be the glucose feed.

6.3 Identifiability analysis

We performed an identifiability analysis of the above control system. For this, Sedoglavic's implementation (see Subsection 3.5) was used as the model has around a hundred parameters which makes calculations by hand very difficult. As the algorithm works only for rational control systems, we approximated any non-integer values of the Hill coefficients by integers. Of course, in general, such approximations can have an important effect on the identifiability of the system. This turns out not to be the case for Rizzi's model, as shown in the next section.

Sedoglavic's algorithm produced the following results - the control system was not identifiable with transcendence degree 2 and the non-identifiable parameters were the kinetic parameters in two of the rate expressions - the expression for the flux $r^C_{\mathrm{RES},2}$ and the one for r^M_{PDH} which have the following form:

$$r^C_{\mathrm{RES},2} = r^{\max}_{\mathrm{RES},2} \frac{\frac{c^C_{\mathrm{NAD}+}}{K_{\mathrm{NAD},7}} A^{n_{1,7}-1} + L_{0,7} \frac{c^C_{\mathrm{NAD}+}}{K'_{\mathrm{NAD},7}} B^{n_{1,7}-1}}{A^{n_{1,7}} + L_{0,7} B^{n_{1,7}}} \frac{{c^C_{\mathrm{GAP}}}^{n_{2,7}}}{K_{\mathrm{GAP},7} + {c^C_{\mathrm{GAP}}}^{n_{2,7}}} , \tag{6.26}$$

where

$$\begin{aligned} A &= 1 + \frac{c^C_{\mathrm{NAD+}}}{K_{\mathrm{NAD},7}} + \frac{c^C_{\mathrm{NADH}}}{K_{\mathrm{NADH},7}} \\ B &= 1 + \frac{c^C_{\mathrm{NAD+}}}{K'_{\mathrm{NAD},7}} + \frac{c^C_{\mathrm{NADH}}}{K'_{\mathrm{NADH},7}} \end{aligned} \tag{6.27}$$

and

$$\begin{aligned} r^M_{\mathrm{PDH}} &= r^{\max}_{\mathrm{PDH}} c^C_{\mathrm{Pyr}} c^M_{\mathrm{NAD+}} / (K_{\mathrm{NAD},13} c^C_{\mathrm{Pyr}} + K_{\mathrm{Pyr},13} c^M_{\mathrm{NAD+}} + \\ &+ \frac{K_{I-\mathrm{Pyr},13} K_{\mathrm{NAD},13}}{K_{I-\mathrm{NADH},13}} c^M_{\mathrm{NADH}} + c^C_{\mathrm{Pyr}} c^M_{\mathrm{NAD+}} + \frac{K_{\mathrm{NAD},13}}{K_{I-\mathrm{NADH},13}} c^C_{\mathrm{Pyr}} c^M_{\mathrm{NADH}}). \end{aligned} \tag{6.28}$$

Observe that the concentrations $C^C_{\mathrm{NAD+}}$ and C^M_{NAD} are used in the rate equations although no differential equations are formulated for them in Rizzi's model. Instead, these are defined in (Rizzi et al., 1997) as:

$$c^C_{\mathrm{NAD+}} = p_1 - c^C_{\mathrm{NADH}} \tag{6.29}$$

$$c^M_{\mathrm{NAD+}} = p_2 - c^M_{\mathrm{NADH}} \quad , \tag{6.30}$$

where p_1 and p_2 are known constants.

More results from the identifiability analysis are shown in Appendix 8.4.

6.4 Symmetry

The results obtained from the Sedoglavic implementation are probabilistic - their validity must be ascertained by the actual finding of symmetries in the model. From the theory described in Section 3, we know that it is necessary to find two derivations that each of them give rise to a symmetry in the model. The fact that the non-identifiable parameters can be separated into two groups, each belonging to a rate equation, suggests the possibility that the symmetries may be found within each rate expression (since the kinetic parameters of the fluxes $r^C_{\mathrm{RES},2}$ and r^M_{PDH} are not used anywhere else in the model). If this is true, the calculation of the symmetries may be greatly simplified - the formal procedure from Subsection 3.4 for the 11 non-identifiable parameters can otherwise be rather cumbersome. In order to verify this hypothesis, we first used Sedoglavic's algorithm on our control system where we added measurements of the fluxes $r^C_{\mathrm{RES},2}$ and r^M_{PDH} to the set of outputs:

$$\begin{cases} \dot{p} = 0 \\ \dot{c} = f(c,p,u) \\ y = \begin{pmatrix} c \\ r^C_{\mathrm{RES},2} \\ r^M_{\mathrm{PDH}} \end{pmatrix} \quad . \end{cases} \tag{6.31}$$

The idea behind this test is that if this system turns out to result in a transcendence degree of 2 and the same non-identifiable parameters as before, then there exist two families of symmetries in these parameters that leave both c and $r^C_{\mathrm{RES},2}$ and r^M_{PDH} invariant (see Subsection 3.4). This means that the rate expressions for $r^C_{\mathrm{RES},2}$ and r^M_{PDH} themselves have symmetries in their kinetic parameters.

We tested the above system in the Sedoglavic implementation and found our hypothesis to be true. The rate expressions for $r^C_{\mathrm{RES},2}$ and r^M_{PDH} were then analysed further to find the symmetries in the parameters. The detailed analysis can be found in Paper I. We found the following families of symmetries (see Subsection 3.4) for $r^C_{\mathrm{RES},2}$:

$$\sigma_\lambda : \left\{ \begin{array}{l} r^{\max}_{\mathrm{RES},2} \\ K_{\mathrm{NAD},7} \\ K'_{\mathrm{NAD},7} \\ K_{\mathrm{NADH},7} \\ K'_{\mathrm{NADH},7} \\ L_{0,7} \end{array} \right\} \to \left\{ \begin{array}{l} \frac{r^{\max}_{\mathrm{RES},2} p_1}{K_{\mathrm{NAD},7}(1-\lambda^{-1/n_{1,7}})+p_1} \\ \frac{K_{\mathrm{NAD},7}\lambda^{-1/n_{1,7}} p_1}{K_{\mathrm{NAD},7}(1-\lambda^{-1/n_{1,7}})+p_1} \\ \frac{(K'_{\mathrm{NAD},7}-K_{\mathrm{NAD},7}(1-\lambda^{-1/n_{1,7}}))p_1}{K_{\mathrm{NAD},7}(1-\lambda^{-1/n_{1,7}})+p_1} \\ \frac{K_{\mathrm{NADH},7}\lambda^{-1/n_{1,7}} p_1}{K_{\mathrm{NADH},7}(1-\lambda^{-1/n_{1,7}})+p_1} \\ \frac{(K'_{\mathrm{NAD},7}-K_{\mathrm{NAD},7}(1-\lambda^{-1/n_{1,7}}))p_1}{K_{\mathrm{NAD},7}(1-\lambda^{-1/n_{1,7}})+\frac{K'_{\mathrm{NAD},7}}{K'_{\mathrm{NADH},7}}p_1} \\ \frac{\lambda L_{0,7}(K'_{\mathrm{NAD},7}-K_{\mathrm{NAD},7}(1-\lambda^{-1/n_{1,7}}))^{n_{1,7}}}{K^{n_{1,7}}_{\mathrm{NAD},7}} \end{array} \right\} \tag{6.32}$$

and for r^M_{PDH}:

$$\sigma_\lambda : \left\{ \begin{array}{l} r^{\max}_{\mathrm{PDH}} \\ K_{\mathrm{NAD},13} \\ K_{\mathrm{PYR},13} \\ K_{I-\mathrm{PYR},13} \\ K_{I-\mathrm{NADH},13} \end{array} \right\} \to \left\{ \begin{array}{l} \lambda r^{\max}_{\mathrm{PDH}} \\ \lambda K_{\mathrm{NAD},13} + p_2(\lambda-1) \\ \lambda K_{\mathrm{PYR},13} \\ \frac{\lambda K_{\mathrm{NAD},13} K_{I-\mathrm{PYR},13}}{(1-\lambda)K_{I-\mathrm{NADH},13}+\lambda K_{\mathrm{NAD},13}} \\ \frac{K_{\mathrm{I-NADH},13}(\lambda K_{\mathrm{NAD},13}+p_2(\lambda-1))}{(1-\lambda)K_{I-\mathrm{NADH},13}+\lambda K_{\mathrm{NAD},13}} \end{array} \right\} \tag{6.33}$$

We see that the constants p_1 and p_2 appear in the symmetries. In fact, it is exactly the equations (6.29) and (6.30) corresponding to the conserved moiety assumptions in the model that cause the system to be unidentifiable. This is discussed in detail in paper I.

We have now verified that the kinetic parameters $r^{\max}_{\mathrm{RES},2}$, $K_{\mathrm{NAD},7}$, $K'_{\mathrm{NAD},7}$, $K_{\mathrm{NADH},7}$, $K'_{\mathrm{NADH},7}$, $L_{0,7}$, $r^{\max}_{\mathrm{PDH}}$, $K_{\mathrm{NAD},13}$, $K_{\mathrm{Pyr},13}$, $K_{I-\mathrm{NADH},13}$, and $K_{I-\mathrm{Pyr},13}$ cannot be identified from any experimental data. If parameter estimation is to be performed on Rizzi's model, for example by a numerical procedure where the error between model predictions and experimental results is minimised, then one of the parameters in each of the groups $r^{\max}_{\mathrm{RES},2}$, $K_{\mathrm{NAD},7}$,

$K'_{\mathrm{NAD},7}$, $K_{\mathrm{NADH},7}$, $K'_{\mathrm{NADH},7}$, $L_{0,7}$ and $r^{\max}_{\mathrm{PDH}}$, $K_{\mathrm{NAD},13}$, $K_{\mathrm{Pyr},13}$, $K_{I-\mathrm{NADH},13}$, $K_{I-\mathrm{Pyr},13}$ must be fixed to a value, while varying the rest of the parameters.

Remark: Using Sedoglavic's algorithm, we investigated the identifiability of this model with other sets of outputs than the ones discussed above (some of the results are shown in Appendix 8.4). As well as including all possible outputs - all concentrations c_j and all fluxes r_j, we also tried to limit the number of measurements by finding a smaller set of outputs which produced the same transcendence degree for the system. One such example is the set C^e_{GLYC}, C^e_{AC}, C^e_{ETOH} and $C^e_{\mathrm{CO_2}}$. Measuring these concentrations should in theory (with perfect error-free measurements) produce the same information on the parameter values as measuring all concentrations and all fluxes. Sedoglavic's algorithm can thus be used in the practical planning of an experiment for the purpose of parameter estimation.

6.5 General features of kinetic models and identifiability

Kinetic models of metabolism can often be described by the following structure, see Chapter 8 of the book on metabolic engineering by Stephanopoulos et al. (1998):

$$\left\{ \dot{c}_j \;=\; u_j + \textstyle\sum_i \nu_{ij} \cdot r_i(c, p_i) - \mu c_j \quad \forall j \quad , \right. \tag{6.34}$$

for each metabolite j. The coefficients ν_{ij} are the stoichiometric coefficients associated to each reaction.

In Rizzi's model, the unidentifiable parameters can be separated into two groups each associated to a single rate equation. The question is whether this is true for all non-identifiable kinetic models of the above form. We will show that this is true when a model is not identifiable from a full-state measurement, that is, from an output set $y = c$, as in the Rizzi case study.

The first step in the construction of a kinetic model for cellular metabolism is usually formulating a network of fluxes which, at steady-state, obey Kirchhoff's law at every node. This underlying steady-state model must be such that the stationary values of all fluxes r_i can be calculated uniquely from the set of linear equations for r obtained at steady-state. Mathematically, this means that the stoichiometric matrix $[\nu_{ij}]$ is of full rank.

If we rewrite the first equation in (6.34), we have:

$$\dot{c}_j - u_j + \mu c_j \;=\; \textstyle\sum_i \nu_{ij} r_i(c, p_i) \quad . \tag{6.35}$$

Since $y = c$, all the quantities on the left-hand side are known, if we assume that we know the value of the specific growth rate μ. The matrix $[\nu_{ij}]$ is

assumed to be of full rank and we thus have unique values of $r_i(c, p_i)$ completely determined by a given set of inputs and outputs. This means that if the system is not identifiable and there exist non-identifiable parameters such that there is a symmetry in them leaving the inputs and outputs invariant, then this symmetry will also leave the r_i:s invariant. Since each flux r_i only involves the parameters in the subset p_i of p, then any symmetry in p leaving r_i invariant must involve only the parameters in p_i. In result, a kinetic model of metabolism of the form (6.34) is not identifiable only if some of the rate expressions are symmetric in their respective kinetic parameters.

The fact that unidentifiability can in such case be found locally in single reaction rate expressions can be used to simplify the identifiability analysis for metabolic models. This is discussed in detail in Paper I, where the simplified analysis is applied to several well-cited kinetic models of glycolysis. The sources of unidentifiable parameters in these models can be traced to the introduction of conserved moiety assumptions. In Paper I, we develop a general method for determining whether a conserved moiety renders a rate expression unidentifiable, as well as a method for reparameterisation into identifiable parameters. The reparameterisation shows which combinations of the original parameters can be uniquely estimated from the data. We provide all symmetry transformations, which leave the output invariant. Furthermore, we show that identifiable rate expressions are enough to ensure identifiability of the entire model, provided that a sufficient set of measurements is available.

7 Discussion

The subject of this report is the problem of investigating the observability of nonlinear control systems, a special case of which is the identifiability problem. We have presented a review of two rather different theoretical approaches to the problem used in literature - the differential-geometric approach and the algebraic one. Each of the two approaches leads to a test for the observability of a class of control systems. The differential-geometric approach covers analytical control systems of the form

$$\Sigma \begin{cases} \dot{x} &= f(x,u) = g^0(x) + g(x)u \\ y &= h(x) \quad , \end{cases}$$

while the algebraic one treats rational systems:

$$\Sigma \begin{cases} \dot{x} &= f(x,u) \\ y &= h(x,u) \quad . \end{cases}$$

Both approaches lead to the so-called observability rank test where the rank of the space spanned by gradients of the Lie-derivatives of the output functions is calculated. In the algebraic approach there is an upper bound derived for the number of Lie-derivatives that have to be considered in the test for rational systems. In Section 4 we derive the same upper bound for the number of Lie-derivatives of the output functions for the class of analytical systems affine in the input variables that are considered in the differential-geometric approach. It remains as future work to investigate the validity of such an upper bound for other classes of control systems.

The identifiability problem is a special case of observability and we can thus use the previously derived rank test in the investigation of identifiability. This has been done in a case study of a dynamic model of the metabolism of *S. cerevisiae*. By finding symmetries in the model, we show that certain model parameters cannot be identified from any set of experimental data. The results from the treatment of this model are generalised to show how the special structure of kinetic models of metabolism considerably simplify the analysis of their identifiability and especially the derivation of symmetries. We show that using conservation laws in metabolic modelling can have an effect on the parameter identifiability of the models.

8 Appendix

8.1 Why do d and L_f commute when f depends on a control variable u?

This appendix contains some calculations that were deferred from Section 4.

Consider a control system of the form:

$$\Sigma \left\{ \begin{array}{rcl} \dot{x} & = & f(x,u) = h^0(x) + h(x)u \\ y & = & g(x) \quad , \end{array} \right. \tag{8.1}$$

where the elements of the n-dimensional vectors h^0 and h are analytic functions of x and u, where u is the single control variable. Denote, as in Section 2, the flow corresponding to the time-dependent vector field $f(x,u)$ by $\Phi_u(t,x)$ which is then an n dimensional vector. Let $\psi(x,u)$ be an analytic function of x,u and u's time derivatives. We will show that:

$$L_f d\psi = dL_f\psi \quad . \tag{8.2}$$

Take the i-th element of $dL_f\psi$. It is:

$$\begin{aligned} \frac{\partial}{\partial x_i} L_f\psi &= \frac{\partial}{\partial x_i}\Big(\sum_{j=1}^{n} f_j \frac{\partial \psi}{\partial x_j} + \sum_{l=0} \frac{\partial \psi}{\partial u^{(l)}} U^{(l+1)}\Big) = \\ &= \sum_{j=1}^{n} \frac{\partial}{\partial x_i}(f_j \frac{\partial \psi}{\partial x_j}) + \sum_{l=0} \frac{\partial}{\partial x_i}\Big(\frac{\partial \psi}{\partial u^{(l)}} U^{(l+1)}\Big) = \\ &= \sum_{j=1}^{n} \frac{\partial f_j}{\partial x_i}\frac{\partial \psi}{\partial x_j} + \sum_{j=1}^{n} \frac{\partial^2 \psi}{\partial x_i \partial x_j} f_j + \sum_{l=0} \frac{\partial}{\partial x_i}(\frac{\partial \psi}{\partial u^{(l)}}) U^{(l+1)} \quad . \end{aligned}$$

Now consider the i-th element of the covector $L_f d\psi$. By the definition of Lie derivative we find,

$$\begin{aligned} &\frac{d}{dt}\Big(\frac{\partial \psi\big(\Phi_u(t,x),u(t)\big)\big)}{\partial x_i}\Big)\Big|_{t=0} = \frac{d}{dt}\Big(\sum_{j=1}^{n} \frac{\partial \psi}{\partial \Phi_{u,j}}\frac{\partial \Phi_{u,j}}{\partial x_i}\Big)\Big|_{t=0} = \\ &= \sum_{j=1}^{n} \frac{d}{dt}\Big(\frac{\partial \psi}{\partial \Phi_{u,j}}\frac{\partial \Phi_{u,j}}{\partial x_i}\Big)\Big|_{t=0} = \\ &= \sum_{j=1}^{n} \left(\frac{d}{dt}\Big(\frac{\partial \psi}{\partial \Phi_{u,j}}\Big)\Big|_{t=0} \frac{\partial \Phi_{u,j}}{\partial x_i}\Big|_{t=0} + \frac{d}{dt}\Big(\frac{\partial \Phi_{u,j}}{\partial x_i}\Big)\Big|_{t=0} \frac{\partial \psi}{\partial \Phi_{u,j}}\Big|_{t=0}\right) = \end{aligned}$$

$$= \sum_{j=1}^{n} \Bigg(\Big(\sum_{k=1}^{n} \frac{\partial}{\partial \Phi_{u,k}} (\frac{\partial \psi}{\partial \Phi_{u,j}}) \Big|_{t=0} \frac{d\Phi_{u,k}}{dt} \Big|_{t=0} +$$
$$+ \sum_{l=0} \frac{\partial}{\partial u^{(l)}} (\frac{\partial \psi}{\partial \Phi_{u,j}}) \Big|_{t=0} U^{(l+1)}) \Big) \delta_{ij} + \frac{\partial \psi}{\partial x_j} \frac{\partial f_j}{x_i} \Bigg) =$$
$$= \sum_{j=1}^{n} \Bigg(\Big(\sum_{k=1}^{n} \frac{\partial^2 \psi}{\partial x_k \partial x_j} f_k + \sum_{l=0} \frac{\partial}{\partial u^{(l)}} (\frac{\partial \psi}{\partial x_j}) U^{(l+1)} \Big) \delta_{ij} + \frac{\partial \psi}{\partial x_j} \frac{\partial f_j}{x_i} \Bigg) =$$
$$= \sum_{k=1}^{n} \frac{\partial^2 \psi}{\partial x_k \partial x_i} f_k + \sum_{l=0} \frac{\partial}{\partial u^{(l)}} (\frac{\partial \psi}{\partial x_i}) U^{(l+1)} + \sum_{j=1}^{n} \frac{\partial \psi}{\partial x_j} \frac{\partial f_j}{x_i} \quad .$$

One can now see that the i-th elements of the covectors $L_f d\psi$ and $dL_f \psi$ are the same, which shows the above equality.

8.2 A basis for derivations on $\mathbb{R}\langle U \rangle(x)$ that are trivial on $\mathbb{R}\langle U \rangle$

This appendix refers to Section 3.4 and is based on pp.371-372 in Lang (1993).

First of all, observe that the x_i form a transcendence basis for $\mathbb{R}\langle U \rangle(x)$ over $\mathbb{R}\langle U \rangle$. The transcendence degree of the field extension is thus n which is also the dimension of $Der_{\mathbb{R}\langle U \rangle}\mathbb{R}\langle U \rangle(x)$. Consider the n derivations $D_i = \frac{\partial}{\partial x_i}$. Clearly, $D_i x_j = \delta_{ij}$. Let D be a derivation in $Der_{\mathbb{R}\langle U \rangle}\mathbb{R}\langle U \rangle(x)$ and let $Dx_i = w_i$ (a derivation is defined by its action on a set of generators of the transcendence basis). Then $D = \sum_i w_i D_i$ and thus the D_i:s form a basis for $Der_{\mathbb{R}\langle U \rangle}\mathbb{R}\langle U \rangle(x)$. Therefore the partial derivatives $\frac{\partial}{\partial x_i}$ form a basis for the derivations on $\mathbb{R}\langle U \rangle(x)$ that are trivial on $\mathbb{R}\langle U \rangle$.

8.3 Nomenclature for Rizzi's model

This appendix refers to Section 6.

For the details of the kinetic model by Rizzi et al we refer to (Rizzi et al., 1997) but in an attempt to make this report somewhat self-sufficient we here provide nomenclature for the abbreviations and parameters used in Section 6.

8.3.1 Superscripts

e	extracellular
C	cytoplasmic
M	mitochondrial

8.3.2 Symbols and abbreviations

a_j	stoichiometric coefficient
S_G	gas supply rate
V_M	volume of mitochondria
V_C	volume of cytoplasm
D	dilution rate
ρ	specific volume
μ	specific growth rate
c_{X}	biomass concentration
m_{ATP}	maintenance coefficient for ATP

8.3.3 Metabolites

AC	acetic acid
ADP	adenosine diphosphate
ALDE	acetaldehyde
AMP	adenosine monophosphate
ATP	adenosine triphosphate
DHAP	dihydroxyacetone phosphate
ETOH	ethanol
FBP	fructose 1,6-bisphosphate
F6P	fructose 6-phosphate
GAP	glyceraldehyde 3-phosphate
GLC	glucose
GLYC	glycerol
G6P	glucose 6-phosphate
NAD^+/NADH	nicotinamide adenine dinucleotide
PEP	phosphoenol pyruvate
PYR	pyruvate

8.3.4 Enzymes and flux indexes

ACETYL	acetate synthetase
ADH	alcohol dehydrogenase
ADK	adenylate kinase
ALDH	acetaldehyde dehydrogenase
ALDO	fructose bisphosphate aldolase
ATP,R	ATP formation via respiratory chain
ATP,T	ATP formation via the citric acid cycle
HK	hexokinase
NADHDH	NADH-dehydrogenase
NADHQR	NADH-Q-reductase
PDC	pyruvate decarboxylase
PDH	pyruvate dehydrogenase
PERM	hexose transporter
PFK	phosphofructo-1-kinase
PGI	phosphoglucose isomerase
PK	pyruvate kinase
RES,1	combination of glycerol-3-phosphate dehydrogenase and glycerol-3-phosphatase
RES,2	combination of glyceraldehyde 3-phosphate dehydrogenase and other enzymes
SYNT,1;SYNT,2	resulting rates for the formation of monomeric building blocks
TIS	triose phosphate isomerase
TR,ADP and TR,ATP	translocases for ADP and ATP respectively.

8.4 Other results on the identifiability of Rizzi's model

This appendix shows some of the results obtained from the identifiability analysis of Rizzi's model. They can be used in choosing a set of measurements in a hypothetical experiment.

Since the parameters c_{X} and ρ always appear together as $\frac{c_{\mathrm{X}}}{\rho}$ one of them must be known - otherwise it is clear that they will not be identifiable. We therefore assume that the value of one of them is measured in any hypothetical experiment when we perform the identifiability analysis, although we do not explicitly write them as outputs. The same applies for the parameters V_M and V_C appearing in $\frac{V_M}{V_C}$.

Outputs	**Transcendence degree**
$c^e_{\mathrm{GLYC}}, c^e_{\mathrm{AC}}, c^e_{\mathrm{ETOH}}, c^e_{\mathrm{CO_2}}$	2
$c^e_{\mathrm{GLYC}}, c^e_{\mathrm{AC}}, c^e_{\mathrm{ETOH}}, c^e_{\mathrm{CO_2}}, r^{\max}_{\mathrm{RES},2}, r^{\max}_{\mathrm{PDH}}$	0
all c_j:s	2
all c_j:s, $r^C_{\mathrm{RES},2}, r^M_{\mathrm{PDH}}$	2
all c_j:s, all r_j:s	2

One conclusion is that measuring c^e_{GLYC}, c^e_{AC}, c^e_{ETOH} and $c^e_{\mathrm{CO_2}}$ produces as much information for the theoretical identification of parameter values as making all possible measurements altogether.

References

Baltes, M. 1996. *Entwurf eines strukturierten Modells zur dynamischen Simulation des Katabolismus von Saccharomyces cerevisiae.* PhD thesis. Universität Stuttgart, Germany, VDI-Verlag Dusseldorf, Reihe, vol. 17, Biotechnik, No.148.

Diop, S. & Fliess, M. (1991a). Nonlinear observability, identifiability, and persistent trajectories. *Proceedings of the 30nd IEEE Conference on Decision and Control.* 714-719. IEEE Press: Brighton England.

Diop, S. & Fliess, M. (1991b). On nonlinear observability. In Commault C. and coll. Editors, *Proceedings of ECC'91, 1*, 152-157. Hermès

Diop, S. & Wang, M. (1993b). Equivalence between algebraic observability and local generic observability. In Commault C. and coll. Editors, *In Proceedings of the 32nd IEEE Conference on Decision and Control*, 2864-2865. IEEE Press: San Antonio Texas.

Evans, N., Chapman, M., Chapell, M., & Godfrey K. (2002). Identifiability of uncontrolled nonlinear rational systems. *Automatica, 38*, 1799-1805.

Hermann, R. & Krener, A.J. (1977). Nonlinear controllability and observability. *IEEE T. Aut. Contr., 22(5)*, 728-740.

Isidori, A. (1995). *Nonlinear Control Systems* (3rd Edition). London: Springer-Verlag.

Jacobsson, N. (1980). *Basic Algebra II.* Freeman and Company: San Fransisco.

Johansson, M.. (2007). *Kinetic Modelling of Glycolysis: Hexose Transporters, Gpd2p and Tdh1p in glucose repression and glycolytic rate control.* PhD thesis. Chalmers University of Technology, Gotheburg, Sweden.

Johnson, J. (1969). Kähler differentials and differential algebra. *Annals of Mathematics, 89*, 92-98.

Krener, J. (1985). $(ad_{f,g})$, $(ad_{f,g})$ and locally $(ad_{f,g})$ invariant and controllability distributions. *SIAM J. Control and Optimization, 23(4)*, 523-549.

Lang, S. (1993). *Algebra.* Addison-Wesley Publishing company, Inc..

Nelson, D. & Cox, M. (2000). *Lehninger principles of biochemistry.* Worth Publishers.

Ljung, L. & Glad, T. (1990). Parametrization of nonlinear model structures as linear regressions. *In Proceedings of the 11th IFAC world congress.* 67-71.

Margaria, G., Riccomagno, E., & White, L.J. (2004). Structural identifiability analysis of some highly structured families of statespace models using differential algebra. *J. Math. Biol. 49(5)*, 433-454.

Ollivier, F. (1990). *Le problème de l'identifiabilité structurelle globale: approche théorique, méthodes effectives et bornes de complexité.* PhD thesis, École polytechnique.

Pohjanpalo, H. (1978). System identifiability based on the power series expansion of the solution. *Math. Biosci., 41*, 21-33.

Rizzi, M., Baltes, M., Theobald, U., & Reuss, M. (1997). In Vivo Analysis of Metabolic Dynamics in *Saccharomyces cerevisiae*: II. Mathematical Model. *Biotechnol. Bioeng., 55*, 592-608.

Ritt, J. (1966). *Differential algebra.* Dover Publications.

Sedoglavic, A. (2002). A probabilistic algorithm to test local algebraic observability in polynomial time. *J. Symb. Comput., 33*, 735-755.

Segel, I.H. (1975). *Enzyme kinetics.* John Wiley & Sons: New York Chichester Brisbane Toronto Singapore.

Sontag, E. 1979. On the observability of polynomial systems, I: Finite-time problems. *SIAM J. Control and Optimisation, 17(1)*, 139-151.

Sontag, E. (1984). A concept of local observability. *Systems&Control Letters, 5*, 41-47.

Sontag, E. (1991). I/O equations for nonlinear systems and observation spaces. *In Proceedings of the 30nd IEEE Conference on Decision and Control.* 720-725. IEEE Press: Brighton England.

Stephanopoulos, G., Aristidou, A., & Nielsen, J. 1998. *Metabolic engineering.* John Wiley & Sons: New York Chichester Brisbane Toronto Singapore

Sussmann, H. (1979). Single-input observability of continuous-time systems. *Math. Systems Theory, 12*, 371-393.

Teusink, B., Passarge, J., Reijenga, C.A., Esgalhado, E., van der Weijden, C.C., Schepper, M., Walsh, M.C., Bakker B., van Dam, K., Westerhoff, H.V., and Snoep, J.L.. (2000). Can yeast glycolysis be understood in terms

of *in vitro* kinetics of the constituent enzymes? testing biochemistry. *Eur. J. Biochem., 267(17)*, 5313-5329.

Theobald, U., Mailinger, W., Baltes, M., Rizzi, M., & Reuss, M. (1997). In vivo analysis of metabolic dynamics in *Saccharomyces cerevisiae*: I. Experimental observations. *Biotechnol. Bioeng. 55*, 305-316.

Vajda, S., Godfrey, K., & Rabitz, H. (1989). Similarity transformation approach to identifiability analysis of nonlinear compartmental models. *Math. Biosci. 93*, 217-248.

van der Put, M. & Singer, M. (2003). *Galois theory of linear differential equations*. Springer Verlag, New York, Berlin, Heidelberg.

Wang, Y. & Sontag, E. (1989). On two definitions of observation spaces. *Systems and control letters, 13*, 279-289.

Xia, X. & Moog, C.H. (2003). Identifiability of nonlinear systems with application to HIV/AIDS models. *IEEE T. Aut. Contr. 48(2)*, 330-336.

Paper I

Conservation laws and unidentifiability of rate expressions in biochemical models

M. Anguelova, G. Cedersund, M. Johansson, C.J. Franzén and B. Wennberg

Abstract: New experimental techniques in bioscience provide us with high-quality data allowing quantitative mathematical modelling. Parameter estimation is often necessary and, in connection with this, it is important to know whether all parameters can be uniquely estimated from available data, (i.e. whether the model is identifiable). Dealing essentially with models for metabolism, we show how the assumption of an algebraic relation between concentrations may cause parameters to be unidentifiable. If a sufficient data set is available, the problem with unidentifiability arises locally in individual rate expressions. A general method for reparameterisation to identifiable rate expressions is provided, together with a Mathematica code to help with the calculations. The general results are exemplified by four well-cited models for glycolysis.

1 Introduction

Throughout the years, a large number of mathematical models have been published that describe smaller or larger parts of the system of chemical reactions in living cells. Such models usually consist of systems of ordinary differential equations describing the reaction rates, and algebraic relations describing, for example conserved moieties.

Models of this kind naturally contain a large number of parameters that must be determined before the model can be used for simulation. For example, in a simple reaction of Michaelis–Menten type, the reaction rate is given by

$$r = \frac{V_{\max} c_S}{K_M + c_S}$$

where c_S denotes the concentration of a substrate, and $V_{\max}$ and K_M are two parameters. The parameter values may be known, but in other cases they must be determined through parameter estimation: by comparing simulated values for some concentrations with an experimentally obtained time series, it may be possible to determine a unique set of parameter values.

On the other hand, infinitely many parameter sets may produce exactly the same simulated values, and hence agree equally well with experimental data. The problem of determining a priori whether the parameters of a model can be uniquely determined from a given experiment is part of the general theory of identifiability analysis, and is the subject matter of this paper.

doi:10.1049/iet-syb:20060081

Paper first received 13th March 2006 and in revised form 9th May 2007

M. Anguelova and B. Wennberg are with Mathematical Sciences, Chalmers University of Technology, Gothenburg SE-412 96, Sweden and also with the Mathematical Sciences, Göteborg University, Gothenburg SE-41296, Sweden

G. Cedersund is with the Department of Cell Biology, Linköping University, Linköping SE-581 85, Sweden

M. Johansson and C. J. Franzén are with Chemical and Biological Engineering, Chemical Reaction Engineering, Chalmers University of Technology, Gothenburg SE-412 96, Sweden

E-mail: milena@math.chalmers.se

More precisely, we consider a family of kinetic models for cell metabolism, where some of the dynamic mass balances are replaced by algebraic relations. Although the results are not restricted to this case, we focus on examples where the algebraic relations express the conservation of $[NAD^+] + [NADH]$. This is a commonly used assumption in biochemical modelling and is only valid under short-term studies. In long-term studies, the synthesis or degradation of the components must be accounted for. We study three well-cited models of yeast metabolism as a in detail [1–3]. Some sets of parameters in these models are unidentifiable as a result of conserved $[NAD^+] + [NADH]$ moiety.

That such assumptions may lead to unidentifiability of parameters in some rate expressions is not new to the mathematical modelling community. A tool of the trade is to recover an identifiable rate expression by non-dimensionalisation or 'divide-through'. However, neither a mathematically stringent analysis of the reasons for such unidentifiability, nor a general method for treating the problem seems to have been published.

We provide a mathematical framework to allow for such unified treatment of a large class of rate expressions, including the particular examples from [1–3]. Thus, we develop a general method for determining whether a conserved moiety may render a rate expression unidentifiable, and a method for reparameterisation into identifiable parameters. The reparameterisation shows which combinations of the original parameters may be uniquely estimated from the data. We provide all symmetry transformations, which leave the output invariant. Furthermore, we show that identifiable rate expressions are enough to ensure identifiability of the entire model, provided that a sufficient set of measurements is available. In other words, all sources of structural unidentifiability can in such case be found locally in single reaction rate expression.

The general method is well adapted for implementation in symbolic programming languages, and we provide a Mathematica notebook that can be used for analysing a class of rate expressions, including all rational ones.

2 Identifiability of kinetic models

Kinetic models of metabolism are usually systems of differential equations, representing mass balances for each

metabolite j [4]. They can normally be written in the form

$$\frac{dc_j}{dt} = \sum_i v_{ji} \cdot r_i(\boldsymbol{c}, \boldsymbol{k}) \tag{1}$$

where c_j denotes the intracellular concentration of metabolite j, and $\boldsymbol{c}$ is the vector containing all such concentrations. The symbol r_i denotes the rate of reaction i, v_{ji} denotes the stoichiometric coefficient of metabolite j in reaction i and $\boldsymbol{k}$ denotes the kinetic parameters in the rate expressions of the model.

The property of structural identifiability guarantees the uniqueness of the parameters $\boldsymbol{k}$ for a given input–output structure corresponding to a set of measurements for the purpose of parameter estimation. This set of measurements usually consists of some metabolic concentrations and/or fluxes or combinations of these, denoted by the vector $\boldsymbol{y}$. The latter can then be written as a vector-valued function $\boldsymbol{g}$ of the variables in $\boldsymbol{c}$ and the parameters $\boldsymbol{k}$. Equation (1) together with

$$\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{c}, \boldsymbol{k}) \tag{2}$$

defines the input–output structure.

For realistic situations, sources of unidentifiability can be an insufficient number of measurements as well as a structural property of the model, which is independent of the input–output structure and renders it unidentifiable even from a perfect data set. In this paper, we focus on the latter and analyse structural identifiability when a perfect data set is available. Clearly, if a model is unidentifiable for such an input–output structure, it will not be identifiable for any realistic measurement set-ups. On the other hand, it should be noted that measuring a single (noise free) metabolic flux can be enough for the identification of all the parameters of an identifiable model.

A perfect data set for a model with a full stoichiometric matrix can be attained by measurements of all metabolic concentrations, as shown by the following calculation. Suppose that a model is not identifiable for an input–output structure of the form

$$\begin{cases} \dfrac{dc_j}{dt} = \sum_i v_{ji} \cdot r_i(\boldsymbol{c}, \boldsymbol{k}) \\ \boldsymbol{y} = \boldsymbol{c} \end{cases} \tag{3}$$

that is, there are at least two sets of parameters, $\boldsymbol{k}$ and $\boldsymbol{K}$, which produce the same output [the same solution $\mathbf{c}(t)$ to the system of differential equations above with given initial conditions]. This can be expressed as

$$\sum_i v_{ji} \cdot r_i(\boldsymbol{c}, \boldsymbol{k}) = \frac{dc_j}{dt} = \sum_i v_{ji} \cdot r_i(\boldsymbol{c}, \boldsymbol{K}) \tag{4}$$

If the underlying steady-state model has a stoichiometric matrix (the matrix with elements v_{ji}) of full rank, we must have $r_i(\boldsymbol{c}, \boldsymbol{k}) = r_i(\boldsymbol{c}, \boldsymbol{K})$ for all the rates r_i. Thus, measurements of metabolic fluxes cannot be used to distinguish between the two sets $\boldsymbol{k}$ and $\boldsymbol{K}$. Consequently, no amount of additional flux measurements will suffice to make (3) identifiable. If the stoichiometric matrix is not of full rank, then the perfect output set must also contain measurements of some metabolic fluxes r_i.

If a model is not identifiable from a perfect output set, then it must contain some over-parameterised rate expression, i.e. the reason for the unidentifiability can be traced back to isolated rate expressions. Indeed, since the parameter sets $\boldsymbol{k}$ and $\boldsymbol{K}$ are by assumption different, the equalities $r_i(\boldsymbol{c}, \boldsymbol{k}) = r_i(\boldsymbol{c}, \boldsymbol{K})$ imply that there is at least one r_i with a rate expression that has the same value for different combinations of its kinetic parameters. This means that structural unidentifiability has its source within individual kinetic expressions.

Consider first a simple example taken from Segel [5]

$$r = \frac{V_{max} c_S}{K_S(1 + (c_P/K_P)) + c_S} \tag{5}$$

where c_S and c_P are the concentrations of the substrate and product respectively. The expression contains three parameters: V_{max}, K_S, and K_P. By measuring $r(t_\ell)$, $c_S(t_\ell)$, and $c_P(t_\ell)$ at three different times t_ℓ, one finds three equations for the three unknowns, if the dynamic variables are varying sufficiently independently, and the problem can in principle be solved.

However, if there are algebraic constraints, it may be impossible to find a sufficient number of independent equations for the unknown rate parameters. In the Segel example, if $c_S + c_P = a$, where a is a constant, then we obtain

$$\begin{aligned} r &= \frac{V_{max} c_S}{K_S(1 + ((a - c_S)/K_P)) + c_S} \\ &= \frac{c_S}{(K_S[1 + (a/K_P)]/V_{max}) + ([1 - (K_S/K_P)]/V_{max}) c_S} \\ &= \frac{c_S}{k_1 + k_2 c_S} \end{aligned} \tag{6}$$

where

$$k_1 = \frac{K_S[1 + (a/K_P)]}{V_{max}} \quad \text{and} \quad k_2 = \frac{[1 - (K_S/K_P)]}{V_{max}} \tag{7}$$

It now becomes clear that the original three parameters K_S, K_P and V_{max} are effectively only two independent ones. One can find infinitely many combinations of K_S, K_P and V_{max} that result in exactly the same rate expression value. For example, if one multiplies V_{max} by any constant λ and then adjusts the other two parameters according to the scheme below, the value of the rate expression does not change at all

$$V_{max}^{new} = \lambda V_{max} \tag{8a}$$

$$K_S^{new} = \lambda K_S - a(1 - \lambda) \tag{8b}$$

$$K_P^{new} = \frac{\lambda K_S - a(1 - \lambda)}{1 - \lambda + \lambda(K_S/K_P)} \tag{8c}$$

In other words, there is a one-parameter family of transformations of the set of parameters, which leave the reaction rate invariant. Consequently, if such a rate equation is included in a metabolic model, not all of its kinetic parameters will be identifiable from experimental data. In (6), one can identify only the parameter combinations k_1 and k_2, and before attempting to fit a model to experimental data, one should first rewrite the model in an identifiable form.

We note here that algebraic constraints may be introduced explicitly in the model, or may be hidden in the formulation of (1). When this is the case, the stoichiometric matrix is not of full rank.

The theory of identifiability is well studied (see for example [6–14] and the references therein), and tools for automated identifiability analysis are available. They are generally based on differential algebraic techniques, and the computational time grows exponentially with the number of variables and parameters [9, 11–13]. However, a probabilistic algorithm that is also useful for large systems has been constructed by Sedoglavic [14]. The

methods have been applied to models for yeast glycolysis as discussed, for example in [15].

There are also published algorithms that efficiently look for conserved quantities in models of the form (1), a recent example being [16]. Three of the models used as examples in this paper are accessible in the JWS Online database [17], where models can be searched for dependencies for the variables, (i.e. conserved moieties).

A final comment here is that methods for practical identifiability must also handle noise and other measurement limitations, and that the actual parameter estimation usually is based on some kind of maximum likelihood optimisation [18].

3 Examples

The main biochemical interactions of glycolysis were characterised over 100 years ago. Some of the earlier modelling attempts occurred in the 1960s [19–21], but these were mostly minimal models trying to explain, for example, the temporal oscillations found in 1957 [22]. Since then the models have grown in size and comprehension, and there have been several attempts at constructing quantitative models. Examples are found, for instance, in the work of Rizzi *et al.* [1], Teusink *et al.* [2], Hynne *et al.* [3] and Lambeth *et al.* [23].

It is from [1, 2] and [3] that we take examples of reaction rate expressions which are unidentifiable because of a conserved pool of, for instance, $NADH^+$ NAD $^+$ and ATP + ADP + AMP. If the model parameters are to be identified, these rate expressions should be reparameterised.

Analysing these models with the method of Sedoglavic [14] would give a list of unidentifiable parameters. However, it will become apparent that in these cases the problem of unidentifiability is located to individual reaction rate expressions. Therefore, the methods explained in Section 4 can be used directly on each and every rate expression and would give all the needed information.

In this section, we discuss the unidentifiable reaction rate expressions from the models in [1, 2] and [3] and propose a general form for reparameterisation. For completeness, we also give examples of similar rate expressions from [23], where the introduction of a conserved quantity does not lead to unidentifiability.

For the four models above, all instances of rate expressions where the variables from a conservation law appear together have been analysed. The results are stated below and the details of the treatment are provided in the supplementary material available online.

3.1 Rizzi model

The first example is a kinetic model of the central metabolic pathways in *Saccharomyces cerevisiae* formulated in [1]. The model includes the reactions of glycolysis, the citric acid cycle, the glyoxylate cycle and the respiratory chain and a simplified description of biosynthesis in growing cells. Subcellular localisation of reactions and metabolites are taken into account by including cytosolic and mitochondrial compartments. We refer to [1] for the complete model description and focus here only on the aspects that are relevant for our analysis.

The concentrations of cytoplasmic and mitochondrial NAD^+ are calculated under the assumption that the total concentrations of nicotinamide nucleotides are constant in each compartment. This results in the following equations for cytoplasmic and mitochondrial NAD^+

$$c^C_{NAD^+} = p_1 - c^C_{NADH} \tag{9a}$$

$$c^M_{NAD^+} = p_2 - c^M_{NADH} \tag{9b}$$

where p_1 and p_2 denote the constant total concentrations of nicotinamide nucleotides in the cytoplasm and mitochondria, respectively.

There are two reaction rates in this model that contain unidentifiable parameters: the pyruvate dehydrogenase (PDH) reaction, and the lumped reactions (RES2) from glyceraldehyde 3-phosphate to phosphoenolpyruvate; these are the only two kinetic expressions that contain all variables from the conserved moieties.

The pyruvate dehydrogenase complex catalyses the reaction

$$\text{Pyruvate} + \text{CoA} + \text{NAD}^+ \longrightarrow \text{AcCoA} + \text{CO}_2 + \text{NADH} + \text{H}^+ \tag{10}$$

In the Rizzi model, its rate equation is formulated

$$r^M_{PDH} = \frac{r^{max}_{PDH} c^C_{Pyr} c^M_{NAD^+}}{\left(\begin{array}{l} K_{NAD,13} c^C_{Pyr} + K_{Pyr,13} c^M_{NAD^+} \\ + \left[\dfrac{(K_{I-Pyr,\,13} K_{NAD,13})}{(K_{I-NADH,13})} \right] c^M_{NADH} + c^C_{Pyr} c^M_{NAD^+} \\ + \left[\dfrac{K_{NAD,13}}{(K_{I-NADH,13})} \right] c^C_{Pyr} c^M_{NADH} \end{array} \right)} \tag{11}$$

where superscripts C and M refer to cytoplasmic and mitochondrial concentrations, respectively. Inserting (9b) into (11) gives

$$r^M_{PDH} = \frac{r^{max}_{PDH} c^C_{Pyr} (p_2 - c^M_{NADH})}{\left(\begin{array}{l} (K_{NAD,13} c^C_{PYR} + K_{PYR,13} (p_2 - c^M_{NADH}) \\ + \left[\dfrac{(K_{I-PYR,13} K_{NAD,13})}{(K_{I-NADH,13})} \right] c^M_{NADH} + c^C_{Pyr} (p_2 - c^M_{NADH}) \\ + \dfrac{K_{NAD,13}}{(K_{I-NADH,13})} c^C_{Pyr} c^M_{NADH} \end{array} \right)} \tag{12}$$

After rearranging terms, the right hand side becomes

$$\frac{c^C_{Pyr} (p_2 - c^M_{NADH})}{\left(\begin{array}{l} (p_2 K_{PYR,13} / r^{max}_{PDH}) + ((K_{NAD,13} + p_2) / r^{max}_{PDH}) c^C_{Pyr} \\ + \dfrac{[(K_{I-Pyr,13} K_{NAD,13}) / (K_{I-NADH,13})] - K_{Pyr,13}}{r^{max}_{PDH}} c^M_{NADH} \\ + \dfrac{[K_{NAD,13} / (K_{I-NADH,13})] - 1}{r^{max}_{PDH}} c^C_{Pyr} c^M_{NADH} \end{array} \right)} \tag{13}$$

Thus, (11) can be rewritten as

$$r^M_{PDH} = \frac{c^C_{Pyr} (p_2 - c^M_{NADH})}{k_1 + k_2 c^C_{Pyr} + k_3 c^M_{NADH} + k_4 c^C_{Pyr} c^M_{NADH}} \tag{14}$$

The five original parameters are thus combined into four new parameters, that can be identified. Solving for r^{max}_{PDH}, $K_{NAD,13}$, $K_{Pyr,13}$, $K_{I-NADH,13}$ and $K_{I-Pyr,13}$, in terms of k_1, k_2, k_3 and k_4, result in a one parameter family of solutions.

The RES2 reaction is a lumped reaction of the glycolytic steps from glyceraldehyde 3-phosphate (GAP) to

phosphoenolpyruvate, with the rate equation

$$r^C_{\mathrm{RES},2} = r^{\max}_{\mathrm{RES},2} \frac{(c^C_{\mathrm{NAD}^+}/K_{\mathrm{NAD},7})A^{n_{1,7}-1} + L_{0,7}(c^C_{\mathrm{NAD}^+}/K'_{\mathrm{NAD},7})B^{n_{1,7}-1}}{A^{n_{1,7}} + L_{0,7}B^{n_{1,7}}} \times \frac{{c^C_{\mathrm{GAP}}}^{n_{2,7}}}{K_{\mathrm{GAP},7} + {c^C_{\mathrm{GAP}}}^{n_{2,7}}} \quad (15)$$

where

$$\begin{aligned} A &= 1 + \frac{c^C_{\mathrm{NAD}^+}}{K_{\mathrm{NAD},7}} + \frac{c^C_{\mathrm{NADH}}}{K_{\mathrm{NADH},7}} \\ B &= 1 + \frac{c^C_{\mathrm{NAD}^+}}{K'_{\mathrm{NAD},7}} + \frac{c^C_{\mathrm{NADH}}}{K'_{\mathrm{NADH},7}} \end{aligned} \quad (16)$$

This expression differs from the other reaction rates considered in this paper in that it is not a rational expression. However, it can be expressed as a (homogeneous) function including the rational expression A/B, and therefore this reaction can be handled in essentially the same way.

The conserved moiety, $c^C_{\mathrm{NAD}^+} = p_1 - c^C_{\mathrm{NADH}}$ appears only in the first part of the expression, and this can be reparameterised in the following way

$$r^{\max}_{\mathrm{RES},2} \frac{\frac{c^C_{\mathrm{NAD}^+}}{K_{\mathrm{NAD},7}}A^{n_{1,7}-1} + L_{0,7}\frac{c^C_{\mathrm{NAD}^+}}{K'_{\mathrm{NAD},7}}B^{n_{1,7}-1}}{A^{n_{1,7}} + L_{0,7}B^{n_{1,7}}} = \frac{(p_1 - c^C_{\mathrm{NADH}})\left(1 + k_3\left\{\frac{(1+k_4 c^C_{\mathrm{NADH}})}{\left[\left(\frac{k_5}{k_3}\right) + \left(\frac{k_2 k_5}{k_1 k_3}\right)c^C_{\mathrm{NADH}}\right]}\right\}^{n_{1,7}-1}\right)}{(k_1 + k_2 c^C_{\mathrm{NADH}})\left(1 + \left\{\frac{(1+k_4 c^C_{\mathrm{NADH}})}{\left[\left(\frac{k_5}{k_3}\right) + \left(\frac{k_2 k_5}{k_1 k_3}\right)c^C_{\mathrm{NADH}}\right]}\right\}^{n_{1,7}}\right)} \quad (17)$$

where we set $k_1 = (K_{\mathrm{NAD},7} + p_1)/r^{\max}_{\mathrm{RES},2}$, $k_2 = (K_{\mathrm{NAD},7} - K_{\mathrm{NADH},7})/K_{\mathrm{NADH},7}r^{\max}_{\mathrm{RES},2}$, $k_3 = (K_{\mathrm{NAD},7}/K'_{\mathrm{NAD},7})\quad L^{\frac{1}{n}}_{0,7}$, $k_4 = (K'_{\mathrm{NAD},7} - K'_{\mathrm{NAD},7})/K'_{\mathrm{NADH},7}(K'_{\mathrm{NAD},7} + p_1)$ and $k_5 = (K_{\mathrm{NAD},7} + p_1)/K'_{\mathrm{NAD},7} + p_1$. The original six kinetic parameters in this part of the rate equation can thus be replaced by the five independent ones, to k_1–k_5. The combinations of the original parameters that they represent are identifiable.

By using the Sedoglavic algorithm [14], it was confirmed that the assumption of conserved moieties of nicotinamide nucleotides was the only source of structural unidentifiability for the Rizzi model.

3.2 Hynne, Teusink and Lambeth models

The Hynne model, originally presented in [3], describes glycolytic oscillations in *S. cerevisiae* during cultivation in a continuous stirred tank reactor. Two conserved moieties are implicitly present in its stoichiometric matrix, one for the nicotinamide nucleotides and one for the adenine nucleotides. It is only the first that leads to unidentifiability. The rate expressions for GAPDH and lpGlyc can both be reparameterised in fewer parameters, as shown in the Supplementary material. In the rate expression for the AK reaction, all the variables from the adenine nucleotide moiety appear together, but its kinetic parameters are nevertheless identifiable.

The Teusink model [2] also describes yeast glycolysis, and just like the Rizzi model, it has two conserved moieties described by $c_{\mathrm{NADH}} + c_{\mathrm{NAD+}} = p_1$ and $c_{\mathrm{ATP}} + c_{\mathrm{ADP}} + c_{\mathrm{AMP}} = p_2$. There is no kinetic expression containing all three adenine nucleotides, and thus the second assumption does not lead to unidentifiability. However, the three dehydrogenase reactions GAPDH, ADH and G3PDH, which include both c_{NADH} and $c_{\mathrm{NAD+}}$ in their rate expressions, were all found to be unidentifiable. The calculation is carried out in the Supplementary material, where also the reparametrisation into identifiable parameters is given. For this model, there are two additional conservation relations because of equilibrium assumptions for [GraP] and [glyceronephosphate] and for the adenine nucleotides. The variables in the first equilibrium assumption, [GraP] and [glyceronephosphate], both appear in the ALD reaction rate expression that was therefore analysed for a potential source of unidentifiability. By applying the general method presented in the next section, it was found that the kinetic parameters in the ALD rate expression were identifiable even after introducing the equilibrium assumption.

The last model that has been tested describes muscle metabolism, and was originally presented by Lambeth *et al.* [23]. It contains the same conserved moieties as the previous models implicitly present in the stoichiometric matrix. The two dehydrogenase reactions included, GAPDH and LDH, are modelled by reversible kinetic expressions that include both $c_{\mathrm{NAD+}}$ and c_{NADH}, and they could potentially have the reported problem with unidentifiability. So could the PFK and the ADK reactions, as their rate expressions include all the variables from the adenine nucleotide moiety. However, by applying the general method, we were able to see that even after introducing the conserved moieties, all parameters in these rate expressions remained identifiable.

4 Unidentifiability as a result of conserved quantities: a linear algebra formalism

Rate expressions like the ones considered here are often rational functions of a number of concentrations. What actually happens in the examples discussed in the previous section is that the presence of a conserved quantity reduces the number of independent coefficients in the rational expression. It is possible to compute an upper bound of the number of independent (and thus identifiable) rate parameters in terms of, for example, the number of concentrations and the degree of the terms in which they appear in the rate expression. In this section, we formalise this procedure, giving a formula for determining cases of unidentifiability and showing how to obtain an identifiable set of parameters with a known relation to the original parameters. We also present a method for constructing the group of transformations of the original parameters that leaves the rate expression invariant.

4.1 Initial example revisited

To introduce the notation, we first demonstrate the procedure on the example of (5). Recall that there were three parameters, $V_{\max}$, K_{S} and K_{P}, and that these were collected in a vector $\boldsymbol{k}$

$$\boldsymbol{k} = \begin{bmatrix} V_{\max} & K_{\mathrm{S}} & K_{\mathrm{P}} \end{bmatrix}^{\mathrm{T}}$$

Introduce the following two symbols c_i, and the following four symbols a_{ij} and b_{ij}

$$c_1 = c_S, \ , \qquad c_2 = c_P, \qquad a_{10} = V_{max}\,,$$
$$b_{00} = K_S\,, \qquad b_{10} = 1\,, \qquad b_{01} = \frac{K_S}{K_P} \tag{18}$$

With this notation, (5) becomes

$$r = \frac{a_{10}c_1}{b_{00} + b_{10}c_1 + b_{01}c_2} \tag{19}$$

This is a rational expression, being the quotient of polynomials in the two variables c_1 and c_2. Collect the coefficients in a vector $\boldsymbol{a}$

$$\boldsymbol{a} = \begin{bmatrix} a_{10} & b_{00} & b_{10} & b_{01} \end{bmatrix}^T \tag{20}$$

Note that even if c_1 and c_2 were independent, we could only estimate $\boldsymbol{a}$ up to a multiplicative factor $\lambda_0 \in \mathbb{R}\backslash\{0\}$, since $r(\boldsymbol{a}) = r(\lambda_0\boldsymbol{a})$. Using (18), we could, of course, directly eliminate λ_0, but it is convenient to keep it, because it reveals the essentially linear structure of the coming transformations. At this point, we see that by inserting the relation $c_1 + c_2 = p$ in (19) we obtain

$$\tilde{r} = \frac{\tilde{a}_1 c_1}{\tilde{b}_0 + \tilde{b}_1 c_1} \tag{21}$$

where $\tilde{a}_1 = a_{10}$, $\tilde{b}_0 = b_{00} + pb_{01}$ and $\tilde{b}_1 = b_{10} - b_{01}$. This is a rational expression of only one variable, c_1, and with three coefficients. The relation between the coefficients in (19) and (21) is linear and is most naturally expressed by

$$\begin{bmatrix} \tilde{a}_1 \\ \tilde{b}_0 \\ \tilde{b}_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & p \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} a_{10} \\ b_{00} \\ b_{10} \\ b_{01} \end{bmatrix} \tag{22}$$

which can also be written as

$$\tilde{\boldsymbol{a}} = \boldsymbol{A}\boldsymbol{a}. \tag{23}$$

Note that the linearity of this transformation is not associated with a property of the underlying dynamical system, but only of the relation between the coefficients $\boldsymbol{a}$ and $\tilde{\boldsymbol{a}}$. There is therefore no approximation associated with this transformation, as is the case, for example, for a linearisation of a dynamical system.

The rank of $\boldsymbol{A}$ shows the number of coefficients $\tilde{\boldsymbol{a}}$ that are independently mapped from $\boldsymbol{a}$. In this example, $\boldsymbol{A}$ is of full (row)rank. However, there is still a common factor that needs to be removed. Let the identifiable parameters be denoted $\tilde{\boldsymbol{k}}$. By dividing all coefficients $\tilde{\boldsymbol{a}}$ by $\tilde{a}_1$, we obtain the following identifiable parameters

$$\tilde{\boldsymbol{k}} = \begin{bmatrix} \tilde{k}_1 & \tilde{k}_2 \end{bmatrix} = \begin{bmatrix} \dfrac{K_S(1 + (p/K_P))}{V_{max}} & \dfrac{(1 - (K_S/K_P))}{V_{max}} \end{bmatrix} \tag{24}$$

Note that this is the same result we obtained in (6) in Section 2. Finally, we now know that the conservation law has introduced an unidentifiability since the resulting number of parameters is less than the original number of parameters (i.e. $2 < 3$).

Equation (24) expresses $\tilde{\boldsymbol{k}}$ as a function of $\boldsymbol{k}$; we now turn to the opposite problem, i.e. expressing $\boldsymbol{k}$ as a function of $\tilde{\boldsymbol{k}}$. Let a superscript 0 denote a specific estimate, and consider a specific estimate of the identifiable parameters $\tilde{\boldsymbol{k}}^0$. The first step is from $\tilde{\boldsymbol{k}}^0$ to $\tilde{\boldsymbol{a}}^0$; in this example it is simply given by

$$\tilde{\boldsymbol{a}}^0 = \begin{bmatrix} \tilde{a}_1^0 & \tilde{b}_0^0 & \tilde{b}_1^0 \end{bmatrix}^T = \begin{bmatrix} \lambda_0 & \lambda_0\tilde{k}_1 & \lambda_0\tilde{k}_2 \end{bmatrix}^T \tag{25}$$

where $\lambda_0 \in \mathbb{R} \setminus \{0\}$ is an unknown scalar. Going from $\tilde{\boldsymbol{a}}^0$ to $\boldsymbol{a}^0$, we use that $\boldsymbol{A}$ is of full rank. This means that there is a vector $\boldsymbol{a}^0$, which satisfies (23), for all estimates $\tilde{\boldsymbol{a}}^0$. One solution is $\boldsymbol{a}_0^0 = \begin{bmatrix} \tilde{a}_1^0 & \tilde{b}_0^0 & \tilde{b}_1^0 & 0 \end{bmatrix}^T$; however, this is not unique because $\boldsymbol{A}$ has a one-dimensional null-space, $\ker(\boldsymbol{A})$, spanned by $\boldsymbol{w}_1 = \begin{bmatrix} 0 & -p & 1 & 1 \end{bmatrix}^T$. The freedom to choose a multiplicative factor λ_0 remains, and hence for any $\lambda_0, \lambda_1 \in \mathbb{R}$,

$$\boldsymbol{a} = \boldsymbol{a}(\lambda_0, \lambda_1) = \lambda_0\boldsymbol{a}_0^0 + \lambda_1\boldsymbol{w}_1 \tag{26}$$

yield a set of coefficients that is consistent with the estimates $\tilde{\boldsymbol{a}}^0$. To translate all the way to the original kinetic parameters $\boldsymbol{k}$, we need to solve (18). In this example, we can explicitly obtain

$$V_{max} = \lambda_0\tilde{a}_1^0 \tag{27a}$$
$$K_S = \lambda_0\tilde{b}_0^0 - \lambda_1 p \tag{27b}$$
$$1 = \lambda_0\tilde{b}_1^0 + \lambda_1 \tag{27c}$$
$$K_P = \frac{\lambda_0\tilde{b}_0^0 - \lambda_1 p}{\lambda_1}. \tag{27d}$$

Note that there is only one real degree of freedom since, for example, the third of these equations can be used to fix λ_0, i.e. $\lambda_0 = (1 - \lambda_1)/\tilde{b}_1^0$, which gives

$$V_{max} = (1 - \lambda_1)\tilde{a}_1^0/\tilde{b}_1^0, \quad K_S = (1 - \lambda_1)\tilde{b}_0^0/\tilde{b}_1^0 - \lambda_1 p$$
$$K_P = (1/\lambda_1 - 1)\tilde{b}_0^0/\tilde{b}_1^0 - p, \tag{28}$$

and finally

$$V_{max} = (1 - \lambda_1)/\tilde{k}_2, \quad K_S = ((1 - \lambda_1)\tilde{k}_1)/\tilde{k}_2 - \lambda_1 p$$
$$K_P = ((1/\lambda_1 - 1)\tilde{k}_1)/\tilde{k}_2 - p. \tag{29}$$

Let us now see how all these calculations can be generalised, following the same procedure as summarised in this diagram

(30)

4.2 Reparameterisation into identifiable parameters

Consider a reaction rate r depending on N concentrations $c_1, \ldots, c_N$ and m parameters $k_1, \ldots, k_m$, and assume that it can be written as a rational function,

$$r = \frac{P(c_1, \ldots, c_N; k_1, \ldots, k_m)}{Q(c_1, \ldots, c_N; k_1, \ldots, k_m)} \tag{31}$$

where P and Q are polynomials in the concentrations $c_1, \ldots, c_N$

$$P = \sum_{j=0}^{n_P} \sum_{r_1+\cdots+r_N=j} a_{r_1 \ldots r_N} c_1^{r_1}, \ldots, c_N^{r_N} \quad (32)$$

$$Q = \sum_{j=0}^{n_Q} \sum_{r_1+\cdots+r_N=j} b_{r_1 \ldots r_N} c_1^{r_1}, \ldots, c_N^{r_N} \quad (33)$$

A general polynomial P in three concentrations, c_1, c_2, and c_3, of second degree would with this notation be written

$$\begin{aligned} P = {} & a_{000} \\ & + a_{100}c_1 + a_{010}c_2 + a_{001}c_3 \\ & + a_{200}c_1^2 + a_{020}c_2^2 + a_{002}c_3^2 + a_{110}c_1c_2 \\ & + a_{101}c_1c_3 + a_{011}c_2c_3 \end{aligned} \quad (34)$$

Note that the sum of the indices of a coefficient $a_{r1}, \ldots, r_N$ is equal to the degree of the corresponding term, and that each coefficient can be a function of the m original kinetic parameters $k_1, \ldots, k_m$. We denote the transformation, which gives the coefficients $a_{r1}, \ldots, r_N$ and $b_{r1}, \ldots, r_N$ as functions of the parameters by Γ, i.e.

$$\Gamma(k_1, \ldots, k_m) = (\ldots, a_{r_1,\ldots,r_N}, \ldots, b_{r_1,\ldots,r_N}, \ldots)^T \quad (35)$$

so that Γ is a function from the m-dimensional parameter space to the M-dimensional space of rational coefficients. We assume here that Γ parameterises an m-dimensional manifold, because otherwise the parameters would be unidentifiable in the original rate expression.

If there is a constant linear combination of the concentrations

$$\sum \alpha_j c_j = p \quad (36)$$

one of the concentrations c_j can be eliminated from the rational expression. With no loss of generality, we may assume that the eliminated concentration is c_N. Replacing all occurrences of c_N in (31) with
$p - \sum_{j<N} \alpha_j c_j$ gives a new rate expression

$$\tilde{r}_X = \frac{\tilde{P}(c_1, \ldots, c_{N-1}; k_1, \ldots, k_m)}{\tilde{Q}(c_1, \ldots, c_{N-1}; k_1, \ldots, k_m)} \quad (37)$$

The new set of coefficients $\tilde{a}_{r_1 \ldots r_{N-1}}$ and $\tilde{b}_{r_1 \ldots r_{N-1}}$ are linear combinations of the coefficients $a_{r_1 \ldots r_N}$ and $b_{r_1 \ldots r_N}$, because each monomial containing c_N is transformed into a polynomial without c_N, which is added to those terms of P (or Q) that do not contain c_N. This relation can again be expressed as

$$\tilde{\boldsymbol{a}} = \boldsymbol{A}\boldsymbol{a} \quad (38)$$

If the matrix $\boldsymbol{A}$ is not of full rank, as in the r_{ADH} example from the Teusink model (cf. the Supplementary material), one can proceed by choosing a set of independent rows from $\boldsymbol{A}$.

It follows that the number of identifiable parameters is given by the rank of $(\boldsymbol{A} - 1)$, and we have the following formulas for calculating a lower bound on the number of unidentifiable parameters.

1. [dim $\boldsymbol{k}$] $-$[rank $\boldsymbol{A}$] $+ 1$ (e.g. $3 - 3 + 1 = 1$), or
2. [dim $\boldsymbol{k}$] $-$ [dim $\boldsymbol{a}$] $+$ [dim ker($\boldsymbol{A}$)] $+ 1$ (e.g. $3 - 4 + 1 + 1 = 1$).

The numbers within the parenthesis refer to the example in Section 1. Since a non-zero number of unidentifiable parameters implies an unidentifiable rate expression, the above formulas are a simple way of detecting unidentifiable rate expressions.

4.3 Back-translation and calculation of the symmetry families

The first step in the back-translation is from $\tilde{\boldsymbol{k}}$ to $\tilde{\boldsymbol{a}}$. Consider a specific estimate $\tilde{\boldsymbol{k}}^0$. Assume, without loss of generality, that it is the first coefficient $\tilde{a}_1$ that was used for the normalisation to the identifiable parameters. We then have the following back-translation formula

$$\boldsymbol{a}^0 = \begin{bmatrix} \tilde{a}_1^0 & a_2^0 & \ldots \end{bmatrix} = \lambda_0 \begin{bmatrix} 1 & \tilde{a}_2(\tilde{\boldsymbol{k}}^0) & \ldots \end{bmatrix} \quad (39)$$

where $\tilde{a}_i(\tilde{\boldsymbol{k}}^0)$ is equal to a single $\tilde{k}_j^0$ if a_i was one of the non-eliminated rows, and equal to the linear combinations mentioned above otherwise (these linear combinations only occur if $\boldsymbol{A}$ is of less than full rank).

The next step is to translate from $\tilde{\boldsymbol{a}}$ to $\boldsymbol{a}$. Let $J = \dim \ker(\boldsymbol{A})$ denotes the dimension of the null space of $\boldsymbol{A}$, and let $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_J$ be a set of vectors that span the null space. Let a specific solution of (23) be given by $\boldsymbol{a}^0$ (such a solution always exists). Then the set of all $\boldsymbol{a}$ that are consistent with $\boldsymbol{k}^0$ is spanned by $\lambda_0 \in \mathbb{R}\{0\}$ and $\lambda_i \in \mathbb{R}$ according to the following formula

$$\boldsymbol{a} = \lambda_0 \boldsymbol{a}^0 + \lambda_1 \boldsymbol{w}_1 + \cdots + \lambda_J \boldsymbol{w}_J \quad (40)$$

The final step in the back-translation is from $\boldsymbol{a}$ to $\boldsymbol{k}$. This is the reversal of the nonlinear mapping Γ, and it is therefore difficult to treat in the general case. Nevertheless, in many cases, explicit solutions should be available through symbolic software packages such as Mathematica and Maple. However, because of these difficulties, we now also present a geometrical interpretation of these results, and an alternative way of calculating the symmetry transformations, i.e. those transformations of $\boldsymbol{k}$ that leave the rate expression v invariant.

Recall that the function Γ parameterises an m-dimensional submanifold of $\mathbb{R}^M$ (which we denote by Γ as well). If we wish to determine whether a parameter set $\boldsymbol{k}^0 = (k_1^0, \ldots, k_m^0)$ is uniquely defined or whether there is a family of parameters $\boldsymbol{k}(s)$ that give the same rate expression, we compute the differential $\Gamma'(\boldsymbol{k}^0)$, and we let $\boldsymbol{u}_1(\boldsymbol{k}^0), \ldots, \boldsymbol{u}_j(\boldsymbol{k}^0)$ be a set of vectors so that $\Gamma'(\boldsymbol{k}^0)\boldsymbol{u}_1(\boldsymbol{k}^0), \ldots, \Gamma'(\boldsymbol{k}^0)\boldsymbol{u}_j(\boldsymbol{k}^0)$ span the intersection of the range of $\Gamma'(\boldsymbol{k}^0)$ and $[\{\lambda \boldsymbol{a}^0, \lambda \in \mathbb{R}\} \cup \ker(\boldsymbol{A})]$. The number of vectors $\boldsymbol{u}(\boldsymbol{k}^0)$ is the dimension of the set of equivalent parameter combinations. With a proper choice of the vectors $\mathbf{u}_i$ (in particular, they should be continuous as functions of $\boldsymbol{k}$), the family of transformations can be found by solving

$$\frac{\mathrm{d}\boldsymbol{k}(s)}{\mathrm{d}s} = \boldsymbol{u}_i[\boldsymbol{k}(s)]$$
$$\boldsymbol{k}(0) = \boldsymbol{k}^0$$

The original parameters that are identifiable correspond to those components that are zero in all vectors $\boldsymbol{u}_i$, $i = 1, \ldots, j$. This calculation is similar to the one in [11].

For the example shown in Section 4.1, the matrix Γ' is given by

$$\Gamma' = \begin{bmatrix} \frac{\partial a_{01}}{\partial V_{max}} & \frac{\partial a_{01}}{\partial K_S} & \frac{\partial a_{01}}{\partial K_P} \\ \frac{\partial b_{00}}{\partial V_{max}} & \frac{\partial b_{00}}{\partial K_S} & \frac{\partial b_{00}}{\partial K_P} \\ \frac{\partial b_{10}}{\partial V_{max}} & \frac{\partial b_{10}}{\partial K_S} & \frac{\partial b_{10}}{\partial K_P} \\ \frac{\partial b_{01}}{\partial V_{max}} & \frac{\partial b_{01}}{\partial K_S} & \frac{\partial b_{01}}{\partial K_P} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1/K_P & -K_S/K_P^2 \end{bmatrix} \tag{41}$$

and the degree of freedom for the choice of parameters is also given by the dimension of the intersection of the range of Γ' with span$\{\boldsymbol{a}_0^0, \boldsymbol{w}_1\}$. Again this gives a one-dimensional set, which is spanned by $\boldsymbol{w}_\cap = \boldsymbol{a}^0 - b_{10}\boldsymbol{w}_1$. The one-dimensional set of parameters that are consistent with a given set of measurements can be obtained by solving the ordinary differential equation

$$\frac{\mathrm{d}[V_{max}, K_S, K_P]^T}{\mathrm{d}s} = (\Gamma')^{-1}\boldsymbol{w}_\cap \tag{42}$$

where $\boldsymbol{w}_\cap$ depends on V_{max}, K_S and K_P, and $(\Gamma')^{-1}$ can be computed because $\boldsymbol{w}_\cap$ is in the range of Γ'.

The Supplementary material contains a Mathematica implementation of the calculations both for the identification of the identifiable parameters and for the calculation of the symmetry relations. The Mathematica implementation includes all the reaction rate expressions that we have discussed in this paper, but it is also easily extensible to other problems.

5 Discussion and conclusions

The problem with unidentifiability as a result of conserved moieties has been observed before. However, there does not seem to exist a systematic treatment of the problem, and the common explanation for the problem does not hold. This explanation is based on the observation that the insertion of conserved moieties in rate expressions usually leads to the situation of one parameter, or combination of parameters, in front of each term. This means that the coefficients in front of the terms are unidentifiable, but that does not necessarily imply unidentifiability of the original parameters. Consider for instance the following rate expression

$$r = \frac{p_1 c}{(p_2/p_1) + p_2 c}$$

with two parameters, p_1 and p_2, and one concentration, c. This rate expression has one coefficient in front of each term, but the parameters are nevertheless identifiable. It is therefore clear that the situation of coefficients in front of each term is not sufficient in itself to guarantee unidentifiability of the underlying parameters (i.e. additional conditions are necessary).

On the other hand, the intuitive method of using the coefficients as new identifiable parameters is a valid method. One coefficient must, however, be removed through division of both the numerator and denominator, and this method is therefore sometimes referred to as the 'divide through' method. The 'divide through' method is similar to our method proposed. However, the 'divide through' method does not involve an analysis of the transformation matrix $\boldsymbol{A}$, and this has some drawbacks. As is shown in the v_{ADH} example in the Supplementary material, the analysis of the linear dependencies in $\boldsymbol{A}$ allows for fewer identifiable parameters than the 'divide through' method would yield. Likewise, it is the analysis of the null-space of $\boldsymbol{A}$ that allows for easy formulas for the translation back to the original parameters [e.g. (27)]. There are also other methods that might be useful when trying to manually rewrite the expression in identifiable parameters, for instance, non-dimensionalisation [24]. However, non-dimensionalisation does not guarantee identifiability in the resulting expressions.

There are also some frameworks to deal with structural unidentifiability for the general situation. One such framework is differential algebra. However, many of its methods are not applicable to realistically large systems. One exception is the probabilistic algorithm by Sedoglavic [14]. Advantages with our method compared with Sedoglavic's algorithm is that our results are exact (i.e. not probabilistic), and that we provide a reparametrisation to identifiable parameters that may be expressed in the original parameters. A general advantage with our method compared with all methods based on differential algebra [11, 14] is that our method is built on a much simpler theory, something which, e.g. yields a more intuitive understanding of the origin of the unidentifiability.

Given all these options, it is also important to discuss how to proceed in different circumstances. In some cases, the problem with unidentifiability in the rate expressions may be left untreated altogether. This is the case if one, for example, is only interested in whether a given model structure is capable of explaining the data. One can also disregard identifiability issues if the parameter values are obtained elsewhere (e.g. from in vitro characterisations), and the model is used as a pure forward simulation model. Finally, if one is only interested in the value of the actual flux, these specific identifiability problems can be disregarded because the flux is identifiable even though some of the parameters describing it are not.

There are, on the other hand, also several scenarios where it is necessary to deal with problems of unidentifiability. If one would, for example, seek to compare the in vivo with the in vitro kinetics for a given enzyme, it is necessary to deal with these issues [18]. Such comparisons are important tools for the understanding of the general differences between in vivo and in vitro kinetics, and this is a central issue for the general understanding of life. Furthermore, if one wants to determine the quality of various model predictions, that is, how well they are characterised by the available data, it is also necessary to handle identifiability problems. For instance, if one wants to analyse the control coefficients in a model (i.e. the sensitivity of a specific model output with respect to perturbations in the model's parameters), then such an analysis will give different results depending on which value of the unidentifiability manifolds one chooses, even though all such choices lead to an identical agreement with the data. In such situations, it is therefore more advantageous to characterise the identifiable core in a model, and only consider the results with respect to this core [18, 25]. One way to find those parameter combinations that describe this core model is the intuitive 'divide through' method. However, as explained above our proposed method has advantages both in terms of yielding fewer parameters and in terms of back-translation of the result.

In conclusion, this article has presented a new framework to treat structural unidentifiability in single, rate expressions caused by conservation relations. The framework provides a valid explanation of the reasons for the problem, a

straightforward method for the detection of it, and a way to choose identifiable parameters if an unidentifiability should be detected. Furthermore, the analysis provides a translation from the identifiable parameters back to the original parameters. The ideas are illustrated by the analysis of some well-cited models of glycolysis, and the proposed method is easy to use via the provided Mathematica implementation.

6 Acknowledgment

We thank all involved reviewers for constructive criticism.

Financial support from the National Research School in Genomics and Bioinformatics, Göteborg University, Sweden is gratefully acknowledged. B. W. was partially supported by the Swedish Research Council and by the Swedish Foundation for Strategic Research via the Gothenburg Mathematical Modelling Center. G. C. carried out most of the work at Fraunhofer-Chalmers Centre for Industrial Mathematics, and was, during this time, supported by the BioSim Network of Excellence and by the Swedish Foundation for Strategic Research.

7 Supplementary data

Supplementary data are available on IET Systems Biology online.

8 References

1 Rizzi, M., Baltes, M., Theobald, U., and Reuss, M.: 'In vivo analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II. Mathematical Model', *Biotechnol. Bioeng.*, 1997, **55**, pp. 592–608

2 Teusink, B., Passarge, J., Reijenga, C.A., Esgalhado, E., van der Weijden, C.C., Schepper, M., Walsh, M.C., Bakker, B., van Dam, K., Westerhoff, H.V., and Snoep, J.L.: 'Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? testing biochemistry', *Eur. J. Biochem.*, 2000, **267**, (17), pp. 5313–5329

3 Hynne, F., Danø, S., and Sørensen, P.G.: 'Full-scale model of glycolysis in *Saccharomyces cerevisiae*', *Biophys. Chem.*, 2001, **94**, pp. 121–163

4 Stephanopoulos, G.N., Aristidou, A.A., and Nielsen, J.: 'Metabolic Engineering – principles and methodologies' (Academic Press, 1998)

5 Segel, I.H.: 'Enzyme kinetics' (John Wiley and Sons, New York, 1975)

6 Pohjanpalo, H.: 'System identifiability based on the power series expansion of the solution', *Math. Biosci.*, 1978, **41**, (1–2), pp. 21–33

7 Vajda, S., Godfrey, K.R., and Rabitz, H.: 'Similarity transformation approach to identifiability analysis of nonlinear comportemental models', *Math. Biosci.*, 1989, **93**, (2), pp. 217–248

8 Walter, E.D.: 'Identifiability of parametric models' (Pergamon Books Ltd, 1987)

9 Ollivier, F.: 'Le problème de l'identifiabilité structurelle globale: approche théorique, méthodes effectives et bornes de complexité', PhD thesis, École polytechnique, France, 1990

10 Diop, S., and Fliess, M.: 'On nonlinear observability'. Proc. First European Control Conf., Grenoble, France, July 1991, pp. 152–157

11 Chappell, M.J., Evans, N.D., Godfrey, K.R., and Chapman, M.J.: 'Structural identifiability of controlled state space systems: a quasi-automated methodology for generating identifiable reparameterisations of unidentifiable systems', *Symbolic Computation for Control (Ref. No. 1999/088), IEE Colloquium on*, 1999, pp. 8/1-812

12 Margaria, G., Riccomagno, E., Chappell, M.J., and Wynn, H.P.: 'Differential algebra methods for the study of the structural identifiability of rational polynomial state-space models in the biosciences', *Math. Biosci.*, 2001, **174**, pp. 1–26

13 Audoly, S., Bellu, G., D'Angio, L., Saccomani, M.P., and Cobelli, C.: 'Global identifiability of nonlinear models of biological systems', *IEEE Trans. Biomed. Eng.*, 2001, **48**, (1), pp. 55–65

14 Sedoglavic, A.: 'A probabilistic algorithm to test local algebraic observability in polynomial time', *J. Symbolic Comput.*, 2002, **33**, pp. 735–755

15 Anguelova, M.: 'Nonlinear observability and identifiability: general theory and a case study of a kinetic models for *S. cerevisiae*', Licentiate thesis, Chalmers University of Technology and Gothenburg University, Sweden, 2004

16 Vallabhajosyula, R., Chickarmane, V., and Sauro, H.: 'Conservation analysis of large biochemical networks', *Bioinformatics*, 2006, **22**, (3), pp. 346–353

17 Olivier, B.G., and Snoep, J.L.: 'Web-based kinetic modelling using JWS Online', *Bioinformatics*, 2004, **20**, pp. 2143–2144

18 Cedersund, G.: 'Core-box modelling – theorical contributions and applications to glucose homeostasis related systems'. PhD thesis, Chalmers University of Technology, Sweden, 2006

19 Higgins, J.: 'A chemical mechanism for oscillation of glycolytic intermediates in yeast cells', *Proc. Natl Acad. Sci. USA*, 1964, **51**, pp. 989–994

20 Sel'kov, E.E.: 'Self-oscillations in glycolysis. A simple kinetic model', *Eur. J. Biochem.*, 1968, **4**, pp. 79–86

21 Goldbeter, A.: 'Biochemical oscillations and cellular rhythms' (Cambridge University Press, Cambridge, UK, 1996)

22 Danø, S., Sorensen, P.G., and Hynne, F.: 'Sustained oscillations in living cells', *Nature*, 1999, **402**, (6759), pp. 320–322

23 Lambeth, M., and Kushmerick, M.J.: 'A computational model for glycogenolysis in skeletal muscle', *Ann. Biomed. Eng.*, 2002, **30**, pp. 808–827

24 Lin, C.C., and Segel, L.A.: 'Mathematics applied to deterministic problems in the natural sciences' (SIAM, 1988)

25 Cedersund, G., Danø, S., Sørensen, P.G., and Jirstrand, M.: 'From in vitro biochemistry to in vivo understanding of the glycolytic oscillations in *Saccharomyces cerevisiae*'. Proc. BioMedSim'2005, Linköping, Sweden, May 2005, pp. 105–114

Conservation laws and non-identifiability of rate expressions in biochemical models

Milena Anguelova[*,**], Gunnar Cedersund[§], Mikael Johansson[#], Carl Johan Franzén[#], Bernt Wennberg[*,**]

[*] Mathematical Sciences, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden

[**] Mathematical Sciences, Göteborg University, SE-412 96 Gothenburg, Sweden

[§] Department of Cell Biology, Linköping University, SE-581 85 Linköping, Sweden

[#] Chemical and Biological Engineering, Chemical Reaction Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden

Keywords
identifiability, structural identifiability, conserved moieties, rate expressions

Correspondence
Milena Anguelova
Address: Mathematical Sciences, Chalmers University of Technology, S-412 96 Gothenburg, Sweden
Email: milena@math.chalmers.se
Phone: +46-(0)31-7725323

SUPPLEMENTARY MATERIAL

A Detailed derivations from Section 3

A.1 The Hynne model

The Hynne model was originally presented in [3]. It consists of 22 states and 60 parameters. The model describes glycolytic oscillations in *Saccharomyces cerevisiae* during cultivation in a continuous stirred tank reactor. There are two compartments in the model, one for the extracellular metabolites and one for the intracellular ones. In the intracellular compartment (corresponding to the total cytosol volume of all the cells) there are two conserved moieties, one for the nicotinamide nucleotides and one for the adenine nucleotides

$$c_{\mathrm{NADH}} + c_{\mathrm{NAD^+}} = p_1 \tag{A1a}$$
$$c_{\mathrm{ATP}} + c_{\mathrm{ADP}} + c_{\mathrm{AMP}} = p_2 \tag{A1b}$$

These conserved moieties are not explicitly formulated in the model, but are implicitly present in the stoichiometric matrix.

There are three reactions that include the inter-conversion of the nicotinamides: glyceraldehyde-3-phosphate-dehydrogenase (GAPDH), a lumped reaction forming glycerol from dihydroxyacetonephosphate (lpGlyc), and the alcohol dehydrogenase (ADH). The ADH reaction is formulated as an irreversible reaction, and only c_{NADH} appears in the kinetic expression. However, both GAPDH and lpGlyc have c_{NADH} and $c_{\mathrm{NAD^+}}$ in their kinetic expressions which were therefore analysed. Both were found to contain unidentifiable parameters by the criterium presented in Section 4. The reaction rate expressions can be reparameterised in fewer identifiable parameters as follows:

The first rate expression is

$$r_{\mathrm{GAPDH}} = \frac{V_{8m}(c_{\mathrm{GAP}}c_{\mathrm{NAD^+}} - \frac{c_{\mathrm{BPG}}c_{\mathrm{NADH}}}{K_{8eq}})}{K_{8\mathrm{GAP}}K_{8\mathrm{NAD}}(1+\frac{c_{\mathrm{GAP}}}{K_{8\mathrm{GAP}}}+\frac{c_{\mathrm{BPG}}}{k_{8\mathrm{BPG}}})(1+\frac{c_{\mathrm{NAD+}}}{K_{8\mathrm{NAD}}}+\frac{c_{\mathrm{NADH}}}{K_{8\mathrm{NADH}}})} \quad , \tag{A2}$$

where there are three unidentifiable kinetic parameters, $V_{8m}, K_{8\mathrm{NAD}}$ and $K_{8\mathrm{NADH}}$, due to the assumption $c_{\mathrm{NADH}} + c_{\mathrm{NAD^+}} = p_1$. These three parameters can be replaced by two new ones k_1 and k_2 to obtain:

$$v_{\mathrm{GAPDH}} = \frac{c_{\mathrm{GAP}}(p - c_{\mathrm{NADH}}) - \frac{c_{\mathrm{BPG}}c_{\mathrm{NADH}}}{K_{8eq}}}{K_{8\mathrm{GAP}}(1+\frac{c_{\mathrm{GAP}}}{K_{8\mathrm{GAP}}}+\frac{c_{\mathrm{BPG}}}{k_{8\mathrm{BPG}}})(k_1 + k_2 c_{\mathrm{NADH}})} \quad , \tag{A3}$$

where $k_1 = \frac{K_{8\mathrm{NAD}}}{V_{8m}}(1+\frac{p_1}{K_{8\mathrm{NAD}}})$ and $k_2 = \frac{K_{8\mathrm{NAD}}}{V_{8m}}(\frac{1}{K_{8\mathrm{NADH}}} - \frac{1}{K_{8\mathrm{NAD}}})$.

The second reaction rate is

$$r_{\mathrm{lpGlyc}} = \frac{V_{15m}c_{\mathrm{DHAP}}}{K_{15\mathrm{DHAP}}\left(1+\frac{K_{15\mathrm{INADH}}}{c_{\mathrm{NADH}}}(1+\frac{c_{\mathrm{NAD^+}}}{K_{15\mathrm{INAD}}})\right) + c_{\mathrm{DHAP}}\left(1+\frac{K_{15\mathrm{NADH}}}{c_{\mathrm{NADH}}}(1+\frac{c_{\mathrm{NAD^+}}}{K_{15\mathrm{INAD}}})\right)} \quad , \tag{A4}$$

with five unidentifiable parameters $V_{15m}, K_{15\mathrm{DHAP}}, K_{15\mathrm{INADH}}, K_{15\mathrm{INAD}}$ and $K_{15\mathrm{NADH}}$. It can be rewritten as

$$r_{\mathrm{lpGlyc}} = \frac{c_{\mathrm{DHAP}}c_{\mathrm{NADH}}}{k_1 + k_2 c_{\mathrm{NADH}} + k_3 c_{\mathrm{DHAP}} + k_4 c_{\mathrm{DHAP}}c_{\mathrm{NADH}}} \quad , \tag{A5}$$

where

$$k_1 = \frac{K_{15\mathrm{DHAP}}K_{15\mathrm{INADH}}}{V_{15m}}(1 + \frac{p}{K_{15\mathrm{INAD}}}) \tag{A6a}$$

$$k_2 = \frac{K_{15\mathrm{DHAP}}}{V_{15m}}(1 - \frac{K_{15\mathrm{INADH}}}{K_{15\mathrm{INAD}}}) \tag{A6b}$$

$$k_3 = \frac{K_{15\mathrm{NADH}}}{V_{15m}}(1 + \frac{p}{K_{15\mathrm{INAD}}}) \tag{A6c}$$

$$k_4 = \frac{1}{V_{15\mathrm{m}}}(1 - \frac{K_{15\mathrm{NADH}}}{K_{15\mathrm{INAD}}}) \quad . \tag{A6d}$$

In the rate expression for the adenylate kinase (AK) reaction, the three variables from the second conserved moiety, c_{ATP}, c_{ADP} and c_{AMP} appear together, and AK has therefore been analysed for a potential source for unidentifiability. The rate expression is

$$r_{\mathrm{AK}} = k_{24f}c_{\mathrm{AMP}}c_{\mathrm{ATP}} - k_{24r}c_{\mathrm{ADP}}^2 \tag{A7}$$

which after insertion of $c_{\mathrm{AMP}} = p_2 - c_{\mathrm{ATP}} - c_{\mathrm{ADP}}$ becomes

$$r_{\mathrm{AK}} = k_{24f}(p_2 - c_{\mathrm{ATP}} - c_{\mathrm{ADP}})c_{\mathrm{ATP}} - k_{24r}c_{\mathrm{ADP}}^2 \tag{A8}$$

$$= k_{24f}p_2 - k_{24f}c_{\mathrm{ATP}}^2 - k_{24f}c_{\mathrm{ADP}}c_{\mathrm{ATP}} - k_{24r}c_{\mathrm{ADP}}^2 \quad . \tag{A9}$$

The two kinetic parameters k_{24f} and k_{24r} are thus identifiable.

The Sedoglavic algorithm showed in this case that the presence of the conserved moiety (A1a) was the only source of unidentifiability if all metabolites can be measured.

In the experimental setup used by Hynne *et al.* only NADH was measured with high time resolution. It is therefore interesting to note that even with $y = k[\mathrm{NADH}]$, the conservation law (A1a) is the only source that leads to structural unidentifiability among the reactions in the cytosol.

A.2 The Teusink model

The Teusink model [2] also describes yeast glycolysis, and just like the Rizzi and Hynne models, it has two conserved moieties described by (A1a) and (A1b). Again there is no kinetic expression containing all three adenine nucleotides. There are three dehydrogenase reactions: GAPDH, ADH, and glycerol-3-phosphate dehydrogenase (G3PDH). All these reactions involve the inter-conversion of the cofactors NAD^+ and NADH, and the corresponding rate equations contain both c_{NAD^+} and c_{NADH} in the denominator. For this model, there are two additional conservation relations due to equilibrium assumptions for [GraP] and [glyceronephosphate] and for the adenine nucleotides. The variables in the first equilibrium assumption, [GraP] and [glyceronephosphate], both appear in the ALD reaction rate expression which must therefore be analysed for a potential source of unidentifiability.

All three dehydrogenase reactions were found to be unidentifiable, due to the conserved moiety. Here we show this by a direct calculation, but the Mathematica code in part B of this supplementary material gives the same result (except that there are, of course, many ways of reparameterising the expressions in order to obtain an identifiable model). The rate equation for the first dehydrogenase reaction, GAPDH, is

$$r_{\mathrm{GAPDH}} = \frac{\frac{V_{\mathrm{maxf}} c_{\mathrm{GAP}} c_{\mathrm{NAD^+}}}{K_{\mathrm{GAP}} K_{\mathrm{NAD}}} - \frac{V_{\mathrm{maxr}} c_{\mathrm{BPG}} c_{\mathrm{NADH}}}{K_{\mathrm{BPG}} K_{\mathrm{NADH}}}}{(1 + \frac{c_{\mathrm{GAP}}}{K_{\mathrm{GAP}}} + \frac{c_{\mathrm{BPG}}}{K_{\mathrm{BPG}}})(1 + \frac{c_{\mathrm{NAD^+}}}{K_{\mathrm{NAD}}} + \frac{c_{\mathrm{NADH}}}{K_{\mathrm{NADH}}})} \quad . \tag{A10}$$

The constraint $c_{\mathrm{NADH}} + c_{\mathrm{NAD^+}} = const = p$ leads to the parameters V_{maxf}, V_{maxr}, K_{NAD} and K_{NADH} being unidentifiable. These four parameters can be replaced by three new ones to obtain the following re-parameterisation of the rate expression

$$r_{\mathrm{GAPDH}} = \frac{p c_{\mathrm{GAP}} - c_{\mathrm{GAP}} c_{\mathrm{NADH}} - \frac{k_1 c_{\mathrm{BPG}} c_{\mathrm{NADH}}}{K_{\mathrm{BPG}}}}{(1 + \frac{c_{\mathrm{GAP}}}{K_{\mathrm{GAP}}} + \frac{c_{\mathrm{BPG}}}{K_{\mathrm{BPG}}})(k_2 + k_3 c_{\mathrm{NADH}})} \tag{A11}$$

where

$$k_1 = \frac{K_{\mathrm{GAP}} K_{\mathrm{NAD}} V_{\mathrm{maxr}}}{V_{\mathrm{maxf}}} \tag{A12a}$$

$$k_2 = \frac{K_{\mathrm{GAP}} K_{\mathrm{NAD}} (1 + \frac{p}{K_{\mathrm{NAD}}})}{V_{\mathrm{maxf}}} \tag{A12b}$$

$$k_3 = (\frac{1}{K_{\mathrm{NADH}}} - \frac{1}{K_{\mathrm{NAD}}}) \frac{K_{\mathrm{GAP}} K_{\mathrm{NAD}}}{V_{\mathrm{maxf}}} \tag{A12c}$$

The rate equation for the second dehydrogenase reaction, G3PDH, is:

$$r_{\mathrm{G3PDH}} = \frac{V_{\mathrm{max}} (\frac{c_{\mathrm{DHAP}} c_{\mathrm{NADH}}}{K_{\mathrm{DHAP}} K_{\mathrm{NADH}}} - \frac{c_{\mathrm{Glycerol}} c_{\mathrm{NAD^+}}}{K_{\mathrm{DHAP}} K_{\mathrm{NADH}} K_{\mathrm{eq}}})}{(1 + \frac{c_{\mathrm{DHAP}}}{K_{\mathrm{DHAP}}} + \frac{c_{\mathrm{Glycerol}}}{K_{\mathrm{Glycerol}}})(1 + \frac{c_{\mathrm{NADH}}}{K_{\mathrm{NADH}}} + \frac{c_{\mathrm{NAD^+}}}{K_{\mathrm{NAD}}})} \tag{A13}$$

The unidentifiable parameters are $V_{\mathrm{max}}, K_{\mathrm{NAD}}$ and K_{NADH}. These three unidentifiable parameters can be replaced by two new ones to obtain the following re-parameterisation of the rate expression:

$$r_{\mathrm{G3PDH}} = \frac{\frac{c_{\mathrm{DHAP}} c_{\mathrm{NADH}}}{K_{\mathrm{DHAP}}} - \frac{c_{\mathrm{Glycerol}} (p - c_{\mathrm{NADH}})}{K_{\mathrm{DHAP}} K_{\mathrm{eq}}}}{(1 + \frac{c_{\mathrm{DHAP}}}{K_{\mathrm{DHAP}}} + \frac{c_{\mathrm{Glycerol}}}{K_{\mathrm{Glycerol}}})(k_1 + k_2 c_{\mathrm{NADH}})} \tag{A14}$$

where

$$k_1 = (1 + \frac{p}{K_{\mathrm{NAD}}}) \frac{K_{\mathrm{NADH}}}{V_{\mathrm{max}}} \tag{A15a}$$

$$k_2 = (\frac{1}{K_{\mathrm{NADH}}} - \frac{1}{K_{\mathrm{NAD}}}) \frac{K_{\mathrm{NADH}}}{V_{\mathrm{max}}} \tag{A15b}$$

The rate equation for the last dehydrogenase reaction, ADH, is:

$$r_{\mathrm{ADH,num}} = V_{\max}\left(\frac{c_{\mathrm{EtOH}}c_{\mathrm{NAD}}}{K_{\mathrm{EtOH}}K_{\mathrm{INAD}}} - \frac{c_{\mathrm{ACA}}c_{\mathrm{NADH}}}{K_{\mathrm{EtOH}}K_{\mathrm{INAD}}K_{\mathrm{eq}}}\right) \tag{A16}$$

$$\begin{aligned} r_{\mathrm{ADH,denom}} = {} & 1 + \frac{c_{\mathrm{NAD^+}}}{K_{\mathrm{INAD}}} + \frac{c_{\mathrm{EtOH}}K_{\mathrm{NAD}}}{K_{\mathrm{INAD}}K_{\mathrm{EtOH}}} + \frac{c_{\mathrm{ACA}}K_{\mathrm{NADH}}}{K_{\mathrm{INADH}}K_{\mathrm{ACALD}}} + \\ & \frac{c_{\mathrm{NADH}}}{K_{\mathrm{INADH}}} + \frac{c_{\mathrm{EtOH}}c_{\mathrm{NAD^+}}}{K_{\mathrm{INAD}}K_{\mathrm{EtOH}}} \\ & + \frac{c_{\mathrm{NAD^+}}c_{\mathrm{ACA}}K_{\mathrm{NADH}}}{K_{\mathrm{INAD}}K_{\mathrm{INADH}}K_{\mathrm{ACALD}}} + \frac{c_{\mathrm{EtOH}}c_{\mathrm{NADH}}K_{\mathrm{NAD}}}{K_{\mathrm{INAD}}K_{\mathrm{INADH}}K_{\mathrm{EtOH}}} + \\ & + \frac{c_{\mathrm{ACA}}c_{\mathrm{NADH}}}{K_{\mathrm{ACALD}}K_{\mathrm{INADH}}} + \frac{c_{\mathrm{EtOH}}c_{\mathrm{NAD^+}}c_{\mathrm{ACA}}}{K_{\mathrm{INAD}}K_{\mathrm{IACALD}}K_{\mathrm{EtOH}}} + \\ & \frac{c_{\mathrm{EtOH}}c_{\mathrm{ACA}}c_{\mathrm{NADH}}}{K_{\mathrm{IEtOH}}K_{\mathrm{INADH}}K_{\mathrm{ACALD}}} \end{aligned} \tag{A17}$$

$$r_{\mathrm{ADH}} = \frac{r_{\mathrm{ADH,num}}}{r_{\mathrm{ADH,denom}}} \tag{A18}$$

The unidentifiable parameters are $V_{\max}$, K_{EtOH}, K_{IEtOH}, K_{NAD}, K_{INAD}, K_{NADH}, K_{INADH}, K_{ACALD}, K_{IACALD}. These nine unidentifiable parameters can be replaced by eight new ones to obtain the following re-parameterisation of the rate expression:

$$r_{\mathrm{ADH,num}} = c_{\mathrm{EtOH}}(p - c_{\mathrm{NADH}}) - \frac{c_{\mathrm{ACA}}c_{\mathrm{NADH}}}{K_{\mathrm{eq}}} \tag{A19a}$$

$$\begin{aligned} r_{\mathrm{ADH,denom}} = {} & k_1 + k_2c_{\mathrm{EtOH}} + k_3c_{\mathrm{NADH}} + k_4c_{\mathrm{ACA}} + k_5c_{\mathrm{EtOH}}c_{\mathrm{NADH}} \\ & + k_6c_{\mathrm{EtOH}}c_{\mathrm{ACA}} + k_7c_{\mathrm{NADH}}c_{\mathrm{ACA}} + k_8c_{\mathrm{EtOH}}c_{\mathrm{NADH}}c_{\mathrm{ACA}} \end{aligned} \tag{A19b}$$

where

$$k_1 = \left(1 + \frac{p}{K_{\mathrm{INAD}}}\right)\frac{K_{\mathrm{EtOH}}K_{\mathrm{INAD}}}{V_{\max}} \tag{A20a}$$

$$k_2 = \left(\frac{K_{\mathrm{NAD}} + p}{K_{\mathrm{INAD}}K_{\mathrm{EtOH}}}\right)\frac{K_{\mathrm{EtOH}}K_{\mathrm{INAD}}}{V_{\max}} \tag{A20b}$$

$$k_3 = \left(\frac{1}{K_{\mathrm{INADH}}} - \frac{1}{K_{\mathrm{INAD}}}\right)\frac{K_{\mathrm{EtOH}}K_{\mathrm{INAD}}}{V_{\max}} \tag{A20c}$$

$$k_4 = \left(\frac{K_{\mathrm{NADH}}}{K_{\mathrm{INADH}}K_{\mathrm{ACALD}}} + \frac{pK_{\mathrm{NADH}}}{K_{\mathrm{INAD}}K_{\mathrm{INADH}}K_{\mathrm{ACALD}}}\right)\frac{K_{\mathrm{EtOH}}K_{\mathrm{INAD}}}{V_{\max}} \tag{A20d}$$

$$k_5 = \left(\frac{K_{\mathrm{NAD}}}{K_{\mathrm{INAD}}K_{\mathrm{INADH}}K_{\mathrm{EtOH}}} - \frac{1}{K_{\mathrm{INAD}}K_{\mathrm{EtOH}}}\right)\frac{K_{\mathrm{EtOH}}K_{\mathrm{INAD}}}{V_{\max}} \tag{A20e}$$

$$k_6 = \frac{p}{K_{\mathrm{IACALD}}V_{\max}} \tag{A20f}$$

$$k_7 = \left(\frac{1}{K_{\mathrm{INADH}}K_{\mathrm{ACALD}}} - \frac{K_{\mathrm{NADH}}}{K_{\mathrm{INAD}}K_{\mathrm{INADH}}K_{\mathrm{ACALD}}}\right)\frac{K_{\mathrm{EtOH}}K_{\mathrm{INAD}}}{V_{\max}} \tag{A20g}$$

$$k_8 = \left(\frac{1}{K_{\mathrm{IEtOH}}K_{\mathrm{INADH}}K_{\mathrm{ACALD}}} - \frac{1}{K_{\mathrm{INAD}}K_{\mathrm{IACALD}}K_{\mathrm{EtOH}}}\right)\frac{K_{\mathrm{EtOH}}K_{\mathrm{INAD}}}{V_{\max}} \tag{A20h}$$

The ALD rate expression containing the variables [GraP] and [glyceronephosphate] from one of the equillibrium assumptions was analysed using the Mathematica

implementation and found to be identifiable. The details can be seen by running the Mathematica code for the ALD rate expression.

B Application examples with the general method

B.1 The r_{ADH} reaction rate from the Teusink model

In this section we consider Teusink's rate expression for r_{ADH} in the framework of Section 4. This is the same reaction as in equation (A18). By use of Sedoglavic's algorithm we learned that nine of the ten parameters are probably not identifiable, whereas K_{eq} is, and hence it is essentially enough to study the denominator of the expression. Below we do not assume any knowledge about which of the parameters that are identifiable, but rather we consider the complete expression, and then have 10 parameters:

$$\begin{array}{lll} c_1 = \text{NAD}^+ & c_2 = \text{ACA} & c_3 = \text{NADH} \\ c_4 = \text{EtOH} & k_1 = V_{\max} & k_2 = K_{\text{INAD}} \\ k_3 = K_{\text{EtOH}} & k_4 = K_{\text{eq}} & k_5 = K_{\text{NADH}} \\ k_6 = K_{\text{INADH}} & k_7 = K_{\text{ACALD}} & k_8 = K_{\text{NAD}} \\ k_9 = K_{\text{IACALD}} & k_{10} = K_{\text{IEtOH}} & \end{array} \tag{B1}$$

With this notation we may write

$$r_{ADH} = \frac{\frac{k_1}{k_3k_2}c_1c_4 - \frac{k_1}{k_3k_2k_4}c_2c_3}{\begin{array}{l}1 + \frac{1}{k_2}c_1 + \frac{k_5}{k_6k_7}c_2 + \frac{1}{k_6}c_3 + \frac{k_8}{k_2k_3}c_4 + \frac{k_5}{k_2k_6k_7}c_1c_2 + \frac{1}{k_7k_6}c_2c_3 + \\ \qquad \frac{1}{k_2k_3}c_1c_4 + \frac{k_8}{k_2k_6k_3}c_3c_4 + \frac{1}{k_2k_9k_3}c_1c_2c_4 + \frac{1}{k_{10}k_6k_7}c_2c_3c_4\end{array}} \tag{B2}$$

We next adapt the notation to that of Section 4, and write

$$\begin{array}{lll} a_{1001} = \frac{k_1}{k_3k_2} & a_{0110} = -\frac{k_1}{k_3k_2k_4} & b_{0000} = 1 \\ b_{1000} = \frac{1}{k_2} & b_{0100} = \frac{k_5}{k_6k_7} & b_{0010} = \frac{1}{k_6} \\ b_{0001} = \frac{k_8}{k_2k_3} & b_{1100} = \frac{k_5}{k_2k_6k_7} & b_{1001} = \frac{1}{k_2k_3} \\ b_{0110} = \frac{1}{k_7k_6} & b_{0011} = \frac{k_8}{k_2k_6k_3} & b_{1101} = \frac{1}{k_2k_9k_3} \\ b_{0111} = \frac{1}{k_{10}k_6k_7} & & \end{array} \tag{B3}$$

to obtain

$$r_{ADH} = \frac{a_{1001}c_1c_4 + a_{0110}c_2c_3}{\begin{array}{l}b_{0000} + b_{1000}c_1 + b_{0100}c_2 + b_{0010}c_3 + b_{0001}c_4 + b_{1100}c_1c_2 + b_{0110}c_2c_3 + \\ \qquad b_{1001}c_1c_4 + b_{0011}c_3c_4 + b_{1101}c_1c_2c_4 + b_{0111}c_2c_3c_4\end{array}} \tag{B4}$$

Introducing the conserved quantity $c_1 = p - c_3$ gives

$$r_{ADH} = \frac{\tilde{a}_{001}c_4 + \tilde{a}_{110}c_2c_3 + \tilde{a}_{011}c_3c_4}{\begin{array}{l}\tilde{b}_{000} + \tilde{b}_{100}c_2 + \tilde{b}_{010}c_3 + \tilde{b}_{001}c_4 + \tilde{b}_{110}c_2c_3 \\ \qquad + \tilde{b}_{101}c_2c_4 + \tilde{b}_{011}c_3c_4 + \tilde{b}_{111}c_2c_3c_4\end{array}}. \tag{B5}$$

The new coefficients are obtained from the old ones by the transformation

$$\begin{bmatrix} \tilde{a}_{001} \\ \tilde{a}_{110} \\ \tilde{a}_{011} \\ \tilde{b}_{000} \\ \tilde{b}_{100} \\ \tilde{b}_{010} \\ \tilde{b}_{001} \\ \tilde{b}_{110} \\ \tilde{b}_{101} \\ \tilde{b}_{011} \\ \tilde{b}_{111} \end{bmatrix} = \begin{bmatrix} p & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & p & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & p & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & p & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} a_{1001} \\ a_{0110} \\ b_{0000} \\ b_{1000} \\ b_{0100} \\ b_{0010} \\ b_{0001} \\ b_{1100} \\ b_{1001} \\ b_{0110} \\ b_{0011} \\ b_{1101} \\ b_{0111} \end{bmatrix} \tag{B6}$$

We write $\tilde{\mathbf{a}} = A\mathbf{a}$ for the expression (B6), and as in Section 4, we denote by the superscript ^{0} an estimated set of coefficients. The rank of A is 10. The number of original kinetic parameters is 10. That means that the counting argument from Section 4

1. [no of parameters] − [rank A] +1 (i.e., $10 - 10 + 1 = 1$)

indicates that there is a degree of freedom in choosing the kinetic parameters.

Since the matrix A is not of full rank, we need to follow the procedure described for the general case in Section 4, if we want to obtain the identifiable parameters. The first row in A can be obtained by multiplying the third row by $-p$. Therefore: replace $\tilde{a}_{001}$ with $-p\tilde{a}_{011}$ and divide both the numerator and denominator with $\tilde{a}_{011}$ in equation (B5). Identifying the resulting coefficients in the denominator as the new identifiable parameters, gives an expression that is structurally identical to (A19).

We will end by giving a geometrical picture of the kinetic parameters giving identical reaction rates. Consider $\mathbf{k} = (k_1, ...k_{10})$. Each coefficient $\mathbf{a} \in \mathbb{R}^{13}$ is a function of $\mathbf{k}$, and we have

$$\mathbf{k} \quad \mapsto \quad \mathbf{a}(\mathbf{k}) \in \Gamma \subset \mathbb{R}^{13} \tag{B7}$$

where Γ typically would be a 10-dimensional sub-manifold of $\mathbb{R}^{13}$. The three dimensional null space of A is spanned by

$$\begin{array}{l} \mathbf{w}_1 = [\,0 \quad 0 \quad 0 \quad 0 \quad 0 \quad -p \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0\,]^{tr} \\ \mathbf{w}_2 = [\,0 \quad 0 \quad 0 \quad -p \quad 0 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0\,]^{tr} \\ \mathbf{w}_3 = [\,0 \quad 0 \quad -p \quad 1 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0\,]^{tr} \end{array} \tag{B8}$$

Hence, given $\mathbf{a}^0$ such that $\tilde{\mathbf{a}}^0 = A\mathbf{a}^0$, every $\mathbf{a}$ of the form

$$\mathbf{a} = \lambda_0\mathbf{a}^0 + \lambda_1\mathbf{w}_1 + \lambda_2\mathbf{w}_2 + \lambda_3\mathbf{w}_3 \tag{B9}$$

is also a coefficient vector that agrees with the estimated $\tilde{\mathbf{a}}^0$ (where λ_0 is the degree of freedom due to the coefficients being determined only up to a constant

factor). This is a four-dimensional subset of $\mathbb{R}^{13}$. The set of kinetic parameters that agree with the estimated $\tilde{\mathbf{a}}^0$ is given by

$$K = \left\{ \mathbf{k} \in \mathbb{R}^{10} \quad \text{such that} \quad \mathbf{a}(\mathbf{k}) \in \text{span}(\mathbf{a}^0, \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3) \right\} \tag{B10}$$

and if $\mathbf{k} \mapsto \mathbf{a}(\mathbf{k})$ is an injective map, K is at least one-dimensional.

B.2 The RES2-reaction in the Rizzi model

The RES2-reaction rate from the Rizzi model (equation (13, 14) involves the concentrations of NAD^+, NADH and GAP. However, for the purpose of this discussion, it is enough to consider the following part of the expression, involving only NAD and NADH:

$$r^{max}_{\text{RES},2} \frac{\frac{1}{K_{\text{NAD},7}} A^{n_{1,7}-1} + L_{0,7} \frac{1}{K'_{\text{NAD},7}} B^{n_{1,7}-1}}{A^{n_{1,7}} + L_{0,7} B^{n_{1,7}}} \tag{B11}$$

where

$$\begin{aligned} A &= 1 + \frac{c^C_{\text{NAD}^+}}{K_{\text{NAD},7}} + \frac{c^C_{\text{NADH}}}{K_{\text{NADH},7}} \\ B &= 1 + \frac{c^C_{\text{NAD}^+}}{K'_{\text{NAD},7}} + \frac{c^C_{\text{NADH}}}{K'_{\text{NADH},7}} \end{aligned} \tag{B12}$$

This can be expressed as (we set $n_{1,7} = \eta$ for notational convenience, and for the argument here, we assume that η is a constant)

$$\frac{\gamma_1 \left(b_{00} + b_{10}c_1 + b_{01}c_2\right)^{\eta-1} + \gamma_2 \left(a_{00} + a_{10}c_1 + a_{01}c_2\right)^{\eta-1}}{(b_{00} + b_{10}c_1 + b_{01}c_2)^{\eta} + \gamma_3 \left(a_{00} + a_{10}c_1 + a_{01}c_2\right)^{\eta}}, \tag{B13}$$

where

$$\begin{array}{lll} c_1 = c^C_{NAD^+} & c_2 = c^C_{\text{NADH}} & \\ \gamma_1 = r^{\max}_{\text{RES},2}/K_{\text{NAD},7} & \gamma_2 = r^{\max}_{\text{RES},2} L_{0,7}/K'_{\text{NAD},7} & \gamma_3 = L_{0,7}\,, \\ a_{00} = 1 & a_{10} = 1/K_{\text{NAD},7} & a_{01} = 1/K_{\text{NADH},7} \\ b_{00} = 1 & b_{10} = 1/K'_{\text{NAD},7} & b_{01} = 1/K'_{\text{NADH},7} \end{array} \tag{B14}$$

While a rational expression is left invariant by the multiplication of all coefficients by the same constant, this does not hold here.

Replacing c_2 by $p - c_1$ gives the new expression

$$\begin{aligned} &\frac{\gamma_1 \left(\tilde{b}_0 + \tilde{b}_1 c_1\right)^{\eta-1} + \gamma_2 \left(\tilde{a}_0 + \tilde{a}_1 c_1\right)^{\eta-1}}{(\tilde{b}_0 + \tilde{b}_1 c_1)^{\eta} + \gamma_3 \left(\tilde{a}_0 + \tilde{a}_1 c_1\right)^{\eta}} \\ &= \frac{\bar{\gamma}_1 \left(1 + \bar{b}_1 c_1\right)^{\eta-1} + \bar{\gamma}_2 \left(1 + \bar{a}_1 c_1\right)^{\eta-1}}{(1 + \bar{b}_1 c_1)^{\eta} + \bar{\gamma}_3 \left(1 + \bar{a}_1 c_1\right)^{\eta}}, \end{aligned} \tag{B15}$$

where

$$\begin{array}{lll} \bar{b}_1 = \dfrac{b_{10} - b_{01}}{1 + p b_{01}} & & \bar{a}_1 = \dfrac{a_{10} - a_{01}}{1 + p a_{01}} \\ \bar{\gamma}_3 = \gamma_3 \left(\dfrac{1 + p a_{01}}{1 + p b_{01}}\right)^{\eta} & \bar{\gamma}_1 = \dfrac{\gamma_1}{1 + p b_{01}} & \bar{\gamma}_2 = \gamma_2 \dfrac{(1 + p a_{01})^{\eta-1}}{(1 + p b_{01})^{\eta}} \end{array} \tag{B16}$$

In fact, the actual form of the coefficients may not be very important, but it is interesting to see that there are only five of them, and hence the six physical parameters cannot be determined uniquely; the expression is not identifiable. This example cannot be as easily treated with the Mathematica code as the other ones.

In the Segel example in Section 4, the linear transformation of the coefficients have a one-dimensional null-space, which can then be identified with the degree of freedom in choosing physical parameters. Here the null-space of the transformation

$$(\gamma_1, \gamma_2, \gamma_3, a_{10}, a_{01}, b_{10}, b_{01}) \rightarrow (\bar{\gamma}_1, \bar{\gamma}_2, \bar{\gamma}_3, \bar{a}_1, \bar{b}_1) \tag{B17}$$

has a two-dimensional null-space. There is, however, still just a one-dimensional set of kinetic parameters that give identical rate values. This set can be geometrically understood as the intersection of the two-dimensional null-space to (B17) with the six-dimensional image of the physical parameters in the space of coefficients $(\gamma_1, \gamma_2, \gamma_3, a_{10}, a_{01}, b_{10}, b_{01})$.

C A brief description of the Mathematica implementation

Although the mathematics used in this paper is elementary, the calculations soon become cumbersome, and some software for symbolic calculation may be helpful. We have created a very simple Mathematica-notebook of the computations described in this paper.

The implementation consists of three notebooks, which should be run sequentially:

- `symmBerData.nb`
- `symmBerStart.nb`
- `symmBer.nb`

Before starting, the notebook `symmBerData.nb` should be edited to contain information about the actual rate equation, as shown in Figure 1. The provided file contains data for all the rate expressions discussed in the article, and it should be easy to see how to modify for any rational rate expression. A little more work is needed to deal with meromorphic rate expressions.

For reasonably large expressions, all calculations are carried out symbolically, and in this case one can normally execute `symmBerStart` without any modifications. There are some parameters that can be changed, in particular with respect to how much of the intermediary results one wishes to see.

Very complicated rate expressions cannot be treated directly without a very large amount of memory in the computer. Hence part of the calculation can be carried out in a semi-numerical fashion. This is much faster, but in that case the present code can only verify that an expression is identifiable. The parameters that decide whether a full symbolic calculation should be attemted or not, can also be set in `symmBerStart`, together with special parameters related to the semi-numerical computaiton.

After executing the notebook `symmBer.nb`, the result is presented as shown in Figure 3.

A remark is in its place here: in terms of which parameters are identifiable, and the degree of freedom in choosing the unidentifable ones, the results obtained with the Mathematica code agree with the calculations done by hand and presented in the main part of the paper, and in Supplement B. However, due to the implentation, the suggested reparametrisations are not necessarily the same, nor are the corresponding matrices A.

The code works well in the cases we have tried, but of course we cannot guarantee that there are no cases where it would fail. There is certainly much space for improvement of the code, both in terms of efficiency and readability.

The Mathematica notbooks are also accessible from the web-page
`http://www.math.chalmers.se/~wennberg/Code`
We will attempt to keep that code up to date. Also, new versions of the code will be available from that web page.

Data for the Teusink model -- GAPDH-reaction

k1= Vmaxf k2=Kgap k3=Knad k4=Vmaxr k5=Kbpg
k6=Knadh

c1 = Cgap c2=Cnad c3=Cnadh c4=Cbpg

In[2032]:=

```
parLista = {k1, k2, k3, k4, k5, k6}
varLista = {c1, c2, c3, c4}
consMoi = {c2 → p - c3}
va = ( (k1 / (k2 k3)) c1 c2 - (k4 / (k6 k5)) c3 c4 ) /
  ((1 + c1 / k2 + c4 / k5) (1 + c2 / k3 + c3 / k6))
```

Out[2032]= {k1, k2, k3, k4, k5, k6}

Out[2033]= {c1, c2, c3, c4}

Out[2034]= {c2 → -c3 + p}

Out[2035]= $$\frac{\frac{c1\,c2\,k1}{k2\,k3} - \frac{c3\,c4\,k4}{k5\,k6}}{\left(1 + \frac{c1}{k2} + \frac{c4}{k5}\right)\left(1 + \frac{c2}{k3} + \frac{c3}{k6}\right)}$$

Figure 1: Data input to Mathematica code for analysing the identifiable of a rate expression.

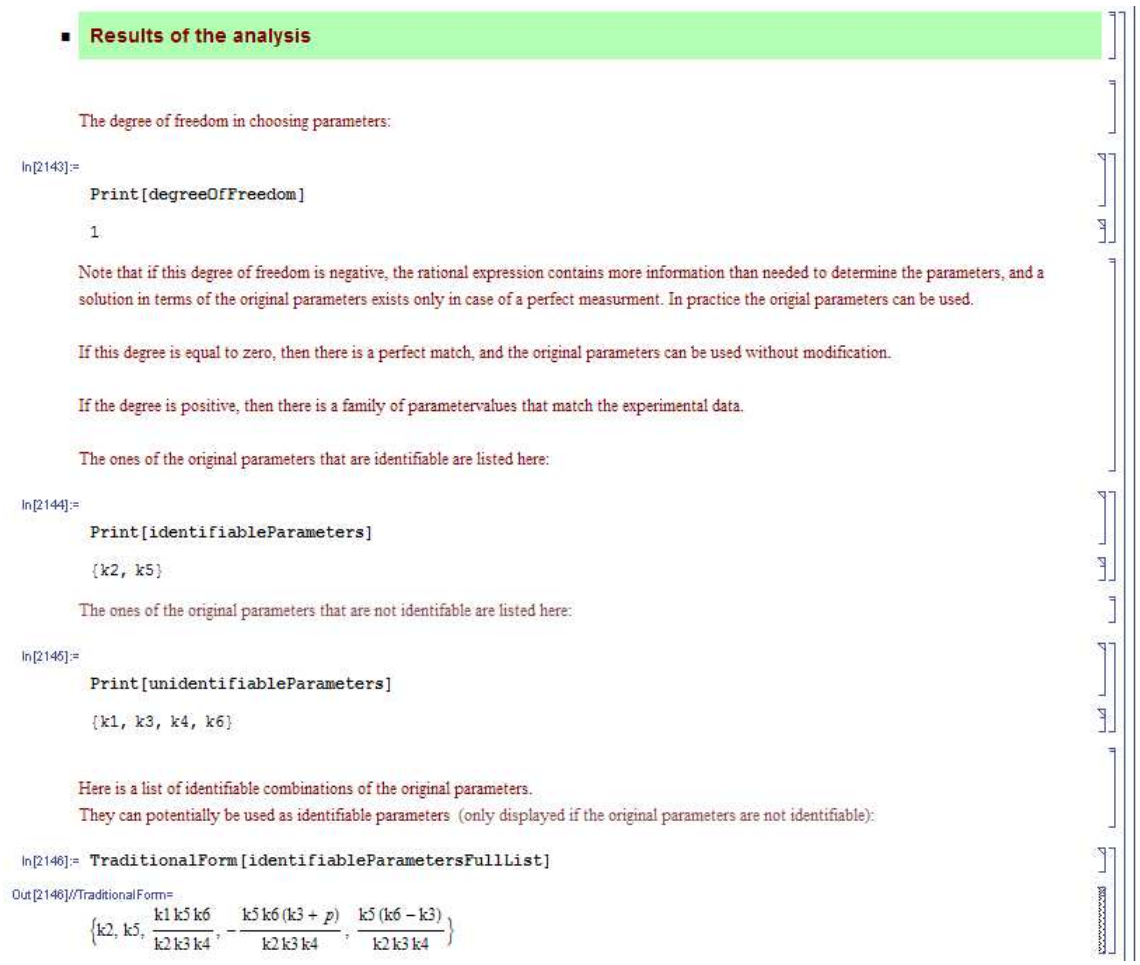


Figure 2: Result from Mathematica code for analysing the identifiable of a rate expression, part1.

Finally, in cases when there is a degree of freedom greater than zero, we present a table with parameters and the vector fields that generate the symmetries. Note that the identifiable parameters are the ones for which (all) the corresponding field components are zero .

```
In[2147]:= If[ degreeOfFreedom > 0,
             Print[TableForm[Transpose[simplifiedFields], TableHeadings → { parLista, {}}]] ]
```

k1	k1 k3
k2	0
k3	k3 (k3 + p)
k4	k4 k6
k5	0
k6	$k6^2 + k6\,p$

Here we present a reparameterisation that attempts to keep the original parameters as much as possible. This depends on an explicit solution of an system of nonlinear equations, and may fail to give useful results. N.B. This calculation is not completely reliable. It depends on the possibility for *Mathematica* to solve for the old parameters in terms on the new parameters. If the result looks strange, look in the calculation above, some manual touch may be needed.

Here is a list of the new parameters.

```
In[2148]:=
If[degreeOfFreedom > 0, Print[nyParLista]]
```

```
{kt1, kt2, kt3, kt4, kt5}
```

The next output describes the relation between the old and new parameters

```
In[2149]:=
If[degreeOfFreedom > 0, TraditionalForm[newToOldSubs]]
```

Out[2149]//TraditionalForm=

$$\left\{k2 \to kt1,\ k5 \to kt2,\ k1 \to \frac{k6\,kt3 + p\,kt3}{k6\,kt5 - kt4},\ k4 \to -\frac{k5\,(k6 + p)}{k2\,(kt4 + kt5\,p)},\ k3 \to \frac{k6\,kt4 + k6\,kt5\,p}{kt4 - k6\,kt5}\right\}$$

This is the rate expression expressed in the new parameters. Note that one of the original parameters, remains in the expression, and can be chosen arbitrarily. Another comment is that

```
In[2150]:= TraditionalForm[symmetryParameter]
           TraditionalForm[newRateExpression]
```

Out[2150]//TraditionalForm=

$$\{k6\}$$

Out[2151]//TraditionalForm=

$$\frac{(c3\,c4\,k5\,kt1 - c1\,c2\,k2\,kt2\,kt3)\,(k6 + p)}{k2\,(c4\,kt1 + (c1 + kt1)\,kt2)\,(c2\,(kt4 - k6\,kt5) + (c3 + k6)\,(kt4 + kt5\,p))}$$

Figure 3: Result from Mathematica code for analysing the identifiable of a rate expression, part2.

Part II

Observability and identifiability for nonlinear systems of delay-differential equations with discrete time-delays

Milena Anguelova

Abstract

The properties of observability and identifiability for nonlinear systems of delay-differential equations with discrete time-delays have been analysed.

The property of weak observability for nonlinear delay systems with known discrete time-delays has previously been characterised in the setting of modules over noncommutative rings. We show that the observability problem can be reduced to the analysis of the Jacobian of a system of algebraic equations, allowing for the application of existing computation algorithms for rank calculation.

New results are presented on state elimination for delay systems and characterisation of the identifiability of time-lag parameters. We show the existence of an input-output representation for systems with multiple discrete time-delays. The form of the input-output equations is shown to decide the identifiability of the time-lag parameters. Their identifiability is not directly related to the well-characterized identifiability/observability of the other model parameters/state variables and an independent analysis must be performed. The values of the time lags can be found directly from the input-output equations, if these can be obtained explicitly. Linear-algebraic criteria are formulated to decide the identifiability of the delay parameters which eliminate the need for explicit computation of the input-output relations. The criteria are applied in the analysis of biological models from the literature.

Keywords: Delay systems, nonlinear systems, time-delay model, observability, identifiability, time delay, state elimination, input-output representation, signalling pathways.

NOTATION

$C = C([-r, 0], \mathbb{R}^n)$	the space of continuous functions mapping the interval[-r,0] into $\mathbb{R}^n$ with the topology of uniform convergence
$x \in \mathbb{R}^n$	state variables
$u \in \mathbb{R}^m$	input/control variables
$y \in \mathbb{R}^p$	output
f	a function describing the time-derivative of the state variables
h	a function relating the output to the state variables
$h^{(j)}$	the j-th time-derivative of h along the system dynamics
φ	a function of initial conditions

Systems with a single time delay

τ	time delay
δ	time-shift operator corresponding to τ
$\mathcal{K}$	the field of meromorphic functions of a finite number of variables from $\{x(t-k\tau), u(t-k\tau), \ldots, u^{(l)}(t-k\tau), \quad k, l \in \mathbb{Z}^+\}$
$\phi(\delta, x, u)$	a meromorphic function of a finite number of variables from $\{x(t-k\tau), u(t-k\tau), \ldots, u^{(l)}(t-k\tau), \quad k, l \in \mathbb{Z}^+\}$
$\mathcal{K}(\delta]$	the noncommutative ring of polynomials in δ with coefficients from $\mathcal{K}$
$\mathcal{K}\langle\delta)$	the fraction field for $\mathcal{K}(\delta]$
$\mathcal{M}$	$\text{span}_{\mathcal{K}(\delta]}\{d\xi : \xi \in \mathcal{K}\}$

Systems with multiple time delays

$\boldsymbol{\tau}$	a vector containing all time-delays, $(\tau_1, \ldots, \tau_\ell)$
δ_i	time-shift operator corresponding to τ_i
$\boldsymbol{i\tau}$	$i_1\tau_1 + \cdots + i_\ell\tau_\ell$
$\mathcal{K}$	be the field of meromorphic functions of a finite number of variables from $x(t-\boldsymbol{i\tau}), u(t-\boldsymbol{i\tau}), \ldots, u^{(l)}(t-\boldsymbol{i\tau}), \quad , \boldsymbol{i} = (i_1, \ldots, i_\ell), \quad i_j, l \in \mathbb{Z}^+$
$\mathcal{K}(\boldsymbol{\delta}]$	the noncommutative ring of polynomials in δ_i with coefficients from $\mathcal{K}$
$\mathcal{K}\langle\boldsymbol{\delta})$	the fraction field for $\mathcal{K}(\boldsymbol{\delta}]$
$\mathcal{M}$	$\text{span}_{\mathcal{K}(\boldsymbol{\delta}]}\{d\xi : \xi \in \mathcal{K}\}$
$x(t-\boldsymbol{i\tau})$	$x(t-(i_1\tau_1 + \cdots + i_\ell\tau_\ell))$
$x_{[i]}(t)$	all the variables of type $x(t-\sum_{j=1}^{\ell} i_j\tau_j), \quad i_j \in \mathbb{Z}^+, \quad \sum_{j=1}^{\ell} i_j \leq i$

Contents

1 Introduction to time-delay systems

Often, in modelling the dynamics of physical and biological processes, the future behaviour of a system is assumed to depend only on its present state. In such cases, a deterministic model of the system can consist of ordinary or partial differential equations, where knowledge of the state at a discrete time-point determines the future state of the system. It can be the case, however, that a more realistic model must include the influence of the past. To exemplify, we will use the modelling of the JAK-STAT signalling pathway in the cell from the paper by Timmer et al. (2004).

The JAK-STAT signalling pathway is one of the signalling pathways in which signals from cell surface receptors binding the hormone erythropoietin (Epo) are transduced to the nucleus where the respective genes are activated, see (Timmer et al., 2004) and the references therein. When Epo is bound to the extracellular part of the receptor, Janus kinase (JAK) at the intracellular, cytoplasmic domain of the receptor is activated by phosphorylation. This, in its turn, leads to the phosphorylation of monomeric STAT-5, a member of the signal transduction- and activator of transcription family of transcription factors. When monomeric STAT-5 is phosphorylated, it forms dimers which migrate into the nucleus. There they bind to the promotor region of the DNA and initiate gene transcription. In the established view, the active role of STAT-5 ends in the nucleus and it is dedimerised, dephosphorylated and exported to the cytoplasm for degradation. An alternative suggestion is that after being dedimerised and dephosphorylated in the nucleus, it reenters the cytoplasm and is involved into another round of activation. In mathematical terms, the first alternative is modelled by a system of ordinary differential equations and is a so-called feed-forward cascade. The second alternative involves a feedback allowing for nuclear-cytoplasmic cycling of STAT-5. Without the possibility to measure the components of STAT-5 in the nucleus, one effective way to model this cycling is to include a time-delay in the state component for STAT-5 corresponding to the time spent by STAT-5 in the nucleus and this model displayed a much better fit to measurement data than the established feed-forward cascade model. Let us look at the two models in detail.

Denoting the amount of activated Epo-receptors by $EpoR_A$, unphosphorylated monomeric STAT-5 by x_1, phosphorylated monomeric STAT-5 by x_2, phosphorylated dimeric STAT-5 in the cytoplasm by x_3 and phosphorylated

dimeric STAT-5 in the nucleus by x_4, the feed-forward cascade model is:

$$\begin{cases} \dot{x}_1(t) &= -k_1 x_1(t) EpoR_A(t) \\ \dot{x}_2(t) &= k_1 x_1(t) EpoR_A - k_2 x_2^2(t) \\ \dot{x}_3(t) &= -k_3 x_3(t) + k_2 x_2^2(t) \\ \dot{x}_4(t) &= k_3 x_3(t) \quad . \end{cases} \tag{1.1}$$

In the STAT-5 cycling model, equations 1 and 4 are replaced by

$$\dot{x}_1(t) = -k_1 x_1(t) EpoR_A(t) + 2k_4 x_3(t-\tau) \tag{1.2}$$

$$\dot{x}_4(t) = k_3 x_3(t) - k_4 x_3(t-\tau) \quad , \tag{1.3}$$

to reflect the assumption that the rate of change in unphosphorylated monomeric STAT-5 in the cytoplasm depends also on the reentering of STAT-5 from the nucleus after it has been dedimerised and dephosphorylated. The latter process is assumed to take time τ and so it is the amount of phosphorylated dimeric STAT-5 in the nucleus at time $t-\tau$ which will be added to the unphosphorylated monomeric STAT-5 in the cytoplasm at time t. It is clear that in this model, the state of the system at a given moment depends also on its past state, which then renders it a *time-delay system*.

Time-delay systems have been used to model a wide range of phenomena in systems biology, including metabolic insulin signaling (Sedaghat et al., 2002), phosphorylation-dephosphorylation cycles (Srividhya, 2007), the *lac* operon (Mahaffy and Savev, 1999), the circadian pacemaker (Lema et al., 2000; Smolen et al., 1999), gene expression in cultured mammalian cells (Monk, 2003) and in zebra fish (Lewis, 2003). These delay models usually contain a number of state variables and parameters with values that are often unknown and can only be found by estimation from experimental data, a prerequisite for which are the observability and identifiability properties. The characterisation of these for time-delay systems is the subject of this second part of the thesis.

2 Basic properties of functional differential equations

Time-delay systems, or systems of delay differential equations, belong to a larger class of functional differential equations (FDE). In this section, we describe some basic properties of this class of equations, following (Hale, 1977) and the overview paper by Richard (2003).

2.1 Retarded functional differential equations

If the rate of change of the state of a system involves the current as well as the past values of the state, one obtains a so-called system of retarded functional differential equations. Before describing their general form, we consider a simple example.

2.1.1 A simple example

Consider the following simple delay equation, taken from Chapter 0 in (Diekmann et al., 1995):

$$\dot{x} = -\pi/2 x(t-1) \quad . \tag{2.1}$$

Both $sin(\pi/2(t+1/2))$ and $cos(\pi/2(t+1/2))$ satisfy this equation and they coincide at $t = 0$. This means that specifying an initial condition only at $t = 0$ does not specify the solution of the system, unlike for ODE-systems. Instead, one has to specify an *initial function* on an interval of length 1.

2.1.2 General form

Let $r \geq 0$ and $C = C([-r, 0], \mathbb{R}^n)$ be the space of continuous functions mapping the interval $[-r, 0]$ into $\mathbb{R}^n$ with the topology of uniform convergence. Designate the norm of an element φ in C by $|\varphi| = \sup_{-r \leq \theta \leq 0} |\varphi(\theta)|$.

If $\sigma \in \mathbb{R}, A \geq 0$ and $x \in C([\sigma - r, 0 + A], \mathbb{R}^n)$, then for any $t \in [\sigma, \sigma + A]$, define *the state* $x_t \in C$ by

$$x_t(\theta) = x(t + \theta), \quad -r \leq \theta \leq 0. \tag{2.2}$$

If D is a subset of $\mathbb{R} \times C$ and f is a function, $f : \mathbb{R} \times C \to \mathbb{R}^n$, then the following equation is called a *retarded functional differential equation (RFDE)* on D:

$$\dot{x}(t) = f(t, x_t) \quad . \tag{2.3}$$

A function x is a *solution* of the above equation on $[\sigma - r, 0 + A)$ if there are $\sigma \in \mathbb{R}$ and $A > 0$ such that $x \in C([\sigma - r, 0 + A), \mathbb{R}^n)$, $(t, x_t) \in D$ and

$x(t)$ satisfies (2.3) for $t \in [\sigma, \sigma + A)$. For given $\sigma \in \mathbb{R}, \varphi \in C$, $x(\sigma, \varphi, f)$ is a *solution of* (2.3) *with initial value* φ *at* σ or simply a *solution through* (σ, φ) if there is an $A > 0$ such that $x(\sigma, \varphi, f)$ is a solution of (2.3) on $[\sigma - r, \sigma + A)$ and $x_\sigma(\sigma, \varphi, f) = \varphi$.

The above equation form includes ODE systems ($r = 0$) as well as differential-difference equations

$$\dot{x} = f(t, x(t), x(t - \tau_1(t)), \ldots, x(t - \tau_\ell(t))) \quad , \quad 0 \leq \tau_i(t) \leq r, i = 1, \ldots, \ell$$

and integro-differential equations

$$\dot{x} = \int_{-r}^{0} g(t, \theta, x(t + \theta)) d\theta \quad , \quad 0 \leq \tau_i(t) \leq r, i = 1, \ldots, \ell \quad .$$

Systems of differential-difference equations with constant (discrete) time-delays τ_i are our main interest in this thesis.

2.1.3 Existence, uniqueness and backward continuation

2.1.3.1 Existence of solution

If Ω is an open subset of $\mathbb{R} \times C$ and f is continuous, $f \in C(\Omega, \mathbb{R}^n$, then there is a solution of(2.3) through $(\sigma, \varphi) \in \Omega$.

2.1.3.2 Uniqueness

If Ω is an open subset of $\mathbb{R} \times C$, $f \in C(\Omega, \mathbb{R}^n$ and $f(t, \varphi)$ is Lipschitzian in φ in each compact set in Ω, then there is a unique solution of(2.3) through $(\sigma, \varphi) \in \Omega$.

2.1.3.3 Backward continuation

We now discuss the question of the existence of the solution to the left of the initial time-point, the so-called *backward continuation of a solution.* This is ensured provided that $f(t, \varphi)$ satisfies the so-called *atomicity* property, defined as follows: f is atomic at β on an open set $\Omega \subseteq \mathbb{R} \times C$ if it is continuous together with its first and second Fréchet derivatives w.r.t. φ, and f_φ, the derivative w.r.t. φ is atomic at β on Ω.

If $f : \Omega \to \mathbb{R}^n$ is atomic at $-r$ on Ω, and there is an $\alpha, 0 < \alpha < r$, such that $\dot{\varphi}(\theta)$ is continuous for $\theta \in [-\alpha, 0]$, then there is an $\bar{\alpha} > 0$ and a unique solution of (2.3) on $[\sigma - r - \bar{\alpha}, \sigma]$ through $(\sigma, \varphi) \in \Omega$.

2.2 Neutral functional differential equations

Neutral functional differential equations (NFDE) involve the same highest derivation order for some components of $x(t)$ at both time t and past time(s) $t' < t$. They can be written

$$\dot{x}(t) = f(t, x_t, \dot{x}_t) \quad , \tag{2.4}$$

or, in the form of (Hale, 1977),

$$\frac{d}{dt} D(t, x_t) = f(t, x_t) \quad , \tag{2.5}$$

where D is continuous and atomic at zero and f is continuous. As shown in (Hale, 1977), difference equations are also NFDE.

For a given open set $\Omega \subseteq \mathbb{R} \times C$, there exists a solution to (2.5) through $(\sigma, \varphi) \in \Omega$. If f is Lipschitzian in φ on compact sets of Ω, then this solution is unique.

3 Mathematical framework for control - noncommutative algebra

This section presents a mathematical framework for the treatment of control problems such as observability and identifiability for time-delay systems, more specifically, for systems of differential-difference equations with constant (discrete) time-delays. It is an algebraic framework based on modules over noncommutative rings which was introduced by Moog, Castro-Linares, Velasco-Villa and Márquez-Martínez (2000), and developed by Márquez-Martínez, Moog and Velasco-Villa (2000) and Xia, Márquez, Zagalak and Moog (2002). Some of the theory in these papers, previously formulated only for systems with commensurate time-delays, will now be generalised to multiple (noncommensurate) delays in Subsection 3.2.

3.1 Systems with a single time-delay

This section contains a review of previously published algebraic framework and results found in (Cohn, 1985), (Lam, 1999), (Moog et al., 2000), (Márquez-Martínez et al., 2000) and (Xia et al., 2002) adapted to the present case of unknown single time-delay.

3.1.1 The ring $\mathcal{K}(\delta]$

Consider nonlinear time-delay systems with a single constant time-delay of the form:

$$\left\{\begin{array}{lcl} \dot{x}(t) &=& f\big(x(t), x(t-\tau), u, u(t-\tau)\big) \\ y(t) &=& h\big(x(t), x(t-\tau)\big) \\ x(t) &=& \varphi(t), \quad t \in [-\tau, 0] \\ u(t) &=& u_0(t), \quad t \in [-T, 0] \quad , \end{array}\right. \tag{3.1}$$

where $x \in \mathbb{R}^n$ denotes the state variables, $u \in \mathbb{R}^m$ is the input and $y \in \mathbb{R}^p$ is the output. Any parameters for the system can be written as state variables with time-derivative zero. The unknown constant time-delay is denoted by $\tau \in [0, T)$, $T \in \mathbb{R}$. The entries of f and h are meromorphic in their arguments (quotients of convergent power series with real coefficients) and $\varphi : [-\tau, 0] \to \mathbb{R}^n$ is an unknown continuous function of initial conditions. The set of initial functions for the variables x is denoted by $C := C([-\tau, 0], \mathbb{R}^n)$. A meromorphic function u is called an admissible input if the differential equation above admits a unique solution. The set of all such input functions is denoted by C_U.

Let $\mathcal{K}$ be the field of meromorphic functions of a finite number of variables from $\{x(t-k\tau), u(t-k\tau), \ldots, u^{(l)}(t-k\tau), \quad k,l \in \mathbb{Z}^+\}$. Such functions will also be denoted $\phi(\delta, x, u)$.

Let $\mathcal{E}$ be the vector space over $\mathcal{K}$ given by:

$$\mathcal{E} = \mathrm{span}_{\mathcal{K}}\{d\xi : \xi \in \mathcal{K}\} \quad . \tag{3.2}$$

Then $\mathcal{E}$ is the set of linear combinations of a finite number of one-forms from $dx(t-k\tau), du^{(l)}(t-k\tau)$ with row vector coefficients in $\mathcal{K}$. The time-shift operator δ is defined by:

$$\delta(\xi(t)) = \xi(t-\tau), \quad \xi(t) \in \mathcal{K} \quad , \tag{3.3}$$

and

$$\delta(\alpha(t)d\xi(t)) = \alpha(t-\tau)d\xi(t-\tau), \quad \alpha(t)d\xi(t) \in \mathcal{E} \quad . \tag{3.4}$$

Let $\mathcal{K}(\delta]$ denote the set of polynomials of the form

$$a(\delta] = a_0(t) + a_1(t)\delta + \cdots + a_{r_a}(t)\delta^{r_a} \quad , \tag{3.5}$$

where $a_j(t) \in \mathcal{K}$. Addition in the noncommutative ring $\mathcal{K}(\delta]$ is defined as usual while multiplication is given by

$$a(\delta]b(\delta] = \sum_{k=0}^{r_a+r_b} \sum_{i+j=k}^{i\le r_a, j\le r_b} a_i(t)b_j(t-i\tau)\delta^k \quad . \tag{3.6}$$

The ring $\mathcal{K}(\delta]$ is a (left) integral domain by its definition. It is also *Noetherian* which will be shown next, using the suggestion in (Xia et al., 2002) to follow the proof of the Hilbert Basis Theorem (for example that in Chapter 1 of (Goodearl et al., 2004)).

Proposition 3.1 *(Theorem 1 in (Xia et al., 2002))* $\mathcal{K}(\delta]$ *is a Noetherian ring.*

Proof: Let I be a nonzero left ideal of $\mathcal{K}(\delta]$. Let J be the set of leading coefficients of the elements of I, together with 0:

$$J = \{r \in \mathcal{K} : r\delta^d + r_{d-1}\delta^{d-1} + \cdots + r_0 \in I, \quad r_{d-1}, \ldots, r_0 \in \mathcal{K}\} \quad . \tag{3.7}$$

We will now verify that $J = \mathcal{K}$. Clearly, $0 \in J$. For any nonzero $r \in J$, there exists $p \in I$, such that p has leading coefficient r. Then $r^{-1}p$ is an element of I with leading coefficient 1, and thus $1 \in J$ and $J = \mathcal{K}$.

Consider the nonzero polynomials in I with least degree n and choose one with leading coefficient 1, denoted by p. Set $I_0 = \mathcal{K}(\delta]p$. We will prove

that $I = I_0$. First we will show that I_0 contains all elements of I of degree n. Let w be an element of I with degree n and let s be its leading coefficient. Then, $sp - w \in I$ and if $sp - w$ is nonzero, it has degree less than n which is a contradiction. Thus, $sp - w = 0$ and $w \in I_0$. Assume that I_0 contains all elements of I with degree less than m for some $m \geq n$. Let u be any element of I with degree $m \geq n$ and let r be its leading coefficient which can be assumed to be 1 w.l.o.g. Let $v = \delta^{m-n}p$, then $v \in I_0$ and v has degree m and leading coefficient 1. Now $u - v \in I$ and $u - v$ has degree less than m. Thus, $u - v \in I_0$ and therefore $u \in I_0$. Hence, I_0 contains all elements of I with degree less than $m + 1$. By the principle of induction, $I = I_0$ and I is thus a finitely generated left ideal of $\mathcal{K}(\delta]$.

Therefore, $\mathcal{K}(\delta]$ is a Noetherian ring. ■

Alternatively, we can prove the following more general statement which will be used later for the multiple-delay case:

Proposition 3.2 *If R is a left Noetherian ring, so is $R(\delta]$.*

Proof: The proof is exactly the same as for the usual polynomial ring case, see for instance (Cox et al., 1997). Nevertheless, we recall it here for completeness:

For $f \in R(\delta]$, if $f = \sum_{k=0}^{n} a_k \delta^k$ with a_n not equal to 0, then the degree of f, $\deg f$ is n and a_n is its leading coefficient. Let I be an ideal in $R(\delta]$ and construct a sequence $f_1, f_2, \ldots$ of elements of I such that f_{i+1} has minimal degree among elements of $I \setminus J_i$, where J_i is the ideal generated by $f_1, \ldots, f_i$. Let a_i be the leading coefficient of f_i and consider the chain of ideals

$$(a_1) \subset (a_1, a_2) \subset (a_1, a_2, a_3) \subset \ldots$$

contained in R. Since R is Noetherian, this chain must terminate for some N. We will show that $I = (f_1, \ldots, f_N)$. For if not then, $f_{N+1} \in I \setminus (f_0, f_1, \ldots, f_N)$ and $a_{N+1} = \sum_{i=1}^{N} u_i a_i$ for some $u_1, \ldots, u_N \in R$. Consider $g = \sum_{i=1}^{N} u_i f_i \delta^{n_i}$ where $n_i = \deg f_N - \deg f_i$. Then $\deg g = \deg f_{N+1}$ and their leading coefficients agree, and so $f_{N+1} - g$ has degree strictly less than $\deg f_{N+1}$ and $f_{N+1} - g \in I$, $f_{N+1} - g \notin (f_0, f_1, \ldots, f_N)$, contradicting the choice of f_{N+1}. Thus I is finitely generated. Since I was an arbitrary ideal in $R(\delta]$, every ideal in $R[\delta)$ is finitely generated and $R(\delta]$ is therefore Noetherian. ■

Another important property of $\mathcal{K}(\delta]$ is that it is a *left Ore domain*, defined as follows (p.38 in (Cohn, 1985)):

Definition 3.1 *A non-zero integral domain R is called a left Ore domain if it satisfies the left Ore condition: $Ra \cap Rb \neq 0$ for all nonzero $a, b \in R$.*

This property of $\mathcal{K}(\delta]$ can be checked directly (see (Xia et al., 2002)), but it also follows from it being a left Noetherian integral domain, see Corollary 8.10 in Chapter 0 of (Cohn, 1985). Let us briefly describe the argument used there: a left integral domain R that is not a left Ore domain must contain two non-zero elements a, b that are left incommensurable: $Ra \cap Rb = 0$. Thus, the left ideal $Ra + Rb$ is a free left R-module of rank 2. The elements $ba^n (n = 0, 1, 2, \ldots$ are left linearly independent, for if $\sum c_i b a^i = 0$, then by factoring out a we obtain $c_0 b + c_1 ba + \cdots + c_r ba^r = 0 (c_0 \neq 0$ and thus $c_0 b \in Ra \cap Rb$, which is a contradiction. Thus, we obtain free left ideals of R of countable rank which are not finitely generated (by Lemma 3.1)and R is therefore not Noetherian.

3.1.2 The fraction field $\mathcal{K}\langle\delta)$

Since $\mathcal{K}(\delta]$ is a left Ore domain, it can be embedded in a *left field of fractions* (Corollary 8.7 in (Cohn, 1985)), that will be denoted $\mathcal{K}\langle\delta)$. Elements of $\mathcal{K}\langle\delta)$ are denoted $b^{-1}(\delta]a(\delta]$.

To define the operations of addition and multiplication in $\mathcal{K}\langle\delta)$, (Xia et al., 2002), let $S = \{(a, b) : a, b \in \mathcal{K}(\delta], b \neq 0$ and define an equivalence relation $\sim$ on S by setting $(a, b) \sim (c, d)$ if $b'a = d'c$ and $b'b = d'd$. Let $b^{-1}a$ denote the equivalence class associated to (a, b) , $\mathcal{K}\langle\delta)$ is then the set of equivalence classes in S.

Addition is defined by

$$b^{-1}a + d^{-1}c = (b'b)^{-1}(b'a + d'c) \quad , \tag{3.8}$$

where $b'b = d'd$ and multiplication is defined by

$$b^{-1}ad^{-1}c = (a'b)^{-1}(d'c) \quad , \tag{3.9}$$

where $a'a = d'd$.

3.1.3 Rank and basis of modules over $\mathcal{K}(\delta]$

The fact that $\mathcal{K}(\delta]$ is a noncommutative ring makes it necessary to elaborate on the definitions of rank and basis of modules over it.

For any ring R, an R-module N is said to be *free* iff it has a basis, i.e. a set $\{e_i : i \in I\} \subset N$ such that any element of N is a unique finite (left) linear combination of the $e_i : s$. A basis is a minimal generating set and a maximal linearly-independent set. The rank of N is the cardinal of the basis, $|I|$. The following lemma implies that the rank of a module, if infinite, is uniquely determined:

Lemma 3.1 *(Generation Lemma, p.2 in (Lam, 1999)) Let R be a ring and N any R-module. Let $\{e_i : i \in I\} \subset N$ be a minimal generating set of N where the cardinality $|I|$ is infinite. Then N cannot be generated by fewer than $|I|$ elements.*

Thus, a free module with a basis of infinite cardinality cannot be finitely generated.

Let ${}^{X}R$ be the direct sum of copies of R indexed by X. Then we have the isomorphism ${}^{X}R \to N$ which maps $(a_x) \mapsto \sum xa_x$. We will say that ${}^{n}R$ has unique rank if it is not isomorphic to ${}^{m}R$ for any $m \neq n$. If N is a free module of unique rank r then we write $\mathrm{rank}_R N = r$.

For a finitely generated free module the rank need not be unique, a trivial counter example is the zero ring for which ${}^{m}0 =^{n} 0$ for all m, n.

A ring R is said to have the *invariant basis property or the invariant basis number (IBN)* if every free left R-module has unique rank. Any non-trivial Noetherian ring has the IBN (Propositions (1.8) and (1.13) in Lam (1999)) and therefore, by Corollary 3.1 so does the ring $\mathcal{K}(\delta]$.

We define the following left module over $\mathcal{K}(\delta]$:

$$\mathcal{M} = \mathrm{span}_{\mathcal{K}(\delta]}\{d\xi : \xi \in \mathcal{K}\} \quad , \tag{3.10}$$

where the elements of the ring $\mathcal{K}(\delta]$ act on elements $d\xi$ of the module according to (3.4). The module $\mathcal{M}$ contains the same elements as $\mathcal{E}$. All bases for a free submodule $\mathcal{N}$ of $\mathcal{M}$ have the same cardinality due to the IBN property of $\mathcal{K}(\delta]$.

Due to the fact that direct inversion of elements of $\mathcal{K}(\delta]$ is not possible, we need to use the concept of closure introduced by Conte and Perdon (1984) and generalized to the present case in Xia et al. (2002). The closure of a submodule $\mathcal{N}$ in $\mathcal{M}$ is the submodule

$$\overline{\mathcal{N}} = \{w \in \mathcal{M} : \exists a(\delta] \in \mathcal{K}(\delta], \quad a(\delta] \neq 0, \quad a(\delta]w \in \mathcal{N}\} \quad . \tag{3.11}$$

For example, $dx(t) \notin \mathrm{span}_{\mathcal{K}(\delta]}\{dx(t-\tau)\}$ but $dx(t) \in \overline{\mathrm{span}_{\mathcal{K}(\delta]}\{dx(t-\tau)\}}$ since $\delta(dx(t)) = dx(t-\tau) \in \mathrm{span}_{\mathcal{K}(\delta]}\{dx(t-\tau)\}$.

If $\mathcal{N}$ coincides with $\overline{\mathcal{N}}$, then $\mathcal{N}$ is said to be closed in $\mathcal{M}$. The following result on closure and closedness of modules from Xia et al. (2002) (Lemma 2) will be needed later:

Proposition 3.3 *(Lemma 2 in Xia et al. (2002))*

1. *The closure $\overline{\mathcal{N}}$ of $\mathcal{N}$ in $\mathcal{M}$ is the smallest closed submodule of $\mathcal{M}$ containing $\mathcal{N}$;*

2. *For any finitely generated submodule* $\mathcal{N}$ *of* $\mathcal{M}$*, one has* $\text{rank}_{\mathcal{K}(\delta]}\mathcal{N} = \text{rank}_{\mathcal{K}(\delta]}\overline{\mathcal{N}}$*;*

3. *For a free submodule* $\mathcal{N}$ *of* $\mathcal{M}$*,* $\overline{\mathcal{N}}$ *is the largest submodule of* $\mathcal{M}$ *containing* $\mathcal{N}$ *and having a rank equal to* $\text{rank}_{\mathcal{K}(\delta]}\mathcal{N}$*.*

Proof: We supply our own proof for this proposition.

1. Let $\mathcal{P}$ be the smallest closed submodule of $\mathcal{M}$ containing $\mathcal{N}$. Choose an element $u \in \mathcal{N} \subseteq \mathcal{P}$. Since $\mathcal{P}$ is closed, it contains $\mathcal{K}(\delta]u$. This is true for all $u \in \mathcal{N}$ and thus $\overline{\mathcal{N}} = \mathcal{K}(\delta]\mathcal{N} \subseteq \mathcal{P}$.

2. Let $\mathcal{N}$ be generated by $w_1, \ldots, w_s$. Choose an element $u \in \overline{\mathcal{N}}$. Then $\exists a(\delta] \in \mathcal{K}(\delta]$ such that $a(\delta]u \in \mathcal{N}$ and thus $a(\delta]u = b_1 w_1 + \cdots + b_s w_s$ for some $b_i \in \mathcal{K}(\delta]$. Thus, u is linearly dependent on $w_1, \ldots, w_s$. This is true for all elements in $\overline{\mathcal{N}}$ and thus $\overline{\mathcal{N}}$ has the same rank as $\mathcal{N}$.

3. Let $\mathcal{P}$ be the largest submodule of $\mathcal{M}$ containing $\mathcal{N}$ and having a rank equal to $\text{rank}_{\mathcal{K}(\delta]}\mathcal{N}$. Let w_I be a basis for $\mathcal{N}$, where I is an index set. Choose an element $u \in \mathcal{P}$, then u must be linearly dependent on w_I and $\exists a(\delta], b_i \in \mathcal{K}(\delta]$ such that $a(\delta]u = \sum_{i \in I} b_i w_i$. Thus, $u \in \overline{\mathcal{N}}$ and therefore $\mathcal{P} \subseteq \mathcal{N}$. ■

3.1.4 The Poincaré lemma

An element w of $\mathcal{M}$, also called a 1-form, is said to be exact if there exists a function $\phi \in \mathcal{K}$ such that $\omega = d\phi$. Since any 1-form $\omega \in \mathcal{M}$ is also an element of $\mathcal{E}$, the following form of Poincaré's Lemma holds (Lemma 3 in Márquez-Martínez et al. (2000)):

Lemma 3.2 (Poincaré) *(Lemma 3 in Márquez-Martínez et al. (2000)) Consider a 1-form* $\omega \in \mathcal{M}$*. Then there exists a function* $\xi(t) \in \mathcal{K}$ *such that (locally)* $\omega = d\xi(t)$ *if and only if* $d\omega = 0$*.*

3.1.5 Differentiation of functions and one-forms

Differentiation with respect to time along the system dynamics for functions $\phi(x(t-i\tau), u(t-j\tau), \ldots, u^{(l)}(t-j\tau))$, $0 \leq i, j \leq k$, $l \geq 0$ in $\mathcal{K}$ and one-forms $\omega = \sum_i \kappa_x^i dx(t-i\tau) + \sum_{ij} \nu_i du^{(j)}(t-i\tau)$ in $\mathcal{M}$ is defined in the natural way (see for example Xia et al. (2002) and Zhang et al. (2006)):

$$\dot{\phi} = \sum_{i=0}^{k} \frac{\partial \phi}{\partial x(t-i\tau)} \delta^i f + \sum_{r=0}^{l} \sum_{j=0}^{k} \frac{\partial \phi}{\partial u^{(r)}(t-j\tau)} u^{(r+1)}(t-j\tau) \quad (3.12)$$

$$\begin{aligned}\dot{\omega} &= \sum_i \dot{\kappa}_x^i dx(t-i\tau) + \sum_{ij} \dot{\nu}_i du^{(j)}(t-i\tau) + \sum_i \kappa_x^i d\delta^i f + \\ &\quad + \sum_{ij} \nu_i du^{(j+1)}(t-i\tau) \quad . \end{aligned} \tag{3.13}$$

3.2 Systems with multiple time-delays

In this subsection we review the generalisation of the mathematical framework introduced above to time-delay systems with multiple time-delays as first introduced by Moog et al. (2000), and developed by Márquez-Martínez et al. (2000). We extend the results of Xia et al. (2002) from the previous subsection to the present case of (not necessarily commensurate) multiple time-delays.

3.2.1 The ring $\mathcal{K}(\boldsymbol{\delta}]$

The general form of the nonlinear time-delay systems considered is:

$$\left\{ \begin{array}{lcl} \dot{x}(t) &=& f\big(x(t), x(t-\tau_1), \ldots, x(t-\tau_\ell), u, u(t-\tau_1), \ldots, u(t-\tau_\ell)\big) \\ y(t) &=& h\big(x(t), x(t-\tau_1), \ldots, x(t-\tau_\ell)\big) \\ x(t) &=& \varphi(t), \quad t \in [-\max_i \tau_i, 0] \\ u(t) &=& u_0(t), \quad t \in [-T, 0] \quad , \end{array} \right. \tag{3.14}$$

where $x \in \mathbb{R}^n$ denotes the state variables, $u \in \mathbb{R}^m$ is the input and $y \in \mathbb{R}^p$ is the output. The unknown constant time-delays are denoted by the vector $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_\ell)$, $\tau_i \in [0, T)$, $\quad T \in \mathbb{R}$. The entries of f and h are meromorphic in their arguments and $\varphi : [-\max_i(\tau_i), 0] \to \mathbb{R}^n$ is an unknown continuous function of initial conditions. The set of initial functions for the variables x is denoted by $C := C([-\max_i(\tau_i), 0], \mathbb{R}^n)$. A meromorphic input function $u(t)$ is called an admissible input if the differential equation above admits a unique solution. The set of all such input functions is denoted by C_U.

Let $\mathcal{K}$ be the field of meromorphic functions of a finite number of variables from $\{x(t - \boldsymbol{i\tau}), u(t - \boldsymbol{i\tau}), \ldots, u^{(l)}(t - \boldsymbol{i\tau}), \quad , \boldsymbol{i} = (i_1, \ldots, i_\ell), \quad i_j, l \in \mathbb{Z}^+\}$, where we have denoted $i_1\tau_1 + \cdots + i_\ell\tau_\ell$ by $\boldsymbol{i\tau}$.

Let $\mathcal{E}$ be the vector space over $\mathcal{K}$ given by

$$\mathcal{E} = \mathrm{span}_{\mathcal{K}}\{d\xi : \xi \in \mathcal{K}\} \quad . \tag{3.15}$$

Then $\mathcal{E}$ is the set of linear combinations of a finite number of one-forms from $\{dx(t - \boldsymbol{i\tau}), du(t - \boldsymbol{i\tau}), \ldots, du^{(l)}(t - \boldsymbol{i\tau})\}$ with row vector coefficients in $\mathcal{K}$. The time-shift operator δ_i is defined by

$$\delta_i(\xi(t)) = \xi(t - \tau_i), \quad \xi(t) \in \mathcal{K} \quad , \tag{3.16}$$

and

$$\delta_i(\alpha(t)d\xi(t)) = \alpha(t-\tau_i)d\xi(t-\tau_i), \quad \alpha(t)d\xi(t) \in \mathcal{E} \quad . \tag{3.17}$$

Let $\mathcal{K}(\boldsymbol{\delta}]$ denote the set of polynomials in $\delta_1, \ldots, \delta_\ell$ with coefficients from $\mathcal{K}$. This set is a noncommutative ring where addition is defined as usual, while multiplication is defined as follows. Let $a(\boldsymbol{\delta}]$ be a polynomial in $\mathcal{K}(\boldsymbol{\delta}]$, $a(\boldsymbol{\delta}] = \sum_{\boldsymbol{k}} a_{\boldsymbol{k}}\boldsymbol{\delta}^{\boldsymbol{k}}$, where we have denoted $\delta_1^{k_1}\ldots\delta_\ell^{k_\ell}$ by $\boldsymbol{\delta}^{\boldsymbol{k}}$, $\boldsymbol{k} = (k_1, \ldots, k_\ell)$. Order the different powers $\boldsymbol{k}$ according to the largest k_1, then k_2, etc. According to this order, let the highest degree of $a(\boldsymbol{\delta}]$ be $\boldsymbol{r_a}$, where $\boldsymbol{r_a} = (r_{a,1}, \ldots, r_{a,\ell})$ and analogously for another polynomial $b(\boldsymbol{\delta}]$ in $\mathcal{K}(\boldsymbol{\delta}]$. Multiplication of $a(\boldsymbol{\delta}]$ and $b(\boldsymbol{\delta}]$ is then given by

$$a(\boldsymbol{\delta}]b(\boldsymbol{\delta}] = \sum_{\boldsymbol{k}=\boldsymbol{0}}^{\boldsymbol{r_a}+\boldsymbol{r_b}} \sum_{\boldsymbol{i}+\boldsymbol{j}=\boldsymbol{k}}^{\boldsymbol{i}\leq\boldsymbol{r_a},\boldsymbol{j}\leq\boldsymbol{r_b}} a_i(t)b_j(t-\boldsymbol{i\tau})\boldsymbol{\delta}^{\boldsymbol{k}} \quad . \tag{3.18}$$

We will now show that $\mathcal{K}(\boldsymbol{\delta}]$ is Noetherian.

Proposition 3.4 *$\mathcal{K}(\boldsymbol{\delta}]$ is a Noetherian ring.*

Proof: We have $\mathcal{K}(\boldsymbol{\delta}] = \Big(\big(\mathcal{K}(\delta_1]\big)(\delta_2]\ldots\Big)(\delta_\ell]$. The statement follows by iterative application of Proposition 3.2. ■

It follows that $\mathcal{K}(\boldsymbol{\delta}]$ is also a left Ore domain by Corollary 8.10 in Chapter 0 of (Cohn, 1985).

3.2.2 The fraction field $\mathcal{K}\langle\boldsymbol{\delta})$

The left field of fractions, $\mathcal{K}\langle\boldsymbol{\delta})$, is defined as in the single time delay case. Elements of $\mathcal{K}\langle\delta)$ are denoted $b^{-1}(\boldsymbol{\delta}]a(\boldsymbol{\delta}]$.

3.2.3 Rank and basis of modules over $\mathcal{K}(\boldsymbol{\delta}]$

This subsection is a direct repetition of the corresponding one for the single delay case, Subsection 3.1.3, recounted here for completeness. The ring $\mathcal{K}(\boldsymbol{\delta}]$, being Noetherian, has the invariant basis number (IBN). A free left module over $\mathcal{K}(\boldsymbol{\delta}]$ thus has uniquely defined rank and all its bases have the same cardinality.

Define the following left module

$$\mathcal{M} = \mathrm{span}_{\mathcal{K}(\boldsymbol{\delta}]}\{d\xi : \xi \in \mathcal{K}\} \quad , \tag{3.19}$$

where the elements of the ring $\mathcal{K}(\boldsymbol{\delta}]$ act on elements $d\xi$ of the module according to (3.17). The module $\mathcal{M}$ contains the same elements as $\mathcal{E}$. As in the

case of single time delay, we need to use the concept of closure. The closure of a submodule $\mathcal{N}$ in $\mathcal{M}$ is the submodule

$$\overline{\mathcal{N}} = \{w \in \mathcal{M} : \exists a(\boldsymbol{\delta}] = a(\delta_1, \ldots, \delta_\ell] \in \mathcal{K}(\boldsymbol{\delta}], \quad a(\boldsymbol{\delta}]w \in \mathcal{N}\} \quad .$$

For example, $dx(t) \notin \text{span}_{\mathcal{K}(\boldsymbol{\delta}]}\{dx(t-\tau_i)\}$ but $dx(t) \in \overline{\text{span}_{\mathcal{K}(\boldsymbol{\delta}]}\{dx(t-\tau_i)\}}$ since $\delta_i dx(t) = dx(t-\tau_i) \in \text{span}_{\mathcal{K}(\boldsymbol{\delta}]}\{dx(t-\tau_i)\}$.

If $\mathcal{N}$ coincides with $\overline{\mathcal{N}}$, then $\mathcal{N}$ is said to be closed in $\mathcal{M}$. The result on closure and closedness of modules from Xia et al. (2002) (Lemma 2) and the previous section is now generalised to the present case:

Proposition 3.5 *1. The closure $\overline{\mathcal{N}}$ of $\mathcal{N}$ in $\mathcal{M}$ is the smallest closed submodule of $\mathcal{M}$ containing $\mathcal{N}$;*

2. For any finitely generated submodule $\mathcal{N}$ of $\mathcal{M}$, one has $\text{rank}_{\mathcal{K}(\boldsymbol{\delta}]}\mathcal{N} = \text{rank}_{\mathcal{K}(\boldsymbol{\delta}]}\overline{\mathcal{N}}$*;*

3. For a free submodule $\mathcal{N}$ of $\mathcal{M}$, $\overline{\mathcal{N}}$ is the largest submodule of $\mathcal{M}$ containing $\mathcal{N}$ and having a rank equal to $\text{rank}_{\mathcal{K}(\boldsymbol{\delta}]}\mathcal{N}$.

Proof:

1. Let $\mathcal{P}$ be the smallest closed submodule of $\mathcal{M}$ containing $\mathcal{N}$. Choose an element $u \in \mathcal{N} \subseteq \mathcal{P}$. Since $\mathcal{P}$ is closed, it contains $\mathcal{K}(\boldsymbol{\delta}]u$. This is true for all $u \in \mathcal{N}$ and thus $\overline{\mathcal{N}} = \mathcal{K}(\boldsymbol{\delta}]\mathcal{N} \subseteq \mathcal{P}$.

2. Let $\mathcal{N}$ be generated by $w_1, \ldots, w_s$, assumed to be linearly-independent w.l.o.g. Choose an element $u \in \overline{\mathcal{N}}$. Then $\exists a(\boldsymbol{\delta}] \in \mathcal{K}(\boldsymbol{\delta}]$ such that $a(\boldsymbol{\delta}]u \in \mathcal{N}$ and thus $a(\boldsymbol{\delta}]u = b_1 w_1 + \cdots + b_s w_s$ for some $b_i \in \mathcal{K}(\boldsymbol{\delta}]$. Thus, u is linearly dependent on $w_1, \ldots, w_s$. This is true for all elements in $\overline{\mathcal{N}}$ and thus $\overline{\mathcal{N}}$ has the same rank as $\mathcal{N}$.

3. Let $\mathcal{P}$ be the largest submodule of $\mathcal{M}$ containing $\mathcal{N}$ and having a rank equal to $\text{rank}_{\mathcal{K}(\boldsymbol{\delta}]}\mathcal{N}$. Let w_I be a basis for $\mathcal{N}$, where I is an index set. Choose an element $u \in \mathcal{P}$, then u must be linearly dependent on w_I and $\exists a(\boldsymbol{\delta}], b_i \in \mathcal{K}(\boldsymbol{\delta}]$ such that $a(\boldsymbol{\delta}]u = \sum_{i \in I} b_i w_i$. Thus, $u \in \overline{\mathcal{N}}$ and therefore $\mathcal{P} \subseteq \mathcal{N}$. ■

3.2.4 The Poincaré lemma

Just as in the single-delay case, a 1-form $\omega \in \mathcal{M}$ is also an element of $\mathcal{E}$ and the following form of Poincaré's Lemma holds (Lemma 3 in Márquez-Martínez et al. (2000)):

Lemma 3.3 (Poincaré) *Consider a 1-form $\omega \in \mathcal{M}$. Then there exists a function $\xi(t) \in \mathcal{K}$ such that (locally) $\omega = d\xi(t)$ if and only if $d\omega = 0$.*

3.2.5 Differentiation of functions and one-forms

Differentiation with respect to time along the system dynamics for functions $\phi(x(t-\boldsymbol{i\tau}), u(t-\boldsymbol{j\tau}), \ldots, u^{(l)}(t-\boldsymbol{j\tau}))$ in $\mathcal{K}$ and one-forms $\omega = \sum_{\boldsymbol{i}} \kappa_x^{\boldsymbol{i}} dx(t-\boldsymbol{i\tau}) + \sum_{\boldsymbol{j},r} \nu_{\boldsymbol{j}} du^{(r)}(t-\boldsymbol{j\tau})$ in $\mathcal{M}$ is defined in the natural way (see for example Xia et al. (2002) and Zhang et al. (2006)):

$$\dot{\phi} = \sum_{\boldsymbol{i}} \frac{\partial \phi}{\partial x(t-\boldsymbol{i\tau})} \boldsymbol{\delta^i} f + \sum_{\boldsymbol{j},r} \frac{\partial \phi}{\partial u^{(r)}(t-\boldsymbol{j\tau})} u^{(r+1)}(t-\boldsymbol{j\tau}) \quad (3.20)$$

$$\begin{aligned} \dot{\omega} = & \sum_{\boldsymbol{i}} \dot{\kappa}_x^{\boldsymbol{i}} dx(t-\boldsymbol{i\tau}) + \sum_{\boldsymbol{j},r} \dot{\nu}_{\boldsymbol{j}} du^{(r)}(t-\boldsymbol{j\tau}) + \\ & + \sum_{\boldsymbol{i}} \kappa_x^i d\boldsymbol{\delta^i} f + \sum_{\boldsymbol{j},r} \nu_{\boldsymbol{j}} du^{(r+1)}(t-\boldsymbol{j\tau}) \quad . \end{aligned} \quad (3.21)$$

4 Observability

The literature on the observability of nonlinear time-delay systems appears to be scarce. For linear delay systems, different definitions of observability are described and characterised in (Lee and Olbrot, 1981). Most of them are not, however, directly applicable to nonlinear delay systems. The definitions from (Lee and Olbrot, 1981) that can be used directly for the present case are the algebraic ones - observability over the ring of polynomials in $\boldsymbol{\delta}$ and over the corresponding field of fractions. These have been characterised for nonlinear systems by Xia et al. (2002) under the term weak observability which will be reviewed in Subsection 4.2. Another more intuitive definition of observability, relating the state to the derivatives of the output and input, and their forward shifts, is used in connection to observer design in (Márquez-Martínez et al., 2002), but is not characterised. In this section, we will attempt to relate the definition in (Xia et al., 2002) to the one in (Márquez-Martínez et al., 2002) and also to our own geometric definition of observability obtained by adapting the one for ODE-systems from Part I of this thesis to the present case of delay systems.

4.1 Intuitive definitions

Consider once again the general form for a delay system with multiple time-lags

$$\begin{cases} \dot{x}(t) &= f\big(x(t), x(t-\tau_1), \ldots, x(t-\tau_\ell), u(t), u(t-\tau_1), \ldots, u(t-\tau_\ell)\big) \\ y(t) &= h\big(x(t), x(t-\tau_1), \ldots, x(t-\tau_\ell)\big) \\ x(t) &= \varphi(t), \quad t \in [-\max_i \tau_i, 0] \\ u(t) &= u_0(t), \quad t \in [-T, 0] \quad . \end{cases} \tag{4.1}$$

In this section the time delays τ_i, are assumed to be known. The characterisation of observability with unknown time-delays is left to Subsection 6.2.

First, let us attempt a more intuitive definition of observability by adapting the one for ODE-systems from Part I of this thesis to the present case. For this purpose, observe that the state is no longer a point in $\mathbb{R}^n$, but a function x_t corresponding to the past time interval $[t - \max_i \tau_i, t]$, defined by $x_t(\theta) = x(t+\theta)$, $\theta \in [-\max_i \tau_i, 0]$.

Definition 4.1 *Two initial functions φ^0 and φ^1 in C are **U-distinguishable** if there exists a measurable bounded input $u(t)$ defined on the interval [0,T] that generates solutions $x^0(t)$ and $x^1(t)$ of $\dot{x} = f(x(t), x(t-\tau_1), \ldots, x(t-\tau_\ell), u, u(t-\tau_1), \ldots, u(t-\tau_\ell))$ satisfying $x^i(t) = \varphi^i(t)$, $t \in [-\max_i \tau_i, 0]$ such that $x^i_t \in U$ for all $t \in [0,T]$ and $g(x^0(t)) \neq g(x^1(t))$ for some $t \in [0,T]$. We*

denote by $I(x^0, U)$ *all points* $x^1 \in U$ *that are not* ***U-distinguishable*** *from* x^0.

Definition 4.2 *System* (4.1) *is* ***observable*** *at* $\varphi^0 \in C$ *if* $I(\varphi^0, C) = \varphi^0$.

This definition of observability is close to the definition of identifiability in (Nakagiri et al., 1995).

A weaker, local version is the following

Definition 4.3 *The system* (4.1) *has the* ***local distinguishability property*** *at* $\varphi^0 \in C$ *if* x^0 *has an open neighbourhood* V *such that for every open neighbourhood* U *of* φ^0, $I(\varphi^0, U) \cap V = \varphi^0$.

The definition in (Márquez-Martínez et al., 2002) provides a bridge between the above definition and the more algebraic one in (Xia et al., 2002) reviewed in the next section:

Definition 4.4 *p.448 in (Márquez-Martínez et al., 2002)*
The system (4.1) *is observable if the state* $x(t)$ *can be expressed as a function of the derivatives of the output and the input, and their forward shifts:*

$$x(t) = \psi(y^{(k)}(t + \boldsymbol{i\tau}), u^{(l)}(t + \boldsymbol{j\tau})), \quad k, l \in \mathbb{Z}^+ \quad .$$

Definition 4.4 indicates that the state is uniquely defined by the input and output and thus implies Definition 4.2.

4.2 Weak observability

4.2.1 Definition and criteria

We review the definition and results on weak observability from (Xia et al., 2002).

First, we characterise observability over the ring of polynomials $\mathcal{K}(\boldsymbol{\delta}]$. Define

$$\mathcal{X} = \text{span}_{\mathcal{K}(\boldsymbol{\delta}]}\{dx\} \tag{4.2}$$

$$\mathcal{Y}_k = \text{span}_{\mathcal{K}(\boldsymbol{\delta}]}\{dy, d\dot{y}, \ldots, dy^{(k-1)}\} \tag{4.3}$$

$$\mathcal{U} = \text{span}_{\mathcal{K}(\boldsymbol{\delta}]}\{du, d\dot{u}, \ldots\} \quad . \tag{4.4}$$

Then

$$(\mathcal{Y}_1 + \mathcal{U}) \cap \mathcal{X} \subset (\mathcal{Y}_2 + \mathcal{U}) \cap \mathcal{X} \subset \cdots \subset (\mathcal{Y}_k + \mathcal{U}) \cap \mathcal{X} \subset \ldots \tag{4.5}$$

is an increasing chain of submodules of $\mathcal{X}$. By the fact that $\mathcal{X}$ is a finitely generated module over the Noetherian ring $\mathcal{K}(\boldsymbol{\delta}]$, and thus Noetherian, this sequence must terminate. In fact, $(\mathcal{Y}_k+\mathcal{U})\cap\mathcal{X} = (\mathcal{Y}_n+\mathcal{U})\cap\mathcal{X}$ for $k \geq n$ and $\mathrm{rank}_{\mathcal{K}(\boldsymbol{\delta}]}(\mathcal{Y}_n+\mathcal{U})\cap\mathcal{X} \leq n$ (see (Xia et al., 2002) and (Zhang et al., 2006)).

Denote $\mathcal{O} = \overline{(\mathcal{Y}_n+\mathcal{U})}\cap\mathcal{X}$ where, as previously, $\overline{(\mathcal{Y}_n+\mathcal{U})}$ is the closure of $\mathcal{Y}_n+\mathcal{U}$. The module $\mathcal{O}$ is called the polynomial observation submodule for system (3.14). We can now define observability over $\mathcal{K}(\boldsymbol{\delta}]$ as in (Lee and Olbrot, 1981):

Definition 4.5 *System* (4.1) *is observable over* $\mathcal{K}(\boldsymbol{\delta}]$ *if* $\mathrm{rank}_{\mathcal{K}(\boldsymbol{\delta}]}\mathcal{O} = n$.

We proceed with defining observability over the field of fractions, $\mathcal{K}\langle\boldsymbol{\delta}\rangle$, which is also called weak observability in (Xia et al., 2002). Define

$$\begin{aligned} \bar{\mathcal{X}} &= \mathrm{span}_{\mathcal{K}\langle\boldsymbol{\delta}\rangle}\{dx\} && (4.6)\\ \bar{\mathcal{Y}}_k &= \mathrm{span}_{\mathcal{K}\langle\boldsymbol{\delta}\rangle}\{dy, d\dot{y}, \ldots, dy^{(k-1)}\} && (4.7)\\ \bar{\mathcal{U}} &= \mathrm{span}_{\mathcal{K}\langle\boldsymbol{\delta}\rangle}\{du, d\dot{u}, \ldots\} \quad . && (4.8) \end{aligned}$$

Then the corresponding chain of submodules of $\bar{\mathcal{X}}$

$$(\bar{\mathcal{Y}}_1+\bar{\mathcal{U}})\cap\bar{\mathcal{X}} \subset (\bar{\mathcal{Y}}_2+\bar{\mathcal{U}})\cap\bar{\mathcal{X}} \subset \cdots \subset (\bar{\mathcal{Y}}_k+\bar{\mathcal{U}})\cap\bar{\mathcal{X}} \subset \ldots \tag{4.9}$$

must terminate as $\mathcal{X}$ is a finitely generated module over the field $\mathcal{K}\langle\boldsymbol{\delta}\rangle$, and thus Noetherian. Again, $(\bar{\mathcal{Y}}_k+\bar{\mathcal{U}})\cap\bar{\mathcal{X}} = (\bar{\mathcal{Y}}_n+\bar{\mathcal{U}})\cap\bar{\mathcal{X}}$ for $k \geq n$. Denote $\bar{\mathcal{O}} = (\bar{\mathcal{Y}}_n+\bar{\mathcal{U}})\cap\bar{\mathcal{X}}$. $\bar{\mathcal{O}}$ is called the rational observation submodule for system (4.1).

Definition 4.6 *The system is said to be weakly observable if* $\mathrm{rank}_{\mathcal{K}\langle\boldsymbol{\delta}\rangle}\bar{\mathcal{O}} = n$.

Observability according to Definition 4.4 clearly implies weak observability (Definition 4.5) as

$$\begin{aligned} x(t) &= \psi(y^{(k)}(t+\boldsymbol{i\tau}), u^{(l)}(t+\boldsymbol{j\tau})) \\ &\Updownarrow \\ dx_i &= \sum a_{rki_1\ldots i_\ell}\delta_1^{-i_1}\ldots\delta_\ell^{-i_\ell}dy_r^{(k)} + \sum b_{slj_1\ldots j_\ell}\delta_1^{-j_1}\ldots\delta_\ell^{-j_\ell}du_s^{(l)} \end{aligned} \tag{4.10}$$

where $a_{rki_1\ldots i_\ell}, b_{slj_1\ldots j_\ell} \in \mathcal{K}$. On the other hand, the following system is a counterexample showing that the reverse is not true:

$$\begin{cases} \dot{x}(t) &= x(t) \\ y(t) &= x(t)+x(t-\tau) \quad . \end{cases} \tag{4.11}$$

For this system $dx \in \mathrm{span}_{\mathcal{K}\langle\boldsymbol{\delta}\rangle}\{dy\}$ since $dy = (1+\delta)dx \Leftrightarrow dx = (1+\delta)^{-1}dy$. However, $(1+\delta)^{-1}dy$ cannot be written as $\sum_{i\geq 0}^{k} a_i\delta^{-i}dy$ for any $k \in \mathbb{Z}^+$, $a_i \in \mathcal{K}$.

Xia et al. (2002) showed that, just as in the linear delay system case (Lee and Olbrot, 1981), observability over $\mathcal{K}\langle\boldsymbol{\delta}\rangle$ and over $\mathcal{K}(\boldsymbol{\delta}]$ are equivalent:

Proposition 4.1 *(Theorem 3 in (Xia et al., 2002)).*
The system is weakly observable if and only if $\mathrm{rank}_{\mathcal{K}(\boldsymbol{\delta}]}\overline{(\mathcal{Y}_n + \mathcal{U})} \cap \mathcal{X} = n$.

Furthermore, a variable x_i is weakly observable if and only if $dx_i \in \overline{(\mathcal{Y}_n + \mathcal{U})}$, i.e. there exists $a_i(\boldsymbol{\delta}] \in \mathcal{K}(\boldsymbol{\delta}]$ such that $a_i(\boldsymbol{\delta}]dx_i \in \mathcal{Y}_n + \mathcal{U}$.

Whether weak observability is equivalent to the system having the local distinguishability property almost everywhere, Definition 4.3, is still an open question. Weak observability for the state-variables $x(t)$ implies the existence of a meromorphic function ζ of $x(t)$ and its shifts $x(t-\boldsymbol{i\tau})$ which is a known function ϱ of time (as a meromorphic function of y, u and their shifts). The main difficulty lies in showing that the equation $\zeta(x(t), x(t-\boldsymbol{i\tau})) = \varrho(t)$ has a locally unique solution $x(t)$. The answer to this question is left as future work.

4.2.2 Examples

In this subsection we demonstrate how weak observability is analysed on some simple control systems.

4.2.2.1 Single time-delay

It should be noted that for the calculations performed in this subsection (as well as for many other control problems for delay systems), one can use the symbolic computation algorithm by Gárate-García et al. (2006), which can check the observability of nonlinear systems with known single time delay.
Example 1: An observable system

Consider the system

$$\left\{\begin{array}{lcl} \dot{x}_1(t) &=& x_2(t-\tau) + x_1^2(t) \\ \dot{x}_2(t) &=& x_1(t-\tau) \\ y(t) &=& x_1(t-\tau) \quad . \end{array}\right. \tag{4.12}$$

Taking $n-1$ time-derivatives of the output function, we obtain the equations

$$y = \delta x_1 \tag{4.13}$$
$$\dot{y} = \delta^2 x_2 + (\delta x_1)^2 \quad . \tag{4.14}$$

Thus,

$$\begin{bmatrix} dy \\ d\dot{y} \end{bmatrix} = \begin{bmatrix} \delta & 0 \\ 2(\delta x_1)\delta & \delta^2 \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} \quad . \tag{4.15}$$

The above matrix has rank 2 over $\mathcal{K}(\delta]$ and thus $\text{rank}_{\mathcal{K}(\boldsymbol{\delta}]}\overline{(\mathcal{Y}_2 + \mathcal{U})} \cap \mathcal{X} = 2$ and the system is weakly observable. In fact, $x_1(t) = y(t+\tau)$ and $x_2 = \dot{y}(t+2\tau) - y^2(t+2\tau)$.

Example 2: An unobservable system

Consider the system

$$\begin{cases} \dot{x}_1(t) &= x_1(t-\tau) + u(t) \\ \dot{x}_2(t) &= x_2(t-\tau) \\ y(t) &= x_1(t) + x_2(t) \quad . \end{cases} \tag{4.16}$$

We have the equations

$$y = x_1 + x_2 \tag{4.17}$$

$$\dot{y} = \delta x_1 + \delta x_2 + u \quad . \tag{4.18}$$

Thus,

$$\begin{bmatrix} dy \\ d\dot{y} - du \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \delta & \delta \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} \quad . \tag{4.19}$$

The above matrix has rank 1 over $\mathcal{K}(\delta]$ and the system is not weakly observable.

4.2.2.2 Multiple time-delays

Example 1: An observable system

Consider the system

$$\begin{cases} \dot{x}_1(t) &= x_2^2(t-\tau_1)x_1(t-\tau_2) + u(t) \\ \dot{x}_2(t) &= x_1(t-\tau_2) \\ y(t) &= x_1(t-\tau_1) \quad . \end{cases} \tag{4.20}$$

Taking $n-1$ time-derivatives of the output function, we obtain the equations

$$y = \delta_1 x_1 \tag{4.21}$$

$$\dot{y} = (\delta_1^2 x_2)^2 \delta_1 \delta_2 x_1 + \delta_1 u \quad . \tag{4.22}$$

Thus,

$$\begin{bmatrix} dy \\ d\dot{y} - d\delta_1 u \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ (\delta_1^2 x_2)^2 \delta_1 \delta_2 & 2\delta_1^2 x_2 \delta_1 \delta_2 x_1 \delta_1^2 \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} \quad . \tag{4.23}$$

The above matrix has rank 2 over $\mathcal{K}(\delta]$ and thus $\mathrm{rank}_{\mathcal{K}(\boldsymbol{\delta}]}\overline{(\mathcal{Y}_2+\mathcal{U})}\cap\mathcal{X}=2$ and the system is weakly observable. In fact,

$$x_1(t) = y(t+\tau_1) \tag{4.24}$$

$$x_2^2(t) = \frac{\dot{y}(t+2\tau_1)-u(t+\tau_1)}{y(t-\tau_2+2\tau_1))} \quad . \tag{4.25}$$

4.3 Observability analysis for DDE systems reduces to the analysis of the Jacobian of an algebraic system

In this subsection, we show that weak observability for delay systems, as characterised in Section 4.2, can be checked by analysis of the Jacobian of a system of algebraic equations derived from the original delay system (as for ODE-systems), simplifying the problem and allowing for the application of existing computer algorithms for symbolic rank calculation. Based on the results in this section, one could potentially use algorithms like the one by Sedoglavic (2002) to test the observability of large delay systems with many variables (which is so far difficult to do with the symbolic package by Gárate-García et al. (2006); it is for the moment also limited to one time delay).

4.3.1 Single time-delay

Consider once again the system

$$\left\{\begin{array}{lcl} \dot{x}(t) &=& f\big(x(t),x(t-\tau),u,u(t-\tau)\big)\\ y(t) &=& h\big(x(t),x(t-\tau)\big)\\ x(t) &=& \varphi(t), \quad t\in[-\tau,0]\\ u(t) &=& u_0(t), \quad t\in[-T,0] \end{array}\right. \tag{4.26}$$

and the np equations for the state-variables obtained by taking $n-1$ time-derivatives of the p output variables:

$$\begin{aligned} y_1(t) &= h_1(x(t),x(t-\tau))\\ &\vdots\\ y_1^{(n-1)}(t) &= h_1^{(n-1)}(x(t),x(t-\tau),\ldots,x(t-n\tau),u(t),\ldots,u^{(n-2)}(t),\ldots,\\ &\quad u(t-(n-1)\tau),\ldots,u^{(n-2)}(t-(n-1)\tau))\\ &\vdots \end{aligned}$$

$$\begin{aligned} y_p(t) &= h_p(x(t), x(t-\tau)) \\ &\vdots \\ y_p^{(n-1)}(t) &= h_p^{(n-1)}(x(t), x(t-\tau), \ldots, x(t-n\tau)), u(t), \ldots, u^{(n-2)}(t), \ldots, \\ &\quad u(t-(n-1)\tau), \ldots, u^{(n-2)}(t-(n-1)\tau)) \quad . \end{aligned} \tag{4.27}$$

For the $n(n+1)$ variables $x_1(t), \ldots, x_1(t-n\tau), \ldots, x_n(t), \ldots, x_n(t-n\tau)$ occurring above, we can obtain more relations by taking shifts of the equations above to obtain a system with $\frac{pn(n+1)}{2}$ equations. For each $k = 1, \ldots, p$ we have the $\frac{n(n+1)}{2}$ equations:

$$\begin{aligned} y_k(t) &= h_k(x(t), x(t-\tau)) \\ &\vdots \\ \delta^{n-1} y_k(t) &= h_k(x(t-(n-1)\tau), x(t-n\tau) \\ &\vdots \\ y_k^{(n-1)}(t) &= h_k^{(n-1)}(x(t), \ldots, x(t-n\tau)), u(t), \ldots, u^{(n-2)}(t), \\ &\quad \ldots, u(t-(n-1)\tau), \ldots, u^{(n-2)}(t-(n-1)\tau)) \\ \delta y_k^{(n-1)}(t) &= h_k^{(n-1)}(x(t-\tau), \ldots, x(t-n\tau)), u(t-\tau), \ldots, u^{(n-2)}(t-\tau), \\ &\quad \ldots, u(t-(n-1)\tau), \ldots, u^{(n-2)}(t-(n-1)\tau)) \\ y_k^{(n)}(t) &= h_k^{(n)}(x(t), x(t-\tau), \ldots, x(t-n\tau), u(t), \ldots, u^{(n-2)}(t), \ldots, \\ &\quad u(t-(n-1)\tau), \ldots, u^{(n-2)}(t-(n-1)\tau)) \quad . \end{aligned} \tag{4.28}$$

Gather all sets of equations of the form (4.28) for the different $k = 1, \ldots, p$. Consider the Jacobian matrix for this system of algebraic equations for the variables $x_1(t), \ldots, x_1(t-n\tau), \ldots, x_n(t), \ldots, x_n(t-n\tau)$, denoted by J. Let the matrix J_i consist of the columns of J corresponding to $dx_i, \ldots, d\delta^n x_i$ and so $J = [J_1|\ldots|J_n]$.

We can now formulate the main result in the following proposition:

Proposition 4.2 *The variable x_i is weakly observable if and only if*

$$\mathrm{rank}_{\mathcal{K}}[J_1|\ldots|J_{i-1}|J_{i+1}|\ldots|J_n] < \mathrm{rank}_{\mathcal{K}} J \quad . \tag{4.29}$$

Proof: Denote by $\mathcal{S}^{J_y, J_u}$ the set

$$\mathrm{span}_{\mathcal{K}}\{dy_k^{(l)}(t-(j_y-l)\tau), du_r^{(l-1)}(t-(j_u-1)\tau)\} \quad ,$$

for $k = 1, \ldots, p$, $r = 1, \ldots, m$, $l \leq n-1$, $j_y = 0, \ldots, J_y$ and $j_u = 1, \ldots, J_u$, where J_y, $J_u \in \mathbb{N}$.

If $\text{rank}_{\mathcal{K}}[J_1|\dots|J_{i-1}|J_{i+1}|\dots|J_n] < \text{rank}_{\mathcal{K}}[J]$, then there is a (nonzero) linear combination of the one-forms $dx_i(t), \dots, dx_i(t-n\tau)$ which is in $\mathcal{S}^{n,n}$, that is,

$$\sum_{j=0}^{n} A_j dx_i(t-j\tau) = \sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j=0}^{n-1-l} B_{jkl} dy_k^{(l)}(t-j\tau) + \sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j=0}^{n-1} E_{jkl} du_k^{(l)}(t-j\tau) \tag{4.30}$$

for some $A_0, \dots, A_{n+1}$, B_{jkl} and $E_{jkl} \in \mathcal{K}$. Then $a_i(\delta)dx_i \in \mathcal{Y}_n + \mathcal{U}$ where $a_i(\delta) = \sum_{j=0}^{n} A_j \delta^j \in \mathcal{K}(\delta]$. Thus the variable x_i is weakly observable. Suppose now that $\text{rank}_{\mathcal{K}}[J_1|\dots|J_{i-1}|J_{i+1}|\dots|J_n] = \text{rank}_{\mathcal{K}}[J]$ and thus there is no (nonzero) linear combination of the one-forms $dx_i(t), \dots, dx_i(t-n\tau)$ that is in $\mathcal{S}^{n,n}$. Using this as a starting point, we will prove that x_i is not weakly observable by induction. Suppose that no (nonzero) linear combination of the one-forms $dx_i(t), \dots, dx_i(t-N\tau)$, $N \geq n$ is in $\mathcal{S}^{M,M}$ for some $M \geq N$. We will show that no linear combination of $dx_i(t), \dots, dx_i(t-(N+1)\tau)$ is in $\mathcal{S}^{M+1,M+1}$. For this, suppose that there is such a linear combination and thus, there exist $\tilde{A}_0, \dots, \tilde{A}_{N+1} \in \mathcal{K}$ and $\tilde{B}_{jkl}, \tilde{E}_{jkl} \in \mathcal{K}$, such that

$$\sum_{j=0}^{N+1} \tilde{A}_j dx_i(t-j\tau) = \sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j=0}^{M-l} \tilde{B}_{jkl} dy_k^{(l)}(t-j\tau) + \sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j=0}^{M} \tilde{E}_{jkl} du_k^{(l)}(t-j\tau). \tag{4.31}$$

Then,

$$\begin{aligned}
\sum_{j=0}^{N} \tilde{A}_j dx_i(t-j\tau) &= \sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j=0}^{N-1-l} \tilde{B}_{jkl} dy_k^{(l)}(t-j\tau) + \\
&+ \sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j=0}^{M} \tilde{E}_{jkl} du_k^{(l)}(t-j\tau) + \\
&+ \sum_{k=1}^{p}\sum_{l=0}^{n}\sum_{j=N-l}^{M-l} \tilde{B}_{jkl} dy_k^{(l)}(t-j\tau) - \\
&- \tilde{A}_{N+1} dx_i(t-(N+1)\tau) \quad .
\end{aligned} \tag{4.32}$$

Note first that the one-forms $dx_i(t-j\tau)$, $j = 0, \dots, N+1$ and $du_k^{(l)}(t-j\tau)$, $k = 1, \dots, m$, $l = 0, \dots, n-2$, $j = 0, \dots, M$ are linearly independent. The left-hand side in the above equation does not contain one-forms $dx_i(t-j\tau)$ for $j > N$ and thus, the right-hand side may not either. From the formulation of (4.28), we know that the first two sums on the right-hand side may only contain terms $dx_i(t-j\tau)$ for $j \leq N$ and the same must be true for the rest

of the terms, that is,

$$\sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j=N-l}^{M-l}\tilde{B}_{jkl}dy_k^{(l)}(t-j\tau)-\tilde{A}_{N+1}dx(t-(N+1)\tau)= \\ =\sum_{j=0}^{N}\bar{A}_j dx_i(t-j\tau)+\sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j=0}^{M}\bar{E}_{jkl}du_k^{(l)}(t-j\tau) \qquad (4.33)$$

and therefore

$$\sum_{j=0}^{N}(\tilde{A}_j-\bar{A}_j)dx_i(t-j\tau) = \sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j=0}^{N-l}\tilde{B}_{jkl}dy_k^{(l)}(t-j\tau)+ \\ + \sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j=0}^{M}(\tilde{E}_{jkl}+\bar{E}_{jkl})du_k^{(l)}(t-j\tau) \qquad (4.34)$$

Since the term

$$\sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j=0}^{N-1-l}\tilde{B}_{jkl}dy_k^{(l)}(t-j\tau)$$

on the right-hand side may not contain one-forms $du_k^{(l)}(t-M\tau)$, $k=1,\ldots,m$, $l=0,\ldots,n-2$ (because $M\geq N$), the other term on the right-hand side,

$$\sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j=0}^{M}(\tilde{E}_{jkl}+\bar{E}_{jkl})du_k^{(l)}(t-j\tau)$$

may not either and we thus have

$$\sum_{j=0}^{N}(\tilde{A}_j-\bar{A}_j)dx_i(t-j\tau) = \sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j=0}^{N-1-l}\tilde{B}_{jkl}dy_k^{(l)}(t-j\tau)+ \\ + \sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j=0}^{M-1}(\tilde{E}_{jkl}+\bar{E}_{jkl})du_k^{(l)}(t-j\tau) \qquad (4.35)$$

Clearly, if not all of the coefficients $\tilde{A}_j-\bar{A}_j$ are zero, we have a nonzero linear combination of the one-forms $dx_i(t),\ldots,dx_i(t-N\tau)$ which is in $\mathcal{S}^{N,M}\subset\mathcal{S}^{M,M}$, contradicting the assumption. If, instead, they are all zero and thus

$$\sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j=0}^{N-1-l}\tilde{B}_{jkl}dy_k^{(l)}(t-j\tau)+\sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j=0}^{M}(\tilde{E}_{jkl}+\bar{E}_{jkl})du_k^{(l)}(t-j\tau)=0, \qquad (4.36)$$

then,

$$\sum_{j=0}^{N+1} \tilde{A}_j dx_i(t-j\tau) = \sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j=N-l}^{M-l} \tilde{B}_{jkl} dy_k^{(l)}(t-j\tau) - \sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j=0}^{M} \bar{E}_{jkl} du_k^{(l)}(t-j\tau). \tag{4.37}$$

Taking into account that the term

$$\sum_{k=1}^{p}\sum_{l=0}^{n}\sum_{j=N-l}^{M-l} \tilde{B}_{jkl} dy_k^{(l)}(t-j\tau)$$

on the right-hand side does not contain the undelayed one-forms dx_i or $du_j, j = 1, \ldots, m$, we have

$$\sum_{j=1}^{N+1} \tilde{A}_j dx_i(t-j\tau) = \sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j=N-l}^{M-l} \tilde{B}_{jkl} dy_k^{(l)}(t-j\tau) - \sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j=1}^{M} \bar{E}_{jkl} du_k^{(l)}(t-j\tau). \tag{4.38}$$

The above equation can be written

$$\begin{aligned}
&\delta\big(\sum_{j=0}^{N} \delta^{-1}(\tilde{A}_j) dx_i(t-j\tau)\big) = \\
&\quad = \delta\big(\sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j=N-1-l}^{M-1-l} \delta^{-1}(\tilde{B}_{jkl}) dy_k^{(l)}(t-j\tau) - \\
&\quad - \sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j=0}^{M-1} \delta^{-1}(\bar{E}_{jkl}) du_k^{(l)}(t-j\tau)\big) \\
&\Updownarrow \\
&\sum_{j=0}^{N} (\delta^{-1}\tilde{A}_j) dx_i(t-j\tau) = \\
&\quad = \sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j=N-1-l}^{M-1-l} \delta^{-1}(\tilde{B}_{jkl}) dy_k^{(l)}(t-j\tau) - \\
&\quad - \sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j=0}^{M-1} \delta^{-1}(\bar{E}_{jkl}) du_k^{(l)}(t-j\tau),
\end{aligned} \tag{4.39}$$

where all $\delta^{-1}(\tilde{A}_j)$, $\delta^{-1}(\tilde{B}_{jkl})$ and $\delta^{-1}(\bar{E}_{jkl})$ are in $\mathcal{K}$. Thus, we obtain a contradiction to the assumption in this case as well. By the principle of induction, we now have that for the one-forms $dx_i(t), \ldots, dx_i(t-P\tau)$, $P \geq n$,

there can be no nonzero linear combination of them which is in $\mathcal{S}^{Q,Q}$ for $Q \geq P$. Therefore, $dx_i \notin \overline{\mathcal{Y}_n + \mathcal{U}}$ and x_i is not weakly observable. This completes the proof of the proposition. ■

We conclude this section by observing that for analysing weak observability, we need only consider the state-variables and their shifts that occur in the $n-1$ output derivatives (4.27) and these may be a smaller set than the $n(n+1)$ variables $x_1(t), \ldots, x_1(t-n\tau), \ldots, x_n(t), \ldots, x_n(t-n\tau)$ considered in the general case.

We will now demonstrate the calculations on the two simple examples from subsection 4.2.2.1 and also show how the analysis can be automated, allowing for the application of symbolic computation algorithms.

Example 1: An observable system

Consider once again the system

$$\left\{ \begin{array}{lcl} \dot{x}_1(t) & = & x_2(t-\tau) + x_1^2(t) \\ \dot{x}_2(t) & = & x_1(t-\tau) \\ y(t) & = & x_1(t-\tau) \quad . \end{array} \right. \tag{4.40}$$

Taking $n-1 = 1$ time-derivative of the output function, we obtain the equations

$$y = \delta x_1 \tag{4.41}$$

$$\dot{y} = \delta^2 x_2 + (\delta x_1)^2 \quad . \tag{4.42}$$

For the variables occurring above, we can get one more equation by considering a shift of the first equation:

$$\delta y = \delta^2 x_1 \quad . \tag{4.43}$$

We will now rename all x and $y, \dot{y}$ variables occurring in the equations above according to the following scheme:

$$z_{10} := x_1, \qquad z_{11} := \delta x_1, \qquad z_{12} := \delta^2 x_1 \tag{4.44}$$

$$z_{20} := x_2, \qquad z_{21} := \delta x_2, \qquad z_{22} := \delta^2 x_2 \tag{4.45}$$

and

$$y_{00} := y, \qquad y_{01} := \delta y \tag{4.46}$$

$$y_{10} := \dot{y} \quad . \tag{4.47}$$

We therefore have the following algebraic system of 3 equations for the 3 z-variables:

$$y_{00} = z_{11} \tag{4.48}$$

$$y_{01} = z_{12} \tag{4.49}$$

$$y_{10} = z_{22} + z_{11}^2 \quad . \tag{4.50}$$

Consider the Jacobian $J = [J_1|J_2]$. J_i consists of the columns of J corresponding to the one-forms $dz_{i0}, \ldots, dz_{i2}$ occurring in the linear form of the above algebraic system, and thus J_1 includes the columns corresponding to dz_{11}, dz_{12} and J_2 consists of the column corresponding to dz_{22}. We have

$$\begin{bmatrix} dy_{00} \\ dy_{01} \\ dy_{10} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2z_{11} & 0 & 1 \end{bmatrix} \begin{bmatrix} dz_{11} \\ dz_{12} \\ dz_{22} \end{bmatrix} \tag{4.51}$$

and

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2z_{11} & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad . \tag{4.52}$$

We see that the rank of the matrix is 3. If we remove the first 2 columns corresponding to z_{11}, z_{12}, it becomes 1, and if we remove the second group of 1 column, that for z_{22}, we get rank 2. Thus both x_1 and x_2 are observable.

Example 2: An unobservable system

Consider once again the system

$$\begin{cases} \dot{x}_1(t) &= x_1(t-\tau) + u(t) \\ \dot{x}_2(t) &= x_2(t-\tau) \\ y(t) &= x_1(t) + x_2(t) \quad . \end{cases} \tag{4.53}$$

We have the equations

$$y = x_1 + x_2 \tag{4.54}$$

$$\dot{y} = \delta x_1 + \delta x_2 + u \quad . \tag{4.55}$$

Observe that for the variables occurring above, we get one more equation by taking delays of the equations above

$$\delta y = \delta x_1 + \delta x_2 \qquad \delta^2 y = \delta^2 x_1 + \delta^2 x_2 \quad . \tag{4.56}$$

We change variables as in Example 1

$$z_{10} := x_1 \qquad z_{11} := \delta x_1 \qquad z_{12} := \delta^2 x_1 \tag{4.57}$$

$$z_{20} := x_2 \qquad z_{21} := \delta x_2 \qquad z_{22} := \delta^2 x_2 \tag{4.58}$$

and

$$y_{00} := y \qquad y_{01} := \delta y \tag{4.59}$$

$$y_{10} := \dot{y} \quad . \tag{4.60}$$

We therefore have the following algebraic system of 3 equations for the 4 z-variables

$$\begin{aligned} y_{00} &= z_{10} + z_{20} && (4.61)\\ y_{01} &= z_{11} + z_{21} && (4.62)\\ y_{10} &= z_{11} + z_{21} + u \quad . && (4.63) \end{aligned}$$

The Jacobian with respect to the z-variables is given by

$$\begin{bmatrix} dy_{00} \\ dy_{01} \\ dy_{10} - du \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} dz_{10} \\ dz_{11} \\ dz_{20} \\ dz_{21} \end{bmatrix} \quad . \tag{4.64}$$

Here the rank of the matrix does not change when we remove groups of columns corresponding to $z_{i0}, \ldots, z_{i2}$, $i = 1, 2$. Thus, x_1 and x_2 are not observable.

4.3.2 Multiple time-delays

Consider once again the system

$$\begin{cases} \dot{x}(t) &= f\big(x(t), x(t-\tau_1), \ldots, x(t-\tau_\ell), u, u(t-\tau_1), \ldots, u(t-\tau_\ell)\big) \\ y(t) &= h\big(x(t), x(t-\tau_1), \ldots, x(t-\tau_\ell)\big) \\ x(t) &= \varphi(t), \quad t \in [-\max_i \tau_i, 0] \\ u(t) &= u_0(t), \quad t \in [-T, 0] \quad . \end{cases} \tag{4.65}$$

We introduce the following notation: $x_{[i]}(t)$ denotes all the variables of type

$$x(t - \sum_{j=1}^{\ell} i_j \tau_j), \quad i_j \in \mathbb{Z}^+, \quad \sum_{j=1}^{\ell} i_j \leq i$$

and analogously for u and its derivatives. This notation is found to be useful after observing that each time-derivative of an output function y_i introduces only one new delay in each state variable and input derivative, a delay chosen from the τ_i:s, $i = 0, 1, \ldots, \ell$ where τ_0 is used for the identity, or zero delay. With this new notation, the np equations for the state-variables obtained by

taking $n-1$ time-derivatives of the p output variables can be written:

$$\begin{aligned}
y_1(t) &= h_1\big(x_{[1]}(t)\big)\\
&\vdots\\
y_1^{(i)}(t) &= h_1\big(x_{[i+1]}(t), u_{[i]}(t), \ldots, u_{[i]}^{(i-1)}(t)\big)\\
&\vdots\\
y_1^{(n-1)}(t) &= h_1^{(n-1)}\big(x_{[n]}(t), u_{[n-1]}(t), \ldots, u_{[n-1]}^{(n-2)}(t)\big)\\
&\vdots\\
y_p(t) &= h_p\big(x_{[1]}(t)\big)\\
&\vdots\\
y_p^{(n-1)}(t) &= h_p^{(n-1)}\big(x_{[n]}(t), u_{[n-1]}(t), \ldots, u_{[n-1]}^{(n-2)}(t)\big) \quad . \qquad (4.66)
\end{aligned}$$

For the $n\binom{n+\ell}{\ell}$ variables $x_{1,[n]}(t), \ldots, x_{n,[n]}(t)$ occurring above, we can obtain more relations by taking shifts by the different δ_i of the equations above to obtain a system with $p\sum_{i=1}^{n}\binom{i}{\ell}$ equations. For each $k = 1, \ldots, p$ we have the $\sum_{i=1}^{n}\binom{i}{\ell}$ equations:

$$\begin{aligned}
y_k(t) &= h_k\big(x_{[1]}(t)\big)\\
\delta_1 y_k(t) &= h_k\big(x_{[1]}(t-\tau_1)\big)\\
&\vdots\\
\delta_1^{i_1}\ldots\delta_\ell^{i_\ell} y_k(t) &= h_k\big(x_{[n]}(t)\big), \quad i_j \geq 0, \quad \sum_{j=1}^{\ell} i_j \leq n-1\\
&\vdots\\
y_k^{(n-2)}(t) &= h_k^{(n-2)}\big(x_{[n-1]}(t), u_{[n-2]}(t), \ldots, u_{[n-2]}^{(n-3)}(t)\big)\\
\delta_1 y_k^{(n-2)}(t) &= h_k^{(n-2)}\big(x_{[n-1]}(t-\tau_1), u_{[n-2]}(t), \ldots, u_{[n-2]}^{(n-3)}(t-\tau_1)\big)\\
&\vdots\\
\delta_\ell y_k^{(n-2)}(t) &= h_k^{(n-2)}\big(x_{[n-1]}(t-\tau_\ell), u_{[n-2]}(t-\tau_\ell), \ldots, u_{[n-2]}^{(n-3)}(t-\tau_\ell)\big)\\
y_k^{(n-1)}(t) &= h_k^{(n-1)}\big(x_{[n]}(t), u_{[n-1]}(t), \ldots, u_{[n-1]}^{(n-2)}(t)\big) \quad . \qquad (4.67)
\end{aligned}$$

Consider the Jacobian matrix for the above system of algebraic equations for the variables $x_{1,[n]}(t), \ldots, x_{n,[n]}(t)$, denoted by J. Let the matrix J_i consist of the columns of J corresponding to $dx_i, d(\delta_1 x_i), \ldots, d(\delta_1^{i_1}\ldots\delta_\ell^{i_\ell} x_i)$, $\sum_{j=1}^{\ell} i_j \leq n$, (these one-forms are also denoted by $dx_{i,[n]}(t)$) and so $J = [J_1|\ldots|J_n]$.

The main result is the same as for systems with a single time-delay (see the previous subsection):

Proposition 4.3 *The variable x_i is weakly observable if and only if*

$$\mathrm{rank}_{\mathcal{K}}[J_1|\dots|J_{i-1}|J_{i+1}|\dots|J_n] < \mathrm{rank}_{\mathcal{K}} J \quad . \tag{4.68}$$

The proof of this proposition is an extension of the one for a single time-delay in the previous subsection. Nevertheless, we present it here in full detail for completeness:
Proof: Denote

$$\mathrm{span}_{\mathcal{K}}\{dy^{(l)}_{k,[J_y-l]}(t), du^{(l-1)}_{k,[J_u-1]}(t), k=1,\dots,p, l\le n-1\}$$

by $\mathcal{S}^{J_y,J_u}$ where $J_y, J_u \in \mathbb{N}$. As previously, by $dy^{(l)}_{k,[j]}(t)$ we mean all one-forms of type

$$dy^{(l)}_k(t-\sum_{s=1}^{\ell} i_s\tau_j), \quad i_s\in\mathbb{Z}^+, \quad \sum_{s=1}^{\ell} i_s \le j$$

and analogously for u and its derivatives.

If $\mathrm{rank}_{\mathcal{K}}[J_1|\dots|J_{i-1}|J_{i+1}|\dots|J_n] < \mathrm{rank}_{\mathcal{K}}[J]$, then there is a (nonzero) linear combination of the one-forms $dx_{i,[n]}(t)$ which is in $\mathcal{S}^{n,n}$, that is,

$$\sum_{j_s=0}^{n} A_{\boldsymbol{j}} dx_i(t-\boldsymbol{j\tau}) = \tag{4.69}$$
$$= \sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j_s=0}^{n-1-l} B_{\boldsymbol{j}kl} dy^{(l)}_k(t-\boldsymbol{j\tau}) + \sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j_s=0}^{n-1} E_{\boldsymbol{j}kl} du^{(l)}_k(t-\boldsymbol{j\tau}) \quad ,$$

for some $A_{\boldsymbol{j}} \in \mathcal{K}$,$\boldsymbol{j} = (j_1,\dots,j_\ell)$, $0\le j_s\le n$, $B_{\boldsymbol{j}kl}$ and $E_{\boldsymbol{j}kl}\in\mathcal{K}$. Then $a_i(\boldsymbol{\delta})dx_i \in \mathcal{Y}_n + \mathcal{U}$ where $a_i(\boldsymbol{\delta}) = \sum_{j_s=0}^{n} A_{\boldsymbol{j}}\boldsymbol{\delta}^{\boldsymbol{j}} \in \mathcal{K}(\boldsymbol{\delta}]$. Thus the variable x_i is weakly observable.

Suppose now that $\mathrm{rank}_{\mathcal{K}}[J_1|\dots|J_{i-1}|J_{i+1}|\dots|J_n] = \mathrm{rank}_{\mathcal{K}}[J]$ and thus there is no (nonzero) linear combination of the one-forms $dx_{i,[n]}(t)$ that is in $\mathcal{S}^{n,n}$. Using this as a starting point, we will prove that x_i is not weakly observable by induction. Suppose that no (nonzero) linear combination of the one-forms $dx_{i,[N]}(t)$, $N\ge n$ is in $\mathcal{S}^{M,M}$ for some $M\ge N$. We will show that no linear combination of $dx_{i,[N+1]}(t)$ is in $\mathcal{S}^{M+1,M+1}$. For this, suppose that there exist $\tilde{A}_{\boldsymbol{j}}\in\mathcal{K}$, $\boldsymbol{j}=(j_1,\dots,j_\ell), 0\le j_s\le N+1$ and $\tilde{B}_{\boldsymbol{j}kl}, \tilde{E}_{\boldsymbol{j}kl}\in\mathcal{K}$, such that

$$\sum_{j_s=0}^{N+1}\tilde{A}_{\boldsymbol{j}} dx_i(t-\boldsymbol{j\tau}) = \sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j_s=0}^{M-l}\tilde{B}_{\boldsymbol{j}kl} dy^{(l)}_k(t-\boldsymbol{j\tau}) + \sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j_s=0}^{M}\tilde{E}_{\boldsymbol{j}kl} du^{(l)}_k(t-\boldsymbol{j\tau}). \tag{4.70}$$

Then,

$$\begin{aligned}\sum_{j_s=0}^{N}\tilde{A}_{\boldsymbol{j}}dx_i(t-\boldsymbol{j}\boldsymbol{\tau}) &= \sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j_s=0}^{N-1-l}\tilde{B}_{\boldsymbol{j}kl}dy_k^{(l)}(t-\boldsymbol{j}\boldsymbol{\tau})+\\ &+\sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j_s=0}^{M}\tilde{E}_{\boldsymbol{j}kl}du_k^{(l)}(t-\boldsymbol{j}\boldsymbol{\tau})+\\ &+\sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{\max_s j_s=N-l}^{M-l}\tilde{B}_{\boldsymbol{j}kl}dy_k^{(l)}(t-\boldsymbol{j}\boldsymbol{\tau})-\\ &-\sum_{\max_s j_s=N+1}\tilde{A}_j dx_i(t-\boldsymbol{j}\boldsymbol{\tau}). \end{aligned} \tag{4.71}$$

Note first that the one-forms $dx_{i,[N+1]}(t)$ and $du_{k,[M]}^{(l)}(t)$, $k=1,\ldots,m$, $l=0,\ldots,n-2$ are linearly independent. The left-hand side in the above equation does not contain one-forms $dx_i(t-\boldsymbol{j}\boldsymbol{\tau})$ for $\boldsymbol{j}=(j_1,\ldots,j_\ell),\max_s j_s>N$ and thus, the right-hand side may not either. From the formulation of (4.67), we know that the first two sums on the right-hand side may only contain terms $dx_{i,[j]}(t)$ for $j\le N$ and the same must be true for the rest of the terms, that is,

$$\begin{aligned}&\sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{\max_s j_s=N-l}^{M-l}\tilde{B}_{\boldsymbol{j}kl}dy_k^{(l)}(t-\boldsymbol{j}\boldsymbol{\tau})-\sum_{\max_s j_s=N+1}\tilde{A}_j dx_i(t-\boldsymbol{j}\boldsymbol{\tau})=\\ &\quad=\sum_{j=0}^{N}\bar{A}_j dx_i(t-j\tau)+\sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j=0}^{M}\bar{E}_{jkl}du_k^{(l)}(t-j\tau)\end{aligned} \tag{4.72}$$

and therefore

$$\begin{aligned}\sum_{j_s=0}^{N}(\tilde{A}_{\boldsymbol{j}}-\bar{A}_{\boldsymbol{j}})dx_i(t-\boldsymbol{j}\boldsymbol{\tau}) &= \sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j_s=0}^{N-1-l}\tilde{B}_{\boldsymbol{j}kl}dy_k^{(l)}(t-\boldsymbol{j}\boldsymbol{\tau})+\\ &+\sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j_s=0}^{M}(\tilde{E}_{\boldsymbol{j}kl}+\bar{E}_{\boldsymbol{j}kl})du_k^{(l)}(t-\boldsymbol{j}\boldsymbol{\tau})\quad .\end{aligned} \tag{4.73}$$

The first term on the right-hand side,

$$\sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j_s=0}^{N-1-l}\tilde{B}_{\boldsymbol{j}kl}dy_k^{(l)}(t-\boldsymbol{j}\boldsymbol{\tau})$$

may not contain one-forms $du_k^{(l)}(t-\boldsymbol{j}\boldsymbol{\tau})$, $k=1,\ldots,m$, $l=0,\ldots,n-2$ for $\max_s j_s \geq M$ (because $M \geq N$). Therefore, the other term on the right-hand side,

$$\sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j_s=0}^{M}(\tilde{E}_{\boldsymbol{j}kl}+\bar{E}_{\boldsymbol{j}kl})du_k^{(l)}(t-\boldsymbol{j}\boldsymbol{\tau})$$

may not either and we thus have

$$\sum_{j_s=0}^{N}(\tilde{A}_{\boldsymbol{j}}-\bar{A}_{\boldsymbol{j}})dx_i(t-\boldsymbol{j}\boldsymbol{\tau})=\sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j_s=0}^{N-1-l}\tilde{B}_{\boldsymbol{j}kl}dy_k^{(l)}(t-\boldsymbol{j}\boldsymbol{\tau})+$$
$$+\sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j_s=0}^{M-1}(\tilde{E}_{\boldsymbol{j}kl}+\bar{E}_{\boldsymbol{j}kl})du_k^{(l)}(t-\boldsymbol{j}\boldsymbol{\tau}) \quad . \tag{4.74}$$

Clearly, if not all of the coefficients $\tilde{A}_{\boldsymbol{j}}-\bar{A}_{\boldsymbol{j}}$ are zero, we have a nonzero linear combination of the one-forms $dx_{i,[N]}(t)$ which is in $\mathcal{S}^{N,M} \subset \mathcal{S}^{M,M}$, contradicting the assumption. If, instead, they are all zero and thus

$$\sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j_s=0}^{N-1-l}\tilde{B}_{\boldsymbol{j}kl}dy_k^{(l)}(t-\boldsymbol{j}\boldsymbol{\tau})+\sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j_s=0}^{M-1}(\tilde{E}_{\boldsymbol{j}kl}+\bar{E}_{\boldsymbol{j}kl})du_k^{(l)}(t-\boldsymbol{j}\boldsymbol{\tau})=0, \tag{4.75}$$

then,

$$\begin{aligned}\sum_{j_s=0}^{N+1}\tilde{A}_{\boldsymbol{j}}dx_i(t-\boldsymbol{j}\boldsymbol{\tau}) \;&=\; \sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{\max_s j_s=N-l}^{M-l}\tilde{B}_{\boldsymbol{j}kl}dy_k^{(l)}(t-\boldsymbol{j}\boldsymbol{\tau})- \\ &\quad -\sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j_s=0}^{M}\bar{E}_{\boldsymbol{j}kl}du_k^{(l)}(t-\boldsymbol{j}\boldsymbol{\tau}).\end{aligned} \tag{4.76}$$

The above equation can be written

$$\begin{aligned}\sum_{\min_s j_s=1}^{N+1}\tilde{A}_{\boldsymbol{j}}dx_i(t-\boldsymbol{j}\boldsymbol{\tau})&=\sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{\min_s j_s=1}^{M-l}\tilde{B}_{\boldsymbol{j}kl}dy_k^{(l)}(t-\boldsymbol{j}\boldsymbol{\tau})- \\ -\sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{\min_s j_s=1}^{M}\bar{E}_{\boldsymbol{j}kl}du_k^{(l)}(t-\boldsymbol{j}\boldsymbol{\tau})&+\Big[\sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{\min_s j_s=0}\tilde{B}_{\boldsymbol{j}kl}dy_k^{(l)}(t-\boldsymbol{j}\boldsymbol{\tau})- \\ -\sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{\min_s j_s=0}\bar{E}_{\boldsymbol{j}kl}du_k^{(l)}(t-\boldsymbol{j}\boldsymbol{\tau})&-\sum_{\min_s j_s=0}\tilde{A}_{\boldsymbol{j}}dx_i(t-\boldsymbol{j}\boldsymbol{\tau})\Big].\end{aligned} \tag{4.77}$$

The terms in the square brackets must sum up to zero and thus,

$$\sum_{\min_s j_s=1}^{N+1} \tilde{A}_{\boldsymbol{j}} dx_i(t-\boldsymbol{j\tau}) = \sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{\min_s j_s=1}^{M-l} \tilde{B}_{\boldsymbol{j}kl} dy_k^{(l)}(t-\boldsymbol{j\tau}) - \\ - \sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{\min_s j_s=1}^{M} \bar{E}_{\boldsymbol{j}kl} du_k^{(l)}(t-\boldsymbol{j\tau}) \quad . \tag{4.78}$$

This can be written

$$\delta_1 \dots \delta_\ell \Big(\sum_{j_s=0}^{N} \delta_1^{-1} \dots \delta_\ell^{-1}(\tilde{A}_{\boldsymbol{j}}) dx_i(t-\boldsymbol{j\tau}) \Big) = \\ = \delta_1 \dots \delta_\ell \Big(\sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j_s=0}^{M-1-l} \delta_1^{-1} \dots \delta_\ell^{-1}(\tilde{B}_{\boldsymbol{j}kl}) dy_k^{(l)}(t-j\tau) + \\ - \sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j_s=0}^{M-1} \delta_1^{-1} \dots \delta_\ell^{-1}(\bar{E}_{\boldsymbol{j}kl}) du_k^{(l)}(t-j\tau) \Big) \quad ,$$

which is equivalent to

$$\sum_{j_s=0}^{N} \delta_1^{-1} \dots \delta_\ell^{-1}(\tilde{A}_{\boldsymbol{j}}) dx_i(t-j\tau) = \\ = \sum_{k=1}^{p}\sum_{l=0}^{n-1}\sum_{j_s=0}^{M-1-l} \delta_1^{-1} \dots \delta_\ell^{-1}(\tilde{B}_{\boldsymbol{j}kl}) dy_k^{(l)}(t-j\boldsymbol{\tau}) + \\ - \sum_{k=1}^{m}\sum_{l=0}^{n-2}\sum_{j_s=0}^{M-1} \delta_1^{-1} \dots \delta_\ell^{-1}(\bar{E}_{\boldsymbol{j}kl}) du_k^{(l)}(t-j\boldsymbol{\tau}) \quad , \tag{4.79}$$

where all $\delta_1^{-1} \dots \delta_\ell^{-1}(\tilde{A}_{\boldsymbol{j}})$, $\delta_1^{-1} \dots \delta_\ell^{-1}(\tilde{B}_{\boldsymbol{j}kl})$ and $\delta_1^{-1} \dots \delta_\ell^{-1}(\bar{E}_{\boldsymbol{j}kl})$ are in $\mathcal{K}$. Thus, we obtain a contradiction to the assumption in this case as well. By the principle of induction, we now have that for the one-forms $dx_{i,[P]}(t)$, $P \geq n$, there can be no nonzero linear combination of them which is in $\mathcal{S}^{Q,Q}$ for $Q \geq P$. Therefore, $dx_i \notin \overline{\mathcal{Y}_n + \mathcal{U}}$ and x_i is not weakly observable. This completes the proof of the proposition. ■

We will now demonstrate the calculations on the simple example from Paragraph 4.2.2.2 and also show how the analysis can be automated, allowing for the application of symbolic computation algorithms.

Example 1: An observable system

Consider the system

$$\left\{\begin{array}{lcl} \dot{x}_1 & = & x_2^2(t-\tau_1)x_1(t-\tau_2)+u \\ \dot{x}_2 & = & x_1(t-\tau_2) \\ y(t) & = & x_1(t-\tau_1) \quad . \end{array}\right. \tag{4.80}$$

Taking n time-derivatives of the output function, we obtain the equations

$$y = \delta_1 x_1 \tag{4.81}$$

$$\dot{y} = (\delta_1^2 x_2)^2 \delta_1\delta_2 x_1 + \delta_1 u \quad . \tag{4.82}$$

For the variables occurring above, we can get two more equations by considering shifts of the first equation:

$$\delta_1 y = \delta_1^2 x_1, \qquad \delta_2 y = \delta_1\delta_2 x_1 \quad . \tag{4.83}$$

We will now rename all x and $y, \dot{y}$ variables occurring in the equations above according to the following scheme:

$$\begin{array}{lll} z_{1(0,0)} := x_1, & z_{1(1,0)} := \delta_1 x_1, & z_{1(2,0)} := \delta_1^2 x_1 \\ z_{1(0,1)} := \delta_2 x_1, & z_{1(0,2)} := \delta_2^2 x_1, & z_{1(1,1)} := \delta_1\delta_2 x_1 \\ z_{2(0,0)} := x_2, & z_{2(1,0)} := \delta_1 x_2, & z_{2(2,0)} := \delta_1^2 x_2 \\ z_{2(0,1)} := \delta_2 x_2, & z_{2(0,2)} := \delta_2^2 x_2, & z_{2(1,1)} := \delta_1\delta_2 x_2 \end{array} \tag{4.84}$$

and

$$\begin{array}{lll} y_{0(0,0)} := y, & y_{0(1,0)} := \delta_1 y, & y_{0(0,1)} := \delta_2 y \\ y_{1(0,0)} := \dot{y}, & & \end{array} \tag{4.85}$$

and analogously for the occurring u-variables.

We therefore have the following algebraic system of 4 equations for the 4 variables z:

$$\begin{array}{rcl} y_{0(0,0} & = & z_{1(1,0)} \\ y_{0(1,0)} & = & z_{1(2,0)} \\ y_{0(0,1)} & = & z_{1(1,1)} \\ y_{1(0,0)} & = & z_{2(2,0)}^2 z_{1(1,1)} + u_{0(1,0)} \quad . \end{array} \tag{4.86}$$

Consider the Jacobian $J = [J_1|J_2]$ (J_i consists of the columns of J corresponding to $dz_{i()}$ of the algebraic system above

$$\left[\begin{array}{c} dy_{0(0,0} \\ y_{0(1,0)} \\ y_{0(0,1)} \\ y_{1(0,0)} - du_{0(1,0)} \end{array}\right] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & z_{2(2,0)}^2 & 2z_{2(2,0)}z_{1(1,1)} \end{array}\right] \left[\begin{array}{c} dz_{1(1,0)} \\ dz_{1(2,0)} \\ dz_{1(1,1)} \\ dz_{2(2,0)} \end{array}\right] \tag{4.87}$$

We see that the rank of the matrix is 4. If we remove the first 3 columns corresponding to $z_{1(1,0)}, z_{1(2,0)}, z_{1(1,1)}$, it becomes 1, and if we remove the second group of 1 column, the one for $z_{2(2,0)}$ we get rank 3. Thus both x_1 and x_2 are observable.

4.4 Identification of a simple example system

In this subsection we will see how the initial function can be identified from the output for a simple example system. The calculations also give an indication of the least amount of time over which the output has to be observed for the system to be identified, as discussed in Verduyn Lunel (2001). For this subsection, we will restrict the class of initial functions to analytic ones.

Consider a simple example, taken from Verduyn Lunel (2001):

$$\begin{cases} \dot{x}(t) &= ax(t) + bx(t-1) \\ y(t) &= x(t) \\ x(t) &= \varphi(t), \quad t \in [-1, 0] \quad . \end{cases} \tag{4.88}$$

Let us first verify that this system is weakly observable. Taking the first two time-derivatives of the output, we obtain the following equations for the three unknowns:

$$y = x \tag{4.89}$$
$$\dot{y} = ax + b\delta(x) \tag{4.90}$$
$$\ddot{y} = a(ax + b\delta(x)) + b(a\delta(x) + b\delta^2(x)) \quad . \tag{4.91}$$

Thus,

$$\begin{bmatrix} dy \\ d\dot{y} \\ d\ddot{y} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ a + b\delta & x & \delta(x) \\ a^2 + 2ab\delta + b^2\delta^2 & 2ax + 2b\delta(x) & 2b\delta(x) + 2b\delta^2(x) \end{bmatrix} \begin{bmatrix} dx \\ da \\ db \end{bmatrix} \tag{4.92}$$

Row elimination over $\mathcal{K}(\delta]$ gives:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & x & \delta(x) \\ 0 & 2ax + 2b\delta(x) & 2b\delta(x) + 2b\delta^2(x) \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & x & \delta(x) \\ 0 & 0 & 2(b-a)\delta(x) \end{bmatrix} \tag{4.93}$$

The last matrix is generically of full rank and the system is weakly observable. We now proceed to the actual identification of the two parameters a and b and the initial function φ. Evaluating the output time-derivatives at time 1

gives

$$\begin{aligned} y(1) &= x(1) && (4.94)\\ \dot{y}(1) &= ax(1) + bx(0) = ay(1) + by(0) && (4.95)\\ \ddot{y}(1) &= ax(1) + bx(0) = a\dot{y}(1) + b\dot{y}(0) \quad . && (4.96) \end{aligned}$$

and this yields unique values for the parameters a and b:

$$\begin{aligned} a &= \frac{\dot{y}(0)\dot{y}(1) - y(0)\ddot{y}(1)}{y(1)\dot{y}(0) - y(0)\dot{y}(1)} && (4.97)\\ b &= \frac{-\dot{y}(1)^2 + y(1)\ddot{y}(1)}{y(1)\dot{y}(0) - y(0)\dot{y}(1)} \quad . && (4.98) \end{aligned}$$

To obtain the initial function, we use the equations

$$\begin{aligned} y(0) &= \varphi(0) && (4.99)\\ \dot{y}(0) &= ay(0) + b\varphi(-1) && (4.100)\\ \ddot{y}(0) &= a\dot{y}(0) + b\dot{\varphi}(-1) && (4.101)\\ &\vdots \end{aligned}$$

which give,

$$\varphi^{(j)}(-1) = \frac{y^{(j+1)}(0) - ay^{(j)}(0)}{b} \quad , \qquad (4.102)$$

for $j \geq 0$. Since φ is assumed to be analytic, we have

$$\varphi(t) = -\frac{a}{b}y(0) + \frac{1}{b}\sum_{j=1}^{\infty}\Big(\frac{(t+1)^{j-1}}{(j-1)!} - \frac{a(t+1)^j}{j!}\Big)y^{(j)}(0) \qquad (4.103)$$

and the initial function is uniquely defined, too. The last equality also gives a relation for a and b:

$$\begin{aligned} y(0) &= \varphi(0) = -\frac{a}{b}y(0) + \frac{1}{b}\sum_{j=1}^{\infty}\Big(\frac{1}{(j-1)!} - \frac{a}{j!}\Big)y^{(j)}(0)\\ &\Updownarrow\\ b &= -a + \frac{1}{y(0)}\sum_{j=1}^{\infty}(\frac{1}{(j-1)!} - \frac{a}{j!})y^{(j)}(0) \quad . && (4.104) \end{aligned}$$

Observe that the values of the parameters a and b were obtained using values of the output and its derivatives at time 1. In fact, unless the output is observed over an interval with length greater than 1, the system cannot be

identified (see (Verduyn Lunel, 2001) for a more general analysis of linear systems). To illustrate, fix $a^0 = b^0 = 1$, $\varphi^0(t) = 1$. Then, $y(0) = \varphi^0(0) = 1$ and $y^{(j)}(0) = 2$ from (4.89)-(4.91). We will now find another set a^1, b^1 and φ^1 which produces the same output. Equation (4.104) gives

$$\begin{aligned} b^1 &= -a^1 + \sum_{j=1}^{\infty}(\frac{1}{(j-1)!} - \frac{a^1}{j!})2 = -a^1 + 2\sum_{j=1}^{\infty}\frac{1}{(j-1)!} - 2a^1\sum_{j=1}^{\infty}\frac{1}{j!} = \\ &= -a^1 + 2e - 2a^1(e-1) = 2e + a^1(1-2e) \quad . \end{aligned} \tag{4.105}$$

We have

$$\begin{aligned} \varphi^1(t) &= -\frac{a^1}{b^1} + \frac{2}{b^1}\sum_{j=1}^{\infty}\Big(\frac{(t+1)^{j-1}}{(j-1)!} - \frac{a^1(t+1)^j}{j!}\Big) = \\ &= -\frac{a^1}{b^1} + \frac{2}{b^1}(e^{t+1} - a^1(e^{t+1}-1)) = \frac{2(1-a^1)}{b^1}e^{t+1} + \frac{a^1}{b^1} \end{aligned} \tag{4.106}$$

For any value of a^1, choosing b^1 and φ^1 according to the above equations will give exactly the same output in the interval $[0, 1]$. For example, one can choose $a^1 = 0$, $b^1 = 2e$ and $\varphi^1(t) = e^t$ as in Verduyn Lunel (2001), or $a^1 = 2$, $b^1 = 2 - 2e$ and $\varphi^1(t) = \frac{1-e^{t+1}}{1-e}$.

As an illustration, we have simulated the output for this system for two different choices of a, b and φ, see Figure 1.

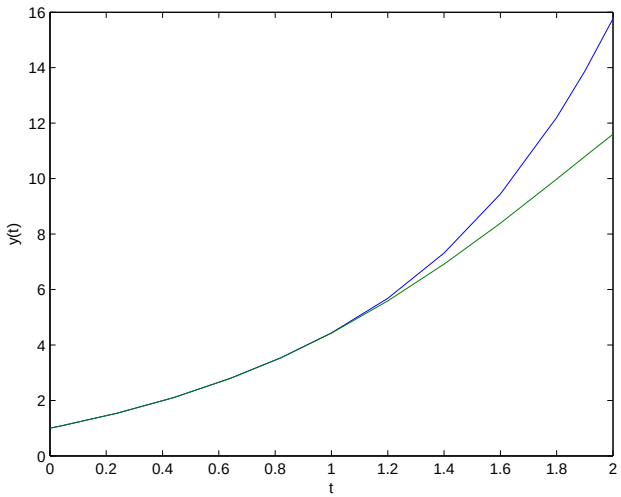


Figure 1: *The output $y(t)$ for the choices $a^0 = b^0 = 1$, $\varphi^0 = 1$ and $a^1 = 2$, $b^1 = 2 - 2e$ and $\varphi^1(t) = \frac{1-e^{t+1}}{1-e}$.*

In conclusion, for delay systems, there are lower limits to the interval of time over which the output has to be observed for the identification of the state-variables and parameters of an observable/identifiable system. This lower limit applies even in the theoretical error-free data case, and thus presents a difference between delay- and ODE systems.

5 State elimination

5.1 Existence of input-output representation

In this section we consider the problem of obtaining an input-output representation from the state-space form of a time-delay control system. This problem has been treated for polynomial systems by Forsman et al. (1994).

We will show that for a system of the form

$$\left\{\begin{array}{rcl} \dot{x}(t) &=& f\big(x(t), x(t-\tau_1), \ldots, x(t-\tau_\ell), u, u(t-\tau_1), \ldots, u(t-\tau_\ell)\big) \\ y(t) &=& h\big(x(t), x(t-\tau_1), \ldots, x(t-\tau_\ell)\big) \\ x(t) &=& \varphi(t), \quad t \in [-\max_i \tau_i, 0] \\ u(t) &=& u_0(t), \quad t \in [-T, 0] \quad , \end{array}\right. \tag{5.1}$$

there always exists, at least locally, a set of input-output differential equations of the form:

$$F(\boldsymbol{\delta}, y, \ldots, y^{(k)}, u, \ldots, u^{(J)}) := \\ = F(y(t - \boldsymbol{i_0\tau}), \ldots, y^{(k)}(t - \boldsymbol{i_k\tau}), u(t - \boldsymbol{j_0\tau}), \ldots, u^{(J)}(t - \boldsymbol{j_l\tau})) = 0\,,$$

such that any pair $(y(t), u(t))$ which solves the original system, also satisfies (5.2). The function F is meromorphic in its arguments.

Theorem 5.1 *There exists an integer $J \geq 0$ and an open dense subset V of $C \times C_U^{J+1}$, such that in the neighborhood of any point of V, there exists an input-output representation of the system of the form* (5.2).

Proof: The proof is an adaptation of the proof of Theorem 2.2.1. in Conte et al. (1999) for the analogous result for ODE-systems.

Let f be an r-dimensional vector with entries $f_j \in \mathcal{K}$. Let $\frac{\partial f}{\partial x}$ denote the $r \times n$ matrix with entries

$$\left(\frac{\partial f}{\partial x}\right)_{j,i} = \sum_{\boldsymbol{k}} \frac{\partial f_j}{\partial x_i(t - \boldsymbol{k\tau})} \boldsymbol{\delta^k} \;\in\; \mathcal{K}(\boldsymbol{\delta}] \quad . \tag{5.2}$$

Denote by s_1 the least nonnegative integer such that

$$\mathrm{rank}_{\mathcal{K}(\delta]} \frac{\partial(h_1, \ldots, h_1^{(s_1-1)})}{\partial x} = \mathrm{rank}_{\mathcal{K}(\boldsymbol{\delta}]} \frac{\partial(h_1, \ldots, h_1^{(s_1)})}{\partial x} \quad . \tag{5.3}$$

If $\frac{\partial h_1}{\partial x} \equiv 0$ then we define $s_1 = 0$. Inductively, for $1 < i \leq p$ denote by s_i the least nonnegative integer such that

$$\mathrm{rank}_{\mathcal{K}(\boldsymbol{\delta}]} \frac{\partial(h_1, \ldots, h_1^{(s_1-1)}, \ldots, h_i, \ldots, h_i^{(s_i-1)})}{\partial x} = \\ = \mathrm{rank}_{\mathcal{K}(\boldsymbol{\delta}]} \frac{\partial(h_1, \ldots, h_1^{(s_1)}, \ldots, h_i, \ldots, h_i^{(s_i)})}{\partial x} \quad . \tag{5.4}$$

Let

$$S = (h_1, \ldots, h_1^{(s_1-1)}, \ldots, h_p, \ldots, h_p^{(s_p-1)}) \quad , \tag{5.5}$$

where h_i does not appear if $s_i = 0$ (unless all h_i:s are constants, at least one s_i will be greater than zero). Then

$$\text{rank}_{\mathcal{K}(\boldsymbol{\delta})} \frac{\partial S}{\partial x} = s_1 + \cdots + s_p = K \leq n \quad . \tag{5.6}$$

If $K < n$, there exists a set of meromorphic functions $g_1(\boldsymbol{\delta}, x), \ldots, g_{n-K}(\boldsymbol{\delta}, x)$ such that $\text{rank}_{\mathcal{K}(\boldsymbol{\delta})} \frac{\partial (S, g_1, \ldots, g_{n-K})}{\partial x} = n$.

For simplicity, introduce the notation

$$\left\{ \begin{array}{lcl} \tilde{x}_1 & = & h_1 \\ & \vdots & \\ \tilde{x}_{s_1} & = & h_1^{(s_1-1)} \\ \tilde{x}_{s_1+1} & = & h_2 \\ & \vdots & \\ \tilde{x}_{s_1+s_2} & = & h_2^{(s_2-1)} \\ & \vdots & \\ \tilde{x}_{s_1+\cdots+s_p} & = & h_p^{(s_p-1)} \\ \tilde{x}_{s_1+\cdots+s_p+1} & = & g_1 \\ & \vdots & \\ \tilde{x}_n & = & g_{n-K} \quad . \end{array} \right. \tag{5.7}$$

The one-forms $d\tilde{x}_i, \quad i = 1, \ldots, n$ now form a basis of $\mathcal{X}$. The definition of s_i (equation (5.4)), together with part (3) of Lemma 2, implies that $\frac{\partial h_i^{(s_i)}}{\partial x}$ is in

$$\overline{\text{span}_{\mathcal{K}(\boldsymbol{\delta})} \left\{ \frac{\partial (h_1, \ldots, h_1^{(s_1-1)}, \ldots, h_i, \ldots, h_i^{(s_i-1)})}{\partial x} \right\}} \quad . \tag{5.8}$$

Thus there exist nonzero polynomials $b_i(\boldsymbol{\delta}] \in \mathcal{K}(\boldsymbol{\delta}]$, $i = 1, \ldots, p$ such that $b_i(\boldsymbol{\delta}] \frac{\partial h_i^{(s_i)}}{\partial x}$ is in

$$\text{span}_{\mathcal{K}(\boldsymbol{\delta})} \left\{ \frac{\partial (h_1, \ldots, h_1^{(s_1-1)}, \ldots, h_i, \ldots, h_i^{(s_i-1)})}{\partial x} \right\} \quad . \tag{5.9}$$

Therefore

$$b_i(\boldsymbol{\delta}] dh_i^{(s_i)} + \sum_{r=1}^{m} \sum_{j=0}^{J} c_{j,r}(\boldsymbol{\delta}] du_r^{(j)} \in \text{span}_{\mathcal{K}(\boldsymbol{\delta}]} \{ d\tilde{x}_1, \ldots, d\tilde{x}_{s_1+\cdots+s_i} \} \quad , \tag{5.10}$$

for some $J \geq 0$, where J is the highest derivative of u appearing in the functions in S and $c_{j,r}(\boldsymbol{\delta}] \in \mathcal{K}(\boldsymbol{\delta}]$.

Hence,

$$b_i(\boldsymbol{\delta}]dh_i^{(s_i)} + \sum_{r=1}^{m}\sum_{j=0}^{\gamma} c_{j,r}(\boldsymbol{\delta}]du_r^{(j)} - \sum_{j=1}^{s_1+\cdots+s_i} a_j(\boldsymbol{\delta}]d\tilde{x}_j = 0 \quad , \tag{5.11}$$

for some $a_j(\boldsymbol{\delta}] \in \mathcal{K}(\boldsymbol{\delta}]$. Since all functions are assumed meromorphic and we have continuous dependence for the output on the input and initial function, the above equality holds on an open dense set of $C \times C_U^{J+1}$.

The left hand side of equation (5.11), being equal to zero, is a closed one-form on $\mathcal{M}$. Applying the Poincaré lemma (see subsection 3.2.4), we obtain functions $\xi_i(t) \in \mathcal{K}$ such that

$$d\xi_i = b_i(\boldsymbol{\delta}]dh_i^{(s_i)} + \sum_{r=1}^{m}\sum_{j=0}^{J} c_{j,r}(\boldsymbol{\delta}]du_r^{(j)} - \sum_{j=1}^{s_1+\cdots+s_i} a_j(\boldsymbol{\delta}]d\tilde{x}_j$$

and

$$\xi_i(\boldsymbol{\delta}, h_i^{(s_i)}, \tilde{x}, u, \ldots, u^{(J)}) = 0 \quad ,$$

for each $i = 1, \ldots, p$.

The function ξ_i does not depend on $\tilde{x}_j$, $j > s_1 + \cdots + s_i$, since $d\xi_i = 0$ would then contain terms $d\tilde{x}_j$, $j > s_1 + \cdots + s_i$ which is impossible by (5.11) due to the variables $d\tilde{x}$ being linearly independent over $\mathcal{K}(\boldsymbol{\delta}]$ by definition. Thus, we have obtained a relation

$$\xi_i(\boldsymbol{\delta}, h_i^{(s_i)}, \tilde{x}_1, \ldots, \tilde{x}_{s_1+\cdots+s_i}, u, \ldots, u^{(J)}) = 0 \quad , \tag{5.12}$$

which together with (5.7) produces an input-output equation

$$\xi_i(\boldsymbol{\delta}, y_i^{(s_i)}, y_1, \ldots, y_1^{(s_1-1)}, y_i, \ldots, y_i^{(s_i-1)}, u, \ldots, u^{(J)}) = 0 \quad . \tag{5.13}$$

This is true for each $i, 1 \leq i \leq p$ resulting in p input-output equations of the form (5.2). ■

5.2 Alternative derivation

Here we obtain the same result using an algebraic system derived from the original delay-differential equations. This has been done for polynomial systems by Forsman et al. (1994).

5.2.1 Single delay

Similarly to what was done in Section 4.3, consider the $(n+1)p$ equations for the state-variables obtained by taking n time-derivatives of the p output variables.

$$\begin{aligned}
y_1(t) &= h_1(x(t), x(t-\tau)) && (5.14)\\
&\vdots\\
y_1^{(n)}(t) &= h_1^{(n)}(x(t), x(t-\tau), \ldots, x(t-(n+1)\tau), && (5.15)\\
&\quad u(t), \ldots, u^{(n-1)}(t), \ldots, u(t-n\tau), \ldots, u^{(n-1)}(t-n\tau))\\
&\vdots\\
y_p(t) &= h_p(x(t), x(t-\tau)) && (5.16)\\
&\vdots\\
y_p^{(n)}(t) &= h_p^{(n)}(x(t), x(t-\tau), \ldots, x(t-(n+1)\tau)), && (5.17)\\
&\quad u(t), \ldots, u^{(n-1)}(t), \ldots, u(t-n\tau), \ldots, u^{(n-1)}(t-n\tau)) \quad .
\end{aligned}$$

For the $n(n+2)$ variables $x_1(t), \ldots, x_1(t-(n+1)\tau), \ldots, x_n(t), \ldots, x_n(t-(n+1)\tau)$ occurring above, we can obtain more relations by taking shifts of the equations above to obtain a system with $\frac{p(n+1)(n+2)}{2}$ equations. For each $k = 1, \ldots, p$ we have the $\frac{(n+1)(n+2)}{2}$ equations:

$$\begin{aligned}
y_k(t) &= h_k(x(t), x(t-\tau)) && (5.18)\\
&\vdots\\
\delta^n y_k(t) &= h_k(x(t-n\tau), x(t-(n+1)\tau) && (5.19)\\
&\vdots\\
y_k^{(n-1)}(t) &= h_k^{(n-1)}(x(t), \ldots, x(t-n\tau)), u(t), \ldots, u^{(n-1)}(t), && (5.20)\\
&\quad \ldots, u(t-(n-1)\tau), \ldots, u^{(n-1)}(t-(n-1)\tau))\\
\delta y_k^{(n-1)}(t) &= h_k^{(n-1)}(x(t-\tau), \ldots, x(t-(n+1)\tau)), && (5.21)\\
&\quad u(t-\tau), \ldots, u^{(n-1)}(t-\tau), \ldots, u(t-n\tau), \ldots, u^{(n-1)}(t-n\tau))\\
y_k^{(n)}(t) &= h_k^{(n)}(x(t), x(t-\tau), \ldots, x(t-(n+1)\tau), && (5.22)\\
&\quad u(t), \ldots, u^{(n-1)}(t), \ldots, u(t-n\tau), \ldots, u^{(n-1)}(t-n\tau)) \quad .
\end{aligned}$$

In equation system (5.18)-(5.23), taking one more shift in $y_k^{(l)}$, $l = 0, \ldots, n$ leads to n new variables (one new shift of each x_i) but adds $n+1$ new equations for them. By choosing q as the least integer larger than $\frac{n^2-n+2}{2}$, and shifting each $y_k^{(l)}(t)$ $q-l$ times, we get a set of equations for the $n(q+2)$

variables $x_i(t), \ldots, x_i(t-(q+1)\tau)$, $i = 1, \ldots, n$ which is larger than the number of unknowns

$$\begin{aligned}
y_k(t) &= h_k(x(t), x(t-\tau)) && (5.23)\\
&\vdots\\
\delta^q y_k(t) &= h_k(x(t-q\tau), x(t-(q+1)\tau)) && (5.24)\\
&\vdots\\
y_k^{(n)}(t) &= h_k^{(n)}(x(t), x(t-\tau), \ldots, x(t-(n+1)\tau), && (5.25)\\
&\quad u(t), \ldots, u^{(n-1)}(t), \ldots, u(t-n\tau), \ldots, u^{(n-1)}(t-n\tau))\\
&\vdots\\
\delta^{q-n} y_k^{(n)}(t) &= h_k^{(n)}(x(t-(q-n)\tau), \ldots, x(t-(q+1)\tau), && (5.26)\\
&\quad u(t-(q-n)\tau), \ldots, u^{(n-1)}(t-(q-n)\tau), \ldots, u(t-q\tau),\\
&\quad \ldots, u^{(n-1)}(t-q\tau)) \quad .
\end{aligned}$$

The number of equations is now $(n+1)(2q-n+2)/2 > n(q+2)$ for the chosen integer q. Thus, there is a linear dependence among the one-forms $dy_k^{(l)}(t-j\tau), l = 0, \ldots, n, j = 0, \ldots, q-l$ and $du_i^{(l)}(t-j\tau), i = 1, \ldots, m, l = 0, \ldots, n-1, j = 0, \ldots, q$ for each k. Applying the Poincaré Lemma we obtain meromorphic functions ξ_k, $k = 1, \ldots, p$ such that

$$\xi_k(u, y) = 0, \tag{5.27}$$

where u, y denotes all the variables $y_k^{(l)}(t-j\tau), l = 0, \ldots, n, j = 0, \ldots, k-l$ and $u_i^{(l)}(t-j\tau), i = 1, \ldots, m, l = 0, \ldots, n-1, j = 0, \ldots, q$ for a given k. We thus obtain p input-output equations for the system (3.1).

5.2.2 Multiple delays

For completeness, we verify that the alternative derivation also works for the multiple delay case. Consider once again the system

$$\left\{\begin{array}{lcl}
\dot{x}(t) &=& f\big(x(t), x(t-\tau_1), \ldots, x(t-\tau_\ell), u, u(t-\tau_1), \ldots, u(t-\tau_\ell)\big)\\
y(t) &=& h\big(x(t), x(t-\tau_1), \ldots, x(t-\tau_\ell)\big)\\
x(t) &=& \varphi(t), \quad t \in [-\max_i \tau_i, 0]\\
u(t) &=& u_0(t), \quad t \in [-T, 0]
\end{array}\right. \tag{5.28}$$

and the $(n+1)p$ equations for the state-variables obtained by taking n time-derivatives of the p output variables:

$$\begin{aligned}
y_1(t) &= h_1\big(x_{[1]}(t)\big) && (5.29)\\
&\vdots\\
y_1^{(i)}(t) &= h_1\big(x_{[i+1]}(t), u_{[i]}(t), u_{[i]}^{(i-1)}(t)\big) && (5.30)\\
&\vdots\\
y_1^{(n)}(t) &= h_1^{(n)}\big(x_{[n+1]}(t), u_{[n]}(t), \ldots, u_{[n]}^{(n-1)}(t)\big) && (5.31)\\
&\vdots\\
y_p(t) &= h_p\big(x_{[1]}(t)\big) && (5.32)\\
&\vdots\\
y_p^{(n)}(t) &= h_p^{(n)}\big(x_{[n+1]}(t), u_{[n]}(t), \ldots, u_{[n]}^{(n-1)}(t)\big) \quad , && (5.33)
\end{aligned}$$

where, as in Section 4.3, $x_{[i]}(t)$ denotes all the variables of type $x(t-\sum_{j=1}^{\ell} i_j\tau_j)$, $i_j \in \mathbb{Z}^+$, $\sum_{j=1}^{\ell} i_j \le i$ and analogously for u, y and their derivatives.

For the $n\binom{n+1+\ell}{\ell}$ variables $x_{1,[n+1]}(t), \ldots, x_{n,[n+1]}(t)$ occurring above, we can obtain more relations by taking shifts by the different δ_i of the equations above to obtain a system with $p\sum_{i=1}^{n+1}\binom{i}{\ell}$ equations. For each $k = 1, \ldots, p$ we have the $\sum_{i=1}^{n+1}\binom{i}{\ell}$ equations

$$\begin{aligned}
y_k(t) &= h_k\big(x_{[1]}(t)\big)\\
\delta_1 y_k(t) &= h_k\big(x_{[1]}(t-\tau_1)\big)\\
&\vdots\\
\delta_1^{i_1}\ldots\delta_\ell^{i_\ell} y_k(t) &= h_k\big(x_{[n+1]}(t)\big), \quad i_j \ge 0, \quad \sum_{j=1}^{\ell} i_j \le n\\
&\vdots\\
y_k^{(n-1)}(t) &= h_k^{(n-1)}\big(x_{[n]}(t), u_{[n-1]}(t), \ldots, u_{[n-1]}^{(n-2)}(t)\big) && (5.34)\\
\delta_1 y_k^{(n-1)}(t) &= h_k^{(n-1)}\big(x_{[n]}(t-\tau_1), u_{[n-1]}(t-\tau_1), \ldots, u_{[n-1]}^{(n-2)}(t-\tau_1)\big)\\
&\vdots\\
\delta_\ell y_k^{(n-1)}(t) &= h_k^{(n-1)}\big(x_{[n]}(t-\tau_\ell), u_{[n-1]}(t-\tau_\ell), \ldots, u_{[n-1]}^{(n-2)}(t-\tau_\ell)\big)\\
y_k^{(n)}(t) &= h_k^{(n)}(x_{[n+1]}(t), u_{[n]}(t), \ldots, u_{[n]}^{(n-1)}(t)\big) \quad .
\end{aligned}$$

Starting from the equation system above, for each $k = 1, \ldots, p$, taking one more shift by each τ_i, $i = 1, \ldots, \ell$ in each $y_k^{(l)}$, $l = 0, \ldots, n$ leads to $n\ell$ new

variables (ℓ new shifts of each x_i) but adds $(n+1)\ell$ new equations for them. Clearly, by repeating this procedure enough times, we will get a system with more equations than unknowns. Hence, for each $k = 1, \ldots, p$, we will then have a relation among the different $y^{(l)}_{k,[j]}(t)$, $l = 0, \ldots, n$, $j \in \mathbb{N}$ and $du^{(r)}_{i,[s]}(t)$, $i = 1, \ldots, m$, $r = 0, \ldots, n-1$, $s \in \mathbb{N}$. Each such relation is an input-output equation for the system (5.28).

6 Identifiability of time-lag parameters

6.1 Identifying delay parameters from explicit input-output equations

In the previous section, it was shown that external input-output representation exists for time-delay systems with discrete time delays. If the input-output equations can be obtained explicitly, the time-lags can in some cases be calculated directly from data of the output and its time-derivatives. In this section we demonstrate how this can be done on simple examples.

6.1.1 Single delay

Example 1:
We start by an example of a simple delay model with a corresponding input-output structure for the purpose of parameter estimation. The model has two state variables, x_1 and x_2, two regular parameters k_1 and k_2, and one time-lag parameter, τ. The controlled input variable is denoted by u and the measured data by y. Initial conditions have to be defined for the input, $u_0(t)$, and the and state variables, $\varphi(t)$. In the case of delay differential equations, these are functions describing the history of the system at least τ units back in time. All parameters and initial conditions for the state variables are unknown, but we assume that τ must take values in the interval $[0, T]$, where T is known.

$$\begin{cases} \dot{x}_1(t) &= k_1 x_2(t-\tau) + u(t) \\ \dot{x}_2(t) &= k_2 x_2(t-\tau) \\ y(t) &= x_1(t) \\ x(t) &= \varphi(t), \quad t \in [-\tau, 0] \\ u(t) &= u_0(t), \quad t \in [-T, 0] \quad . \end{cases} \tag{6.1}$$

By taking time derivatives of the output at a given point in time, we obtain equations for the state variables and parameters:

$$\dot{y}(t) = k_1 x_2(t-\tau) + u(t) \tag{6.2}$$

$$\ddot{y}(t) = k_1 k_2 x_2(t-2\tau) + \dot{u}(t) \tag{6.3}$$

$$y^{(3)}(t) = k_1 k_2^2 x_2(t-3\tau) + \ddot{u}(t) \quad . \tag{6.4}$$

From the above equations, we can extract an external input-output representation of the system, given by the input-output equation

$$\big(\ddot{y}(t) - \dot{u}(t)\big)\big(\ddot{y}(t-\tau) - \dot{u}(t-\tau)\big) - \big(y^{(3)}(t) - \ddot{u}(t)\big)\big(\dot{y}(t-\tau) - u(t-\tau)\big) = 0. \tag{6.5}$$

Since the input and output are known functions of time, evaluating the above equation at a chosen point t_0 in time enables us to calculate the time-lag parameter τ, for example, by numerically finding the zeros of the function

$$\begin{aligned}\xi_{t_0}(\tau) &= \big(\ddot{y}(t_0) - \dot{u}(t_0)\big)\big(\ddot{y}(t_0-\tau)) - \dot{u}(t_0-\tau)\big) - \\ &\quad - \big(y^{(3)}(t_0) - \ddot{u}(t_0)\big)\big(\dot{y}(t_0-\tau) - u(t_0-\tau)\big)\end{aligned}$$

of τ. Note that t_0 should be chosen so as to ensure the continuity of all involved derivatives, in this case $t_0 \geq 3T$.

As an illustration, we have used the `dde23.m` differential equation solver in Matlab (Shampine and Thompson, 2001) to simulate an output for the above system. We chose $k_1 = -2, k_2 = -3, \varphi_1(t) = t + 1, \varphi_2(t) = t^2 + 1, u(t) = t$ and $\tau = 1$ and plotted the output of the system (left-hand side) and the function

$$\xi_6(\tau) = \big(\ddot{y}(6) - \dot{u}(6)\big)\big(\ddot{y}(6-\tau) - \dot{u}(6-\tau)\big) - \big(y^{(3)}(6) - \ddot{u}(6)\big)\big(\dot{y}(6-\tau) - u(6-\tau)\big)$$

for τ in the interval $[0, 2]$, see Figure 2. As expected, this function takes the value zero for $\tau = 1$.

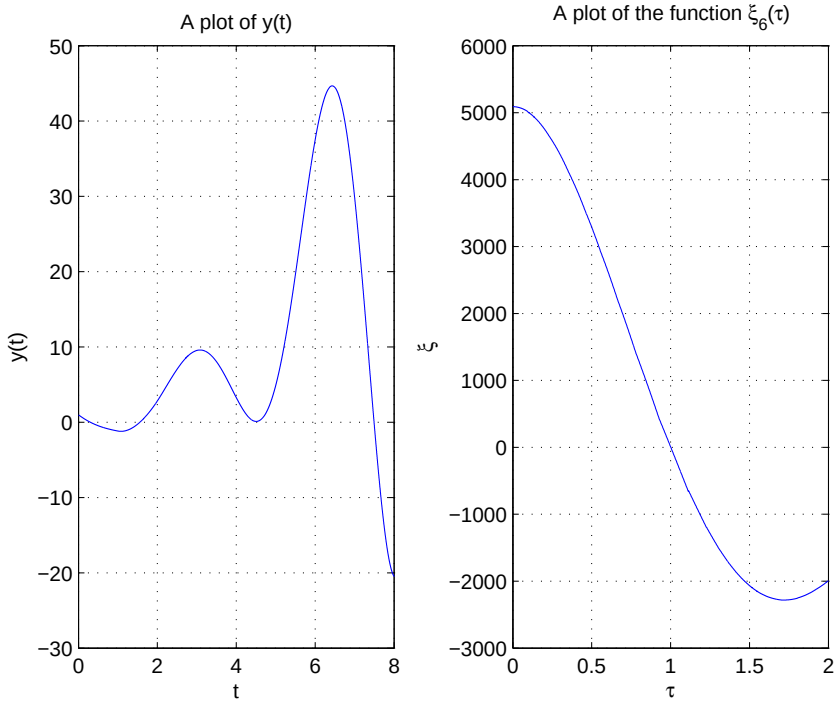


Figure 2: *The output $y(t)$ and the function $\xi(\tau)$.*

Of course, to use this identification procedure, one needs to be able to measure the output time-derivatives with reasonable accuracy. Methods for the estimation of output derivatives can be found, for example in (Mboup et al., 2007).

On the other hand, it can be the case that the input-output equations of a system do not contain τ at all. Then there are infinitely many values that produce the same output and τ cannot be identified from the available data. This is demonstrated by the following example:

Example 2:

$$\left\{ \begin{array}{lcl} \dot{x}_1(t) & = & x_2^2(t-\tau) \\ \dot{x}_2(t) & = & x_2(t) \\ y(t) & = & x_1(t) \\ x(t) & = & \varphi(t), \quad t \in [-\tau, 0] \quad . \end{array} \right. \tag{6.6}$$

Calculating time-derivatives of the output function as above, we obtain

$$\dot{y}(t) = x_2^2(t-\tau) \tag{6.7}$$

$$\ddot{y}(t) = 2(x_2(t-\tau))^2 \quad . \tag{6.8}$$

An output equation of lowest degree (of derivation) for the above system is $\ddot{y}(t) - 2\dot{y}(t) = 0$, which does not involve delays of the variables. We can see that τ is not identifiable for this example, by observing the following symmetry involving the functions of initial conditions φ and τ. For any choice of τ, setting

$$\left\{ \begin{array}{lcl} \varphi_1(t) & = & c \\ \varphi_2(t) & = & e^{t+\tau} \quad , \quad t \in [-\tau, 0] \quad , \end{array} \right. \tag{6.9}$$

where c is a constant, leads to the solution

$$\left\{ \begin{array}{lcl} x_1(t) & = & \frac{e^{2t}}{2} + c - \frac{1}{2} \\ x_2(t) & = & e^{t+\tau} \quad , \end{array} \right. \tag{6.10}$$

for all $t \geq 0$. Since $y(t) = x_1(t)$, it is clear that τ cannot be identified from the output.

6.1.2 Multiple delays

Example 1:
Consider the system

$$\left\{ \begin{array}{lcl} \dot{x}_1(t) & = & -x_2(t-\tau_1) \\ \dot{x}_2(t) & = & x_1(t-\tau_2) \\ y_1(t) & = & x_1(t) \\ y_2(t) & = & x_2(t-\tau_2) \\ x(t) & = & \varphi(t), \quad t \in [-T, 0] \quad . \end{array} \right. \tag{6.11}$$

We have

$$\dot{y}_1(t) = -x_2(t-\tau_1) \tag{6.12}$$
$$\ddot{y}_1(t) = -x_1(t-\tau_1-\tau_2) \tag{6.13}$$

and the explicit input-output equations

$$\ddot{y}_1(t) = -y_1(t-\tau_1-\tau_2) \tag{6.14}$$
$$y_2(t) = -\dot{y}_1(t-\tau_2+\tau_1) \quad , \tag{6.15}$$

from which we can calculate the values of the two time lags. To illustrate, we proceed exactly as in Example 1 of the previous subsection by solving the system numerically for the values $\tau_1 = 1$, $\tau_2 = \sqrt{2}$, $\varphi_1(t) = e^t$ and $\varphi_2(t) = t+1$. We then plot $\mu_1(\tau_1+\tau_2)|_{t_0} := \ddot{y}_1(t_0) + y_1(t_0-\tau_1-\tau_2)$ and $\mu_2(\tau_2-\tau_1)|_{t_0} := y_2(t_0)+\dot{y}_1(t_0-\tau_2+\tau_1)$ for $t_0 = 4$. As expected, the functions are zero for $\tau_1+\tau_2 = 1+\sqrt{2}$ and $\tau_2-\tau_1 = \sqrt{2}-1$ and locally these are the only roots, see Figure 3.

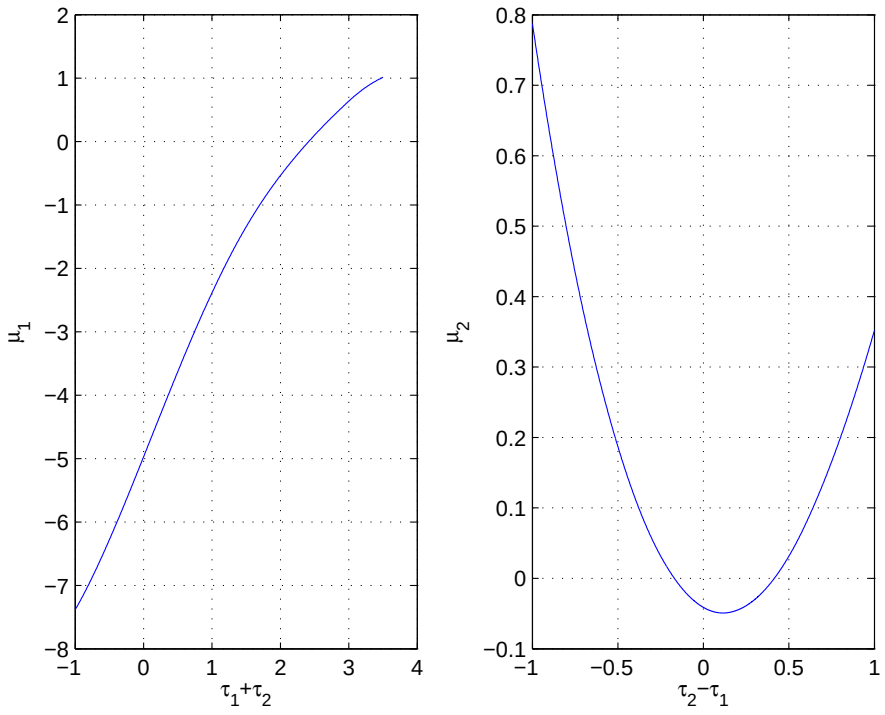


Figure 3: *The functions* $\mu_1(\tau_1+\tau_2)|_4$ *and* $\mu_2(\tau_2-\tau_1)|_4$.

Globally, there are also other roots for $\mu_2(\tau_2 - \tau_1)$ that can be seen in Figure 3. They can be discarded in this case by analysing the zeros of

$$\dot{\mu}_2(\tau_2 - \tau_1)|_{t_0} := \frac{d}{dt}(y_2(t) + \dot{y}_1(t - \tau_2 + \tau_1))|_{t_0}$$

(and other subsequent time-derivatives of μ_2). The correct value of $\tau_2 - \tau_1$ must be a zero to all of these functions. In Figure 4 we have plotted the function

$$\dot{\mu}_2(\tau_2 - \tau_1)|_4 := \frac{d}{dt}(y_2(t) + \dot{y}_1(t - \tau_2 + \tau_1))|_4 \quad .$$

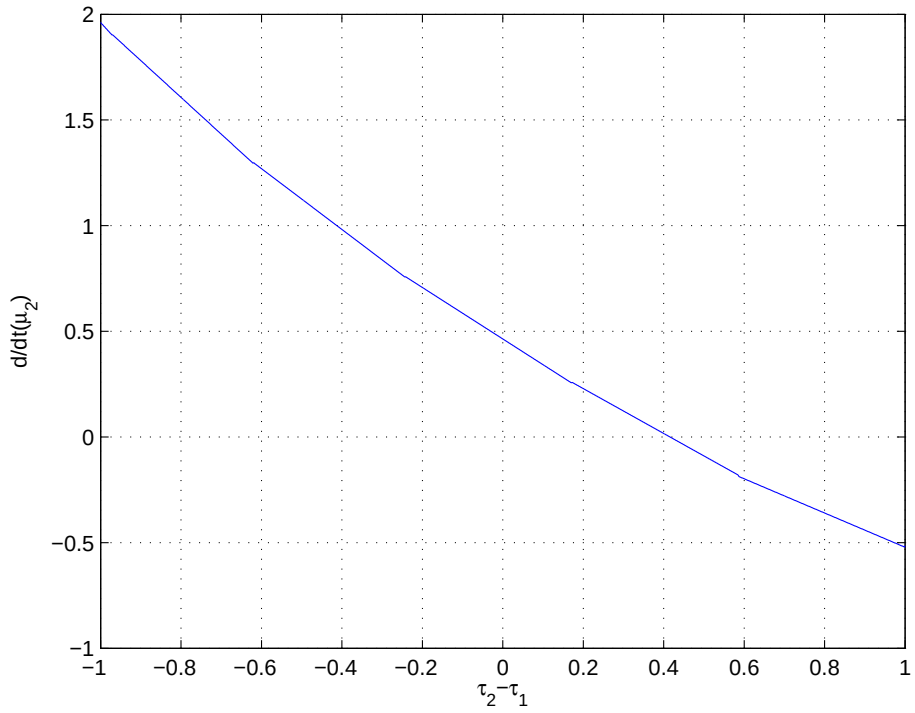


Figure 4: *The function* $\dot{\mu}_2(\tau_2 - \tau_1)$ *for* $t_0 = 4$.

6.2 Linear-algebraic criteria

Clearly, it is only feasible to calculate explicit input-output equations for systems with few variables and parameters. Another method is needed for

analysis of the identifiability of the delay parameters. For the actual calculation of their values, we will rely on other existing methods for parameter estimation, see for example the book by Banks et al. (1989).

In this section, we give a formal definition of the identifiability of the time-delay parameter(s) and present linear-algebraic criteria to test this property without explicit calculation of the input-output representation of the system. The detailed derivations can be found in Paper II for the single delay case and in Paper IV for the multiple delay case.

6.2.1 The single delay case

Let us consider the case of a single time-delay parameter:

$$\begin{cases} \dot{x}(t) &= f\big(x(t), x(t-\tau), u, u(t-\tau)\big) \\ y(t) &= h\big(x(t), x(t-\tau), u, u(t-\tau)\big) \\ x(t) &= \varphi(t), \quad t \in [-\tau, 0] \\ u(t) &= u_0(t), \quad t \in [-T, 0] \quad . \end{cases} \tag{6.16}$$

6.2.1.1 Definition of identifiability

The property of local identifiability of the delay parameter τ is defined as follows:

Definition 6.1 *The delay parameter τ is said to be locally identifiable at $\tau_0 \in (0,T)$ if there exists an open set $W \ni \tau_0$, $W \subset [0,T)$, such that $\forall \tau_1 \in W$: $\tau_1 \neq \tau_0$, $\forall \varphi_0, \varphi_1 \in C$, there exist $t_0 \geq 0$ and $u \in C_U$ such that $y(t_0, \varphi_1, u, \tau_1) \neq y(t_0, \varphi_0, u, \tau_0)$, where $y(t, \varphi, u, \tau)$ denotes the parameterized output for the initial function φ, the admissible input u and delay τ. The delay parameter τ is said to be locally identifiable if it is locally identifiable for all $\tau_0 \in (0,T)$.*

6.2.1.2 Criteria

In this subsection, we show that the local identifiability of τ in (6.16) depends on whether it is present in the input-output representation of the system. The latter can be decided in one of two ways: the first is by the occurrence of a delayed input variable in the time-derivatives of the output functions; the second is by a linear-algebraic criterion involving rank calculation for the set of gradients of the outputs and their time-derivatives over $\mathcal{K}$ and over $\mathcal{K}(\delta]$.

The presence of τ in the input-output representation of the system can be defined formally in terms of δ (using the definitions from the previous section, Subsection 5.1):

Definition 6.2 *An input-output equation* $\phi(\delta, y, \ldots, y^{(l)}, u, \ldots, u^{(k)}) = 0$ *is said to involve* δ *in an essential way if the meromorphic function* $\phi(\ldots)$ *cannot be written as* $c(\delta]\tilde{\phi}(y, \ldots, y^{(l)}, u, \ldots, u^{(k)})$ *with* $c(\delta] \in \mathcal{K}(\delta]$.

We have the following result:

Theorem 6.1 *Given a system of the form* (6.16) *and the set* S *defined in Subsection 5.1, there exists an input-output equation*

$$\phi(\delta, y, \ldots, y^{(l)}, u, \ldots, u^{(k)}) = 0 \quad ,$$

that involves δ *in an essential way if and only if (i) or (ii) below are satisfied.*

i. $\frac{\partial h_i^{(j)}(t)}{\partial u_r^{(k)}(t-s\tau)} \neq 0$ *for some* $1 \leq i \leq p$, $0 \leq j \leq s_i$, $s \geq 1$, $1 \leq r \leq m$ *and* $k \geq 0$, *i.e. a delayed input-variable* $u_r^{(k)}$ *occurs in some of the functions in* $\{S, h_1^{(s_1)}, \ldots, h_p^{(s_p)}\}$;

ii. $\operatorname{rank}_{\mathcal{K}(\delta]} \frac{\partial S}{\partial x} \neq \operatorname{rank}_{\mathcal{K}} \frac{\partial (S, h_1^{(s_1)}, \ldots, h_p^{(s_p)})}{\partial x}$.

The next result connects the essential-way dependence of the input-output representation of the system on the delay parameter to the identifiability of the latter:

Theorem 6.2 *Given a system of the form* (6.16), τ *is locally identifiable if and only if there exists an input-output equation* $\phi(\delta, y, \ldots, y^{(l)}, u, \ldots, u^{(k)}) = 0$, *that involves* δ *in an essential way. If* τ *is not locally identifiable, system* (6.16) *can locally be realised as an ODE-system.*

The proofs of Theorem 6.1 and 6.2 can be found in Paper II. There we also show that if the delay parameter is not identifiable, the input-output equations for (6.16) locally represent a system of ordinary differential equations. Recall the notation from Subsection 5.1:

$$\left\{ \begin{array}{lcl} \tilde{x}_1 & = & h_1 \\ & \vdots & \\ \tilde{x}_{s_1} & = & h_1^{(s_1-1)} \\ \tilde{x}_{s_1+1} & = & h_2 \\ & \vdots & \\ \tilde{x}_{s_1+s_2} & = & h_2^{(s_2-1)} \\ & \vdots & \\ \tilde{x}_{s_1+\cdots+s_p} & = & h_p^{(s_p-1)} \\ \tilde{x}_{s_1+\cdots+s_p+1} & = & g_1 \\ & \vdots & \\ \tilde{x}_n & = & g_{n-K} \end{array} \right. \quad . \tag{6.17}$$

If the delay parameter in system (6.16) is unidentifiable, then the following ODE-system has the same output as the original system (6.16) (modulo the output functions h_i for which $s_i = 0$, hence the $\tilde{p}$ below; they are, however, dependent on the rest and thus also the same):

$$\left\{\begin{array}{lcl} \dot{\tilde{x}}_1 & = & \tilde{x}_2 \\ \dot{\tilde{x}}_2 & = & \tilde{x}_3 \\ & \vdots & \\ \dot{\tilde{x}}_{s_1} & = & \tilde{f}_1(\tilde{x}) \\ \dot{\tilde{x}}_{s_1+1} & = & \tilde{x}_{s_1+2} \\ & \vdots & \\ \dot{\tilde{x}}_{s_1+s_2} & = & \tilde{f}_2(\tilde{x}) \\ & \vdots & \\ \dot{\tilde{x}}_{s_1+\cdots+s_p} & = & \tilde{f}_p(\tilde{x}) \\ \tilde{y}_1 & = & \tilde{x}_1 \\ \tilde{y}_2 & = & \tilde{x}_{s_1+1} \\ & \vdots & \\ \tilde{y}_{\tilde{p}} & = & \tilde{x}_{1+s_1+\cdots+s_{p-1}} \quad . \end{array}\right. \tag{6.18}$$

We will now demonstrate this observation on a simple example. Consider the delay system:

$$\left\{\begin{array}{lcl} \dot{x}_1 & = & x_2(t) + x_2(t-\tau) \\ \dot{x}_2 & = & x_2(t) \\ y(t) & = & x_1(t) \\ \varphi_1(t) & = & e^t \\ \varphi_2(t) & = & t+1 \quad . \end{array}\right. \tag{6.19}$$

We have

$$\dot{y}(t) = x_2(t) + x_2(t-\tau) \tag{6.20}$$

$$\ddot{y}(t) = x_2(t) + x_2(t-\tau) = \dot{y}(t) \tag{6.21}$$

and thus, with notation

$$\left\{\begin{array}{lcl} \tilde{x}_1 & = & h_1 \\ \tilde{x}_2 & = & h_1^{(1)} \quad , \end{array}\right. \tag{6.22}$$

we get the ODE-system

$$\left\{\begin{array}{lcl} \dot{\tilde{x}}_1(t) & = & \tilde{x}_2(t) \\ \dot{\tilde{x}}_2(t) & = & \tilde{x}_2(t) \\ \tilde{y}(t) & = & \tilde{x}_1(t) \\ \tilde{x}_1(t_0) & = & x_1(t_0) \\ \tilde{x}_2(t_0) & = & x_2(t_0) + x_2(t_0-\tau) \quad . \end{array}\right. \tag{6.23}$$

Observe that the initial values for the equivalent ODE-system must be solutions of the original system. Choosing $t_0 = 1$ as a starting point for the ODE-system, we set $\tilde{x}_1(1) = e + 1/2$ and $\tilde{x}_2(1) = e + 1$. Figure 5 presents the outputs $y(t)$ and $\tilde{y}(t)$ for t such that all involved output-derivatives ($\tilde{x}$) are continuous ($t \geq 1$).

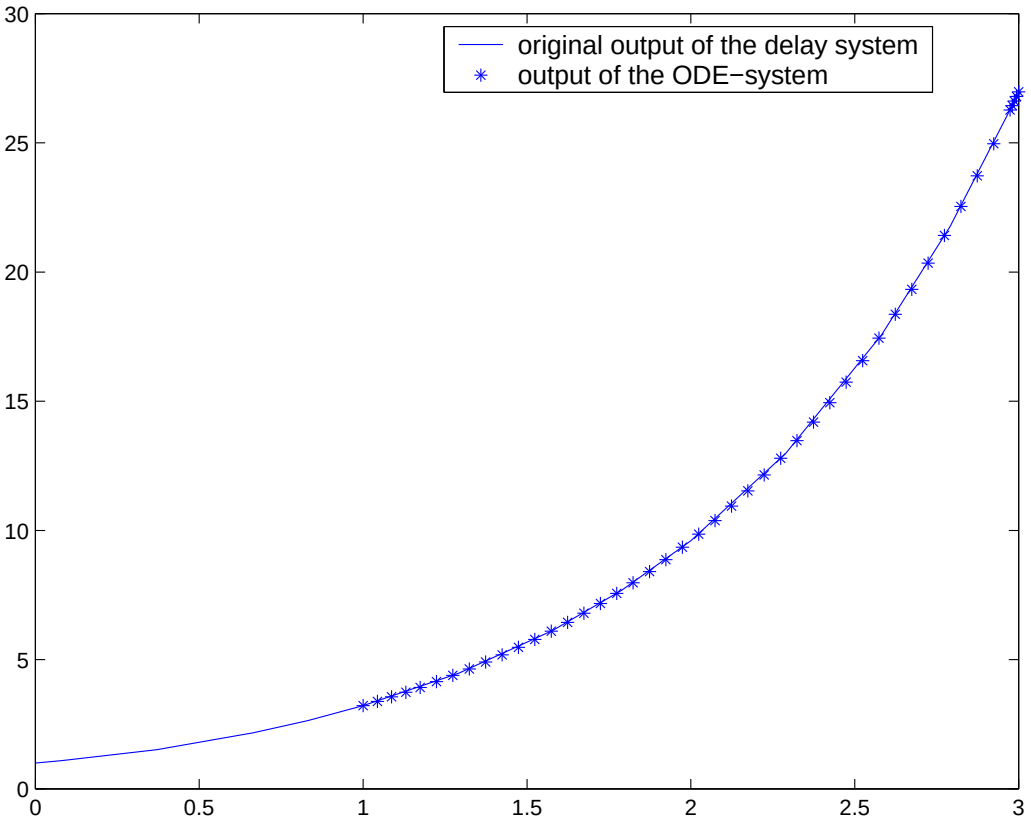


Figure 5: *The outputs $y(t)$ and $\tilde{y}(t)$ for the delay and ODE-systems, respectively.*

6.2.1.3 Examples and relationship between identifiability of the time-lag parameter and observability

We now give simple examples demonstrating the application of the criteria developed in the previous section. For each of the examples we also analyse the observability of the state variables and show that the identifiability of the delay parameter can be a necessary, but not sufficient condition for the observability of the variables (and/or parameter identifiability of the regular model parameters). The already established methods for testing observability

and identifiability for nonlinear delay systems (see Xia et al. (2002) and Zhang et al. (2006)) cannot be used to determine the identifiability of the delay parameter.

Example 1: The full information case.
Consider first a special case of a control system where the output coincides with the undelayed state vector:

$$\begin{cases} \dot{x}(t) &= f\big(x(t), x(t-\tau), u, u(t-\tau)\big) \\ y(t) &= x(t) \quad . \end{cases} \tag{6.24}$$

We can immediately obtain n input-output equations by taking the first time-derivative of each function in the vector y:

$$\dot{y}(t) = f\big(y(t), y(t-\tau), u(t), u(t-\tau)\big) \quad . \tag{6.25}$$

Clearly, unless the given control system is a system of ODEs, there is at least one input-output equation with delays in the variables and the delay parameter τ is identifiable.

Example 2: In this example x_3 and x_4 are in fact parameters which we have incorporated into the general system form used in this work, by setting their time-derivatives to zero and introducing the equalities $\delta x_3 = x_3$ and $\delta x_4 = x_4$. In fact, invariance with respect to shift operators must always be introduced when incorporating parameters in the general form used in this work.

$$\begin{cases} \dot{x}_1(t) &= x_3 x_2(t-\tau) + u(t) \\ \dot{x}_2(t) &= x_4 x_2(t-\tau) \\ \dot{x}_3(t) &= 0 \\ \dot{x}_4(t) &= 0 \\ y(t) &= x_1(t) \quad . \end{cases} \tag{6.26}$$

We have

$$\dot{y} = x_3 \delta x_2 + u \tag{6.27}$$

$$\ddot{y} = x_3 x_4 \delta^2 x_2 + \dot{u} \tag{6.28}$$

$$y^{(3)} = x_3 x_4^2 \delta^3 x_2 + \ddot{u} \tag{6.29}$$

and

$$\frac{\partial (S, h_1^{(s_1)})}{\partial x} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & x_3\delta & \delta x_2 & 0 \\ 0 & x_3 x_4 \delta^2 & x_4 \delta^2 x_2 & x_3 \delta^2 x_2 \\ 0 & x_3 x_4^2 \delta^3 & x_4^2 \delta^3 x_2 & 2 x_3 x_4 \delta^3 x_2 \end{bmatrix} . \tag{6.30}$$

The above matrix has rank 4 over $\mathcal{K}$ and rank 3 over $\mathcal{K}(\delta]$, and τ is locally identifiable by Theorems 6.1 and 6.2. For this simple example, the input-output equation can be obtained explicitly:

$$(\ddot{y} - \dot{u})(\delta\ddot{y} - \delta\dot{u}) - (y^{(3)} - \ddot{u})(\delta\dot{y} - \delta u) = 0 \quad . \tag{6.31}$$

The system as a whole, however, is not weakly observable as the matrix above is not of full rank over $\mathcal{K}(\delta]$. The value of τ can be obtained by numerically finding the zeros of the meromorphic function

$$\big(\ddot{y}(t_0) - \dot{u}(t_0)\big)\big(\ddot{y}(t_0-\tau) - \dot{u}(t_0-\tau)\big) - \big(y^{(3)}(t_0) - \ddot{u}(t_0)\big)\big(\dot{y}(t_0-\tau) - u(t_0-\tau)\big) = 0,$$

where t_0 is a fixed time-point. This delay identification procedure is, of course, sensitive to errors in the computation of the output derivatives.

Example 3: Consider the system

$$\begin{cases} \dot{x}_1(t) &= x_2^2(t-\tau) \\ \dot{x}_2(t) &= x_2(t) \\ y(t) &= x_1(t) \quad . \end{cases} \tag{6.32}$$

We have

$$\dot{y} = (\delta x_2)^2 \tag{6.33}$$

$$\ddot{y} = 2\delta x_2 \delta x_2 = 2(\delta x_2)^2 \tag{6.34}$$

and

$$\frac{\partial(S, h_1^{(s_1)})}{\partial x} = \begin{bmatrix} 1 & 0 \\ 0 & 2\delta(x_2)\delta \\ 0 & 4\delta(x_2)\delta \end{bmatrix} \quad . \tag{6.35}$$

This matrix has rank 2 over both $\mathcal{K}$ and $\mathcal{K}(\delta]$ and τ is not identifiable. This, in turn, means that the state-variable x_2 is not observable. We have the following symmetry involving τ and the function of initial conditions φ_2. For any choice of τ, setting $\varphi_1(t) = c$, $c \in \mathbb{R}$ and $\varphi_2(t) = e^{t+\tau}$, $t \in [-\tau, 0]$, leads to the solution $x_1(t) = \frac{e^{2t}}{2} + c - \frac{1}{2}$, $x_2(t) = e^{t+\tau}$, $\forall t \geq 0$. Since $y(t) = x_1(t)$, it is clear that τ cannot be identified from the output.

6.2.1.4 Simplification of the identifiability criteria

There are some cases when the identifiability of the delay can be decided without any rank calculations. The obvious one is when there is a delayed input variable in the time-derivatives of the (independent) output functions. Then, (*i.*) in Theorem 6.1 is fulfilled and the time-delay is locally identifiable.

Another observation is that if for some output function h_i each derivative $h_i^{(j)}$, $j = 0, \ldots, n$ contains a state-variable that is delayed compared to the previous derivative, then (*ii.*) in Theorem 6.1 is fulfilled. The reason is that for each $j \geq 0$, some $\frac{\partial h_i^{(j)}}{\partial x_k}$, $k = 1, \ldots, n$ is a polynomial in δ of degree higher than for the $j-1$-th derivative and $\text{rank}_{\mathcal{K}} \frac{\partial (S, h_i, \ldots, h_i^{(j)})}{\partial x} = j+1$. Since the latter is true for $j = 0, \ldots, n$ and thus also for $j = s_i$ (because $s_i \leq n-1$), we have

$$\text{rank}_{\mathcal{K}} \frac{\partial (S, h_i, \ldots, h_i^{(s_i)})}{\partial x} = s_i + 1 > s_i = \text{rank}_{\mathcal{K}(\delta)} \frac{\partial S}{\partial x} \quad . \tag{6.36}$$

Thus, τ is identifiable by Theorems 6.1 and 6.2. In practice, it is not even necessary to calculate time-derivatives of the output - one only needs to trace how the variables are delayed in each derivative, which can be done by inspection.

6.2.2 The multiple-delays case

We now consider the case of multiple delays

$$\left\{ \begin{array}{rcl} \dot{x}(t) & = & f\big(x(t), x(t-\tau_1), \ldots, x(t-\tau_\ell), u, u(t-\tau_1), \ldots, u(t-\tau_\ell)\big) \\ y(t) & = & h\big(x(t), x(t-\tau_1), \ldots, x(t-\tau_\ell)\big) \\ x(t) & = & \varphi(t), \quad t \in [-\max_i \tau_i, 0] \\ u(t) & = & u_0(t), \quad t \in [-T, 0] \quad . \end{array} \right. \tag{6.37}$$

6.2.2.1 Definition of identifiability

Intuitively, the τ_i:s are identifiable if any two sets of parameters can be distinguished by the system's input-output behaviour. The property of local identifiability of the delay parameters τ_i, $i = 1, \ldots, \ell$ is formally defined as follows:

Definition 6.3 *The delay parameters $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_\ell)$ are said to be locally identifiable at $\boldsymbol{\tau_0} \in (0, T)^\ell$ if there exists an open set $W \ni \boldsymbol{\tau_0}$, $W \subset [0, T)^\ell$, such that $\forall \boldsymbol{\tau_1} \in W : \boldsymbol{\tau_1} \neq \boldsymbol{\tau_0}$, $\forall \varphi_0, \varphi_1 \in C$, there exist $t \geq 0$ and $u \in C_U$ such that $y(t, \varphi_1, u, \boldsymbol{\tau_1}) \neq y(t, \varphi_0, u, \boldsymbol{\tau_0})$, where $y(t, \varphi, u, \boldsymbol{\tau})$ denotes the parameterized output for the initial function φ, the admissible input u and delays $\boldsymbol{\tau}$.*

6.2.2.2 Criteria

This section deals with the problem of the identifiability of the time-delay parameters τ_i, $i = 1, \ldots, \ell$ in (6.37). Linear-algebraic criteria are

formulated, based on the linear form of the input-output equations (5.11) from the previous section.

The equations (5.11) give

$$a_i(\boldsymbol{\delta}]dy_i^{(s_i)} + \sum_{r=1}^{m}\sum_{j=0}^{\gamma} c_{j,r}(\boldsymbol{\delta}]du_r^{(j)} = \sum_{l=1}^{i}\sum_{j=0}^{s_l-1} a_{i,l,j}(\boldsymbol{\delta}]dy_l^{(j)} \quad . \tag{6.38}$$

With no loss of generality, we assume that the polynomials $a_i(\boldsymbol{\delta}]$ on the left-hand side of the above equations are irreducible. Let $a_i(\boldsymbol{\delta}] = \sum_{\boldsymbol{k}} a_{i_{\boldsymbol{k}}}\boldsymbol{\delta}^{\boldsymbol{k}}$, $c_{j,r}(\boldsymbol{\delta}] = \sum_{\boldsymbol{k}} c_{j,r_{\boldsymbol{k}}}\boldsymbol{\delta}^{\boldsymbol{k}}$ and $a_{i,l,j}(\boldsymbol{\delta}] = \sum_{\boldsymbol{k}} a_{i,l,j_{\boldsymbol{k}}}\boldsymbol{\delta}^{\boldsymbol{k}}$. Denote all the different monomials $\boldsymbol{\delta}^{\boldsymbol{k}}$ appearing above by $\Delta_{i1},\ldots,\Delta_{iq}$ and the (integer) linear combinations of $\tau_1,\ldots,\tau_\ell$ that they represent by $T_{i1},\ldots,T_{iq}$. If some of the terms in $a_i(\boldsymbol{\delta}]$, $c_{j,r}(\boldsymbol{\delta}]$ or $a_{i,l,j}(\boldsymbol{\delta}]$ is a polynomial in $\boldsymbol{\delta}$ of degree zero, that is, the input-output equations contain undelayed variables, then we set Δ_{i0} equal to δ_0, where δ_0 denotes the identity operator, and the corresponding T_{i0} is zero. It is the combinations $T_{i0},\ldots,T_{iq}$ of $\tau_1,\ldots,\tau_\ell$ that determine the local identifiability of the latter.

Let Δ_{i_0} be the monomial $\boldsymbol{\delta}^{\boldsymbol{k}}$ in $a_i(\boldsymbol{\delta}]$ with smallest index $\boldsymbol{k}$ (ordered after $k_1,\ldots,k_\ell$) - it is either equal to Δ_{i0} or is among the $\Delta_{i1},\ldots,\Delta_{iq}$. The input-output equations corresponding to (6.38) can locally be written

$$\begin{aligned} y_i^{(s_i)} = \; & \tilde{f}_i\big(\Delta_{i_0}^{-1}\Delta_{i1},\ldots,\Delta_{i_0}^{-1}\Delta_{iq}, y_1,\ldots,y_1^{(s_1-1)},\ldots,y_i,\ldots,y_i^{(s_i-1)},y_i^{(s_i)}, \\ & u,\ldots,u^{(\gamma)}\big) \end{aligned} \tag{6.39}$$

or

$$\begin{aligned} y_i^{(s_i)}(t) = \tilde{f}_i\big(\; & y_1(t),\ldots,y_i^{(s_i-1)}(t),u(t),\ldots,u^{(\gamma)}(t), \\ & y_1^{(s_1-1)}(t-T_{i1}+T_{i_0}),\ldots,y_i^{(s_i)}(t-T_{i1}+T_{i_0}), \\ & u(t-T_{i1}+T_{i_0}),\ldots,u^{(\gamma)}(t-T_{i1}+T_{i_0}), \\ & \ldots, \\ & y_1^{(s_1-1)}(t-T_{iq}+T_{i_0}),\ldots,y_i^{(s_i)}(t-T_{iq}+T_{i_0}), \\ & u(t-T_{iq}+T_{i_0}),\ldots,u^{(\gamma)}(t-T_{iq}+T_{i_0})\big) \quad . \end{aligned} \tag{6.40}$$

Consider those of the equations (6.40), for which $iq \geq 1$. Evaluated at a fixed

time point $t_0 \geq T$, (6.40) gives an equation for $T_{i1}, \ldots, T_{iq}$:

$$\begin{aligned} y_i^{(s_i)}(t_0) \quad = \xi\big(\quad & y_1(t_0), \ldots, y_i^{(s_i-1)}(t_0), u(t_0), \ldots, u^{(\gamma)}(t_0), \\ & y_1^{(s_1-1)}(t_0 - T_{i1} + T_{i_0}), \ldots, y_i^{(s_i)}(t_0 - T_{i1} + T_{i_0}), \\ & u(t_0 - T_{i1} + T_{i_0}), \ldots, u^{(\gamma)}(t_0 - T_{i1} + T_{i_0}), \\ & \ldots, \\ & y_1^{(s_1-1)}(t_0 - T_{iq} + T_{i_0}), \ldots, y_i^{(s_i)}(t_0 - T_{iq} + T_{i_0}), \\ & u(t_0 - T_{iq} + T_{i_0}), \ldots, u^{(\gamma)}(t_0 - T_{iq} + T_{i_0})\big) \quad . \end{aligned} \tag{6.41}$$

Let the time-point t_0 be chosen large enough to ensure the existence and continuity of all time-derivatives involved. This can be achieved by choosing for example $t_0 \geq \max_i s_i T$). Differentiating (6.41) with respect to time gives new equations for $T_{i1} - T_{i_0}, \ldots, T_{iq} - T_{i_0}$ which are independent, since the one-forms $dy_i^{(j)}$, $j \geq 0$ are linearly independent over $\mathcal{K}$ due to $iq \geq 1$:

$$\begin{aligned} y_i^{(s_i+j)}(t_0) \quad = \tfrac{d^j}{dt^j}\xi\big(\quad & y_1(t), \ldots, y_i^{(s_i-1)}(t), u(t), \ldots, u^{(\gamma+j)}(t), \\ & y_1(t - T_{i1} + T_{i_0}), \ldots, y_i^{(s_i-1)}(t - T_{i1} + T_{i_0}, \\ & u(t - T_{i1} + T_{i_0}), \ldots, u^{(\gamma+j)}(t - T_{i1} + T_{i_0}), \ldots, \\ & y_1(t - T_{iq} + T_{i_0}), \ldots, y_i^{(s_i-1)}(t - T_{iq} + T_{i_0}), \\ & u(t - T_{iq} + T_{i_0}), \ldots, u^{(\gamma+j)}(t - T_{iq} + T_{i_0})\big)_{|t_0}. \end{aligned} \tag{6.42}$$

Unless $y_i^{(s_i+j)}(t)$ is identically zero for some $0 \leq j < iq-1$, the first iq of these equations identify $T_{i1} - T_{i_0}, \ldots, T_{iq} - T_{i_0}$ locally. The following equations can in some cases be used to analyze global identifiability. The question, then, is whether the original time delays $\tau_1, \ldots, \tau_\ell$ are identifiable from the integer linear combinations $T_{i1} - T_{i_0}, \ldots, T_{iq} - T_{i_0}$. Let

$$(T_{11} - T_{1_0}, \ldots, T_{1q} - T_{1_0}, \ldots, T_{p1} - T_{p_0}, \ldots, T_{pq} - T_{p_0})^{tr} = M(\tau_1, \ldots, \tau_\ell)^{tr} \quad , \tag{6.43}$$

where M is a $(1q + \cdots + pq) \times \ell$ integer matrix and $\boldsymbol{v}^{tr}$ denotes the transpose of $\boldsymbol{v}$. We can now formulate the identifiability criteria for $\tau_1, \ldots, \tau_\ell$ in the following proposition:

Proposition 6.1 *If M is defined by* (6.43) *and $y_i^{(s_i+j)}(t)$ is not identically equal to zero for any $0 \leq j < iq-1$, $i = \{1, \ldots, p\}$, then $\tau_1, \ldots, \tau_\ell$ are locally identifiable generically, if and only if* $\operatorname{rank}(M) = \ell$.

The proof of this theorem can be found in Paper IV. Setting $\tilde{\boldsymbol{\tau}} := \{T_{i1} - T_{i_0}, \ldots, T_{iq} - T_{i_0}\}$, $i = 1, \ldots, p$, it is shown that system (6.37) can be rep-

resented locally as a neutral system with time lags $\tilde{\boldsymbol{\tau}}$. This system is obtained from the input-output equations (6.40) using the notation from Subsection 5.1:

$$\left\{\begin{array}{lcl}
\dot{\tilde{x}}_1 & = & \tilde{x}_2 \\
& \vdots & \\
\dot{\tilde{x}}_{s_1-1} & = & \tilde{x}_{s_1} \\
\dot{\tilde{x}}_{s_1} & = & \tilde{f}_1(\dot{\tilde{x}}_{s_1}(t-\tilde{\boldsymbol{\tau}}), \tilde{x}, \tilde{x}(t-\tilde{\boldsymbol{\tau}}), u, \ldots, \\
& & u^{(\gamma)}, u(t-\tilde{\boldsymbol{\tau}}), \ldots, u^{(\gamma)}(t-\tilde{\boldsymbol{\tau}})) \\
\dot{\tilde{x}}_{s_1+1} & = & \tilde{x}_{s_1+2} \\
& \vdots & \\
\dot{\tilde{x}}_{s_1+s_2-1} & = & \tilde{x}_{s_1+s_2} \\
\dot{\tilde{x}}_{s_1+s_2} & = & \tilde{f}_2(\dot{\tilde{x}}_{s_1+s_2}(t-\tilde{\boldsymbol{\tau}}), \tilde{x}, \tilde{x}(t-\tilde{\boldsymbol{\tau}}), u, \\
& & \ldots, u^{(\gamma)}, u(t-\tilde{\boldsymbol{\tau}}), \ldots, u^{(\gamma)}(t-\tilde{\boldsymbol{\tau}})) \\
& \vdots & \\
\dot{\tilde{x}}_{s_1+\cdots+s_p} & = & \tilde{f}_p(\dot{\tilde{x}}_{s_1+\cdots+s_p}(t-\tilde{\boldsymbol{\tau}}), \tilde{x}, \tilde{x}(t-\tilde{\boldsymbol{\tau}}), u, \\
& & \ldots, u^{(\gamma)}, u(t-\tilde{\boldsymbol{\tau}}), \ldots, u^{(\gamma)}(t-\tilde{\boldsymbol{\tau}})) \\
\tilde{y}_1 & = & \tilde{x}_1 \\
\tilde{y}_2 & = & \tilde{x}_{s_1+1} \\
& \vdots & \\
\tilde{y}_{\tilde{p}} & = & \tilde{x}_{1+s_1+\cdots+s_{p-1}} \\
\tilde{x}(t) & = & \tilde{\varphi}(t), \quad t \in [t_0 - \max_j \tilde{\tau}_j, t_0] \quad .
\end{array}\right. \tag{6.44}$$

Any pair $(y(t), u(t))$ which satisfies the original system (6.37) also satisfies the above.
We will exemplify by the following system:

$$\left\{\begin{array}{lcl}
\dot{x}_1(t) & = & x_2^2(t-1) \\
\dot{x}_2(t) & = & x_1(t-\sqrt{2})x_2(t) \\
y(t) & = & x_1(t) \\
x_1(t) & = & \varphi_1(t) = e^t, \quad t \in [-\sqrt{2}, 0] \\
x_2(t) & = & \varphi_2(t) = t+2, \quad t \in [-\sqrt{2}, 0] \quad .
\end{array}\right. \tag{6.45}$$

We have

$$\dot{y} = (\delta_1(x_2))^2 \tag{6.46}$$

$$\ddot{y} = 2\delta_1(x_2)\delta_1\delta_2(x_1)\delta_1(x_2) = 2(\delta_1(x_2))^2\delta_1\delta_2(x_1) \quad . \tag{6.47}$$

Extracting the input-output equation, we get $\ddot{y}(t) = 2(\dot{y}(t))^2 y(t-1-\sqrt{2})$.

Thus the original system corresponds to a system with only one delay:

$$\begin{cases} \dot{\tilde{x}}_1(t) &= \tilde{x}_2(t) \\ \dot{\tilde{x}}_2(t) &= 2(\tilde{x}_2(t))^2 \tilde{x}_1(t-1-\sqrt{(2)}) \\ \tilde{y}(t) &= \tilde{x}_1(t) \\ \tilde{x}(t) &= \tilde{\varphi}(t), \quad t \in [t_0 - 1 - \sqrt{2}, t_0] \quad . \end{cases} \tag{6.48}$$

Plotting the outputs from the two delay systems for t such that all output-derivatives ($\tilde{x}$) involved are continuous ($t \geq 1 + \sqrt{2}$), we see that they are the same:

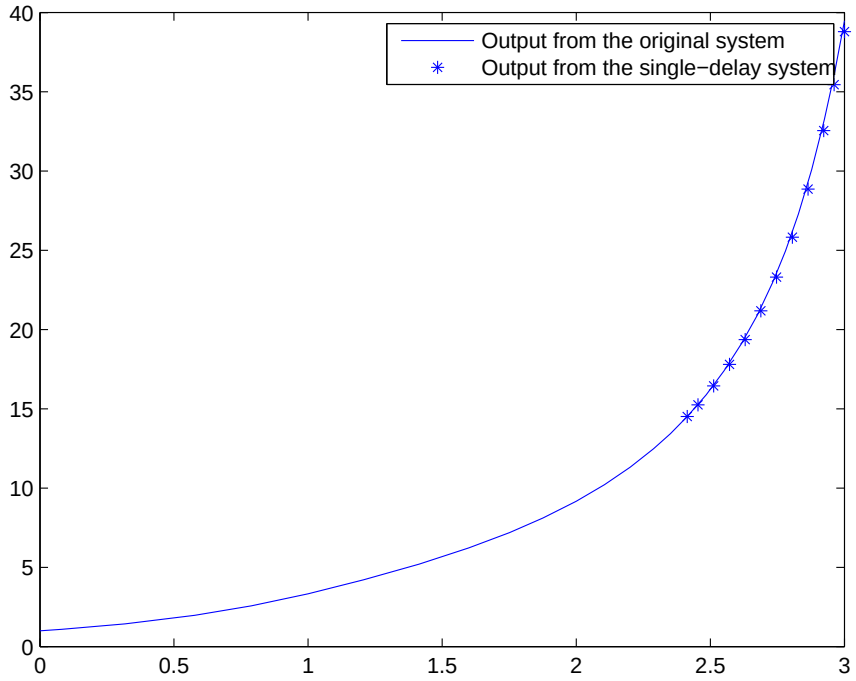


Figure 6: *The outputs $y(t)$ and $\tilde{y}(t)$ for the two-delay and one-delay systems, respectively.*

6.2.2.3 Examples and relationship between identifiability of the time-lag parameters and observability

We now give simple examples demonstrating the application of the criteria developed in the previous section. Similarly to the single time-delay case, the examples show that the identifiability of the delay parameter can be a

necessary, but not sufficient condition for the observability of the variables (and/or parameter identifiability of the regular model parameters).

We start by the example from Subsection 6.1.2:

Example 1:

$$\left\{\begin{array}{lcl} \dot{x}_1(t) &=& -x_2(t-\tau_1) \\ \dot{x}_2(t) &=& x_1(t-\tau_2) \\ y_1(t) &=& x_1(t) \\ y_2(t) &=& x_2(t-\tau_2) \\ x(t) &=& \varphi(t), \quad t \in [-T, 0] \end{array}\right. . \tag{6.49}$$

We have

$$\dot{y_1} = -\delta_1 x_2 \tag{6.50}$$

$$\ddot{y_1} = -\delta_1\delta_2 x_1 \tag{6.51}$$

and

$$\begin{bmatrix} dy_1 \\ d\dot{y}_1 \\ d\ddot{y}_1 \\ dy_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \delta_1 \\ \delta_1\delta_2 & 0 \\ 0 & \delta_2 \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} . \tag{6.52}$$

We obtain the following input-output equations in linear form (6.38):

$$d\ddot{y}_1 = -\delta_1\delta_2 dy_1 \tag{6.53}$$

$$\delta_1 dy_2 = -\delta_2 d\dot{y}_1 \Leftrightarrow dy_2 = -\delta_1^{-1}\delta_2 d\dot{y}_1 . \tag{6.54}$$

We thus have Δ_{10}, $\Delta_{11} = \delta_1\delta_2$, $\Delta_{21} = \delta_1$ and $\Delta_{22} = \delta_2$ and thus, $T_{10} = T_{1_0} = 0$, $T_{11} = \tau_1 + \tau_2$, $T_{2_0} = T_{21} = \tau_1$, $T_{22} = \tau_2$ and

$$M = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ -1 & 1 \end{bmatrix} , \tag{6.55}$$

which is of rank 2. Thus τ_1 and τ_2 are identifiable.

Example 2:

$$\left\{\begin{array}{lcl} \dot{x}_1(t) &=& x_2^2(t-\tau_1) + u \\ \dot{x}_2(t) &=& x_1(t-\tau_2)x_2 \\ y(t) &=& x_1(t) \\ x(t) &=& \varphi(t), \quad t \in [-\tau, 0] \end{array}\right. . \tag{6.56}$$

We have

$$\dot{y} = (\delta_1(x_2))^2 + u \tag{6.57}$$

$$\ddot{y} = 2\delta_1(x_2)\delta_1\delta_2(x_1)\delta_1(x_2) + \dot{u} = 2(\delta_1(x_2))^2\delta_1\delta_2(x_1) + \dot{u} \tag{6.58}$$

and

$$\begin{bmatrix} dy \\ d\dot{y} - du \\ d\ddot{y} - d\dot{u} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2\delta_1(x_2)\delta_1 \\ 2(\delta_1(x_2))^2\delta_1\delta_2 & 4\delta_1\delta_2(x_1)\delta_1(x_2)\delta_1 \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} \quad . \tag{6.59}$$

Clearly, the above matrix has rank 2 over $\mathcal{K}(\boldsymbol{\delta}]$ and so the system is weakly observable according to the definition in Xia et al. (2002), if τ_1 and τ_2 are known.

However, τ_1 and τ_2 are not identifiable. From the last equality we obtain the (input-)output equation (in linear form (6.38)):

$$d\ddot{y} - d\dot{u} = 2(\delta_1(x_2))^2\delta_1\delta_2 dy + 2\delta_1\delta_2(x_1)(d\dot{y} - du) \quad , \tag{6.60}$$

and we see that there are two monomials, Δ_0 and $\Delta_1 = \delta_1\delta_2$ with corresponding combinations $T_0 = 0$ and $T_1 = \tau_1 + \tau_2$ of the two time delays. Thus,

$$M = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \tag{6.61}$$

with rank 1, and the time lags are not identifiable.

This can be confirmed by observing that a change in variables $z(t) = x_2(t - \tau_1)$ leads to the following reformulation of the above system:

$$\begin{cases} \dot{x}_1(t) &= z^2(t) \\ z(t) &= x_1(t - (\tau_1 + \tau_2))z(t) \\ y(t) &= x_1(t) \quad , \end{cases} \tag{6.62}$$

which involves a single delay equal to the sum of τ_1 and τ_2.

6.3 Analysis of time-delay models from systems biology

Delay-differential equations have been used to model a wide range of phenomena in systems biology, including the circadian pacemaker (Lema et al., 2000; Smolen et al., 1999), the *lac* operon (Mahaffy and Savev, 1999), metabolic insulin signaling (Sedaghat et al., 2002), gene expression in cultured mammalian cells (Monk, 2003) and in zebrafish (Lewis, 2003), signal transduction (Timmer et al., 2004) and phosphorylation-dephosphorylation cycles (Srividhya, 2007).

In Paper III, we apply the criteria developed in Section 6.2 to analyse the identifiability of the time-lag parameter in the models by Timmer et al.

(2004) and Monk (2003). The calculations for these systems display the simplifications in the analysis that can be made when testing the identifiability of a single delay parameter in practice, discussed in Paragraph 6.2.1.4.

In this section, we apply the simplified identifiability criteria to one more model from systems biology. We analyse the identifiability of the time-lag parameter in the delay model of genetic regulation for a single transcription factor activating its own transcription by Smolen et al. (1999). The mechanism of macromolecular transport is expected to be important for the process of genetic regulation. In the paper (Smolen et al., 1999), two different methods of modelling the transport of macromolecules between the nucleus and cytoplasm are analysed. The first method assumes diffusive transport of mRNA and protein, the second assumes active transport for which a time delay in the model can be used. The latter method was applied to model a single transcription factor (TF) which when phosphorylated, binds to DNA sequences known as responsive elements (TF-RE) and activates its own transcription. The transcription factor, TF-A, forms a homodimer that can bind to TF-REs present in the *tf-a* gene and increase *tf-a* transcription. A delay in the model appears between the transcription of *tf-a* mRNA and any change in the level of nuclear TF-A and this delay is due to two time-lags. The first is the time it takes for the movement of *tf-a* mRNA from the innermost shell of the nucleus to the outermost shell for translation into TF-A protein. The second is the time required for the movement of TF-A protein to the nucleus. The differential equations are

$$\left\{\begin{array}{lcl} \dfrac{d[\textit{tf-a}\ \text{mRNA}]}{dt} & = & \dfrac{k_{1,f}[\text{TF}-\text{A}]^2}{[\text{TF}-\text{A}]^2+K_d} - k_{1,d}[\textit{tf-a}\ \text{mRNA}] + \text{R}_{\text{bas}} \\ \dfrac{d[\text{TF}-\text{A}]}{dt} & = & k_{2,f}[\textit{tf-a}\ \text{mRNA}](t-\tau) - k_{2,d}[\text{TF}-\text{A}] \\ y & = & [\text{TF}-\text{A}] \quad , \end{array}\right. \tag{6.63}$$

where the two state-variables in $x = ([\text{TF}-\text{A}] \quad [\textit{tf-a}\ \text{mRNA}])$ describe the nuclear concentration of TF-A and the concentration of *tf-a* mRNA, respectively. The 6 parameters p (not counting the delay itself) are the rate constants $k_{i,d}$, $k_{i,f}$, $i = 1, 2$, the dissociation constant of dimer from TF-REs K_d and the constant synthesis rate at negligible dimer concentration, R_{bas}. We thus need to take 7 output derivatives to apply the delay identifiability criteria. Fortunately, we only actually need to observe that in each output derivative, a new delay appears in one of the variables, according to the simplified criteria from Paragraph 6.2.1.4.

$$\begin{array}{rcl} \dot{y} & = & k_{2,f}[\textit{tf-a}\ \text{mRNA}](t-\tau) - k_{2,d}[\text{TF}-\text{A}] = h^{(1)}(\delta([\textit{tf-a}\ \text{mRNA}]), \dots) \\ \ddot{y} & = & h^{(2)}(\delta([\text{TF}-\text{A}]), \delta([\textit{tf-a}\ \text{mRNA}]), \dots) \end{array}$$

$$\begin{aligned} y^{(3)} &= h^{(3)}([\delta^2(tf\text{-}a \text{ mRNA}]), \delta([\text{TF} - \text{A}]), \delta([tf\text{-}a \text{ mRNA}], \dots) \\ y^{(4)} &= h^{(4)}(\delta^2([\text{TF} - \text{A}]), [\delta^2(tf\text{-}a \text{ mRNA}]), \delta([\text{TF} - \text{A}]), \\ &\quad \delta([tf\text{-}a \text{ mRNA}], \dots) \\ &\vdots \quad . \end{aligned} \tag{6.64}$$

For each $j \geq 1$, either $\frac{\partial h^{(j)}}{\partial[\text{TF}-\text{A}]}$ or $\frac{\partial h^{(j)}}{\partial[tf\text{-}a \text{ mRNA}]}$ is a polynomial in δ of degree higher than for $j-1$. This means that $\text{rank}_{\mathcal{K}} \frac{\partial(S, h_1, \dots, h_1^{(8)})}{\partial x} = 9$, which is then greater than $\text{rank}_{\mathcal{K}(\delta)} \frac{\partial S}{\partial x, p}$ as the latter is limited by the number of variables and parameters, eight in this case. Thus, τ is identifiable.

7 Discussion and future work

The second part of this thesis presents both previously published and new results on the observability and identifiability of systems of nonlinear delay-differential equations with discrete time-delays. New criteria have been developed to analyse the identifiability of the time-lag parameters via the input-output characterisation of the system.

The literature on the observability of nonlinear delay systems is scarce. The only definitions, to the best of our knowledge, are algebraic ones and no analysis has yet been made as to how they relate to geometric and more intuitive definitions. A problem for future work is thus to find relations between the different concepts in the manner of the paper on linear delay systems by Lee and Olbrot (1981).

As usual for nonlinear systems, the application of observability and identifiability criteria becomes increasingly difficult with the size of the models. For instance, all of the biological delay models we were able to find in the literature were too complex to be analysed for observability, since the criteria involve the calculation of as many time-derivatives of the output as the total number of variables and parameters (minus one). For nonlinear models, this is practically impossible to do by hand. Symbolic packages exist for nonlinear delay systems which can handle relatively small systems with a single time delay. For ODE systems, even large nonlinear models can be analysed efficiently by existing computer algorithms based on symbolic computation, such as the one by Sedoglavic (2002). The results from this thesis show that just as for ODE systems, the observability problem for delay systems reduces to analysis of an algebraic system of equations, and it should be possible to use similar symbolic computation algorithms even in this case. The same applies to the analysis of the identifiability of the time-lag parameters. The development of such algorithms is an important issue left for future work.

References

Alaoui, H. T., & Yafia, R. (2004). Stability and Hopf bifurcations in a haematopoietic stem cells model. *Electronic Journal of Differential Equations, 2004-Fez Conference on Differential Equations and Mechanics, Conference 11*, 167-173.

Banks, H., & Kunisch, K.. (1989). *Estimation Techniques for Distributed Parameter Systems*. Boston: Birkhauser.

Belkoura, L., Dambrine, M., & Richard, J. P. (2000). Identifiability of delay systems: a distribution approach. *In Proceedings of 14th MTNS 2000, Mathematical Theory of Networks and Systems*. France-Perpignan.

Belkoura, L., & Orlov, Y.. (2002). Identifiability analysis of linear delay-differential systems. *IMA J. Math. Control I., 19*, 73-81.

Chappell, M. J., Evans, N. D., Godfrey, K. R., & Chapman, M. J.. (1999). Structural identifiability of controlled state space systems: a quasi-automated methodology for generating identifiable reparameterisations of unidentifiable systems. *IEE Colloquium on Symbolic Computation for Control (Ref. No. 1999/088)*, 8/1-12.

Cohn, P. M. (1985). *Free rings and their relations*. 2nd edition. London: Academic Press.

Conte, G., & Perdon, A. M. (1984). The disturbance decoupling problem for systems over a ring. *SIAM J. Control Optim., 33*, 750-764.

Conte, G., Moog, C. H., & Perdon, A. M. (1999). *Nonlinear control systems: An algebraic setting, Lecture notes in control and information sciences, 242*. London: Springer-Verlag.

Cox, Little, & O'Shea. (1997).*Ideals, Varieties, and Algorithms*. Springer-Verlag.

Diekmann, O., van Gils, S. A., Verduyn Lunel, S. M., Walther H.-O. (1995). *Delay equations: Functional-, complex-, and nonlinear analysis*. Springer-Verlag.

Fliess, M. (1993). Generalized controller canonical forms for linear and nonlinear dynamics. *IEEE T. Aut. Contr., 35*, 994-1001.

Forsman, K., & Habets, L.. (1994). Input-Output Equations and Observability for Polynomial Delay Systems. *In Proceedings of the 33rd IEEE Conference on Decision and Control, December 1994*, 880 - 882.

Gárate-García, A., Cuesta-Garcia, J. R., & Márquez-Martínez, L. A. (2006). Symbolic analysis and control package for nonlinear time-delay systems. *In Proceedings of Congreso Nacional de Control Automático 2006, October 2006*, 381-386.

Goodearl, K. R., & Warfield, R. B. Jr. (2004). An introduction to noncommutative noetherian rings. 2nd edition. Cambridge University Press.

Hale, J. K. (1977). *Theory of functional differential equations*. Springer-Verlag.

Hermann, R., & Krener, A. J. (1977). Nonlinear controllability and observability. *IEEE T. Aut. Contr., 22*, 728-740.

Horbelt, W., Timmer, J., & Voss, H. U.. (2002). Parameter estimation in nonlinear delayed feedback systems from noisy data. *Phys. Lett. A., 299*, 513-521.

Isidori, A. (1995) *Nonlinear control systems*. Springer-Verlag.

Lam, T. Y. (1999). *Lectures on modules and rings*. Springer-Verlag.

Lee, E. B., & Olbrot, A. (1981). Observability and related structural results for linear hereditary systems. *Int. J. Control, 34(6)*, 1061-1078.

Lema, M. A., Golombek, D. A., & Echave, J.. (2000). Delay model of the circadian pacemaker. *J. Theor. Biol., 204*, 565-573.

Lewis, J. Autoinhibition with transcriptional delay: a simple mechanism for the zebrafish somitogenesis oscillator. (2003). *Curr. Biol., 13*, 1398-1408.

Mahaffy, J. M., & Savev, E. S.. (1999). Stability analysis for a mathematical model of the *lac* operon. *Q. Appl. Math., LVII*, 37-53.

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

Márquez-Martínez, L. A., Moog, C. H., & Velasco-Villa, M. (2000). The structure of nonlinear time delay systems. *Kybernetika, 36*, 53-62.

Márquez-Martínez, L. A., Moog, C. H., & Velasco-Villa, M. (2002). Observability and observers for nonlinear systems with time delays. *Kybernetika, 38(4)*, 445-456.

Mboup, M., Join, C., & Fliess,M. (2007). *In Proceedings of 15th Mediterrean Conference on Control and Automation (MED'2007), June 2007.* Athens, Greece.

Monk, N. A. M.. (2003). Oscillatory expression of Hes1, p53, and NF-κB driven by transcriptional time delays. *Curr. Biol., 13*, 1409-1413.

Moog, C. H., Castro-Linares, R., Velasco-Villa, M.,& Márquez-Martínez, L. A. (2000). The disturbance decoupling problem for time-delay nonlinear systems. *IEEE T. Aut. Contr., 45*, 305-309.

Nakagiri, S., & Yamamoto, M. (1995). Unique identification of coefficient matrices, time delays and initial functions of functional differential equations. *Journal of Mathematical Systems, Estimation and Control, 5*, 1-22.

Ollivier, F. (1990). *Le problème de l'identifiabilité structurelle globale: approche théorique, méthodes effectives et bornes de complexité.* . PhD thesis, École polytechnique.

Ore, O. (1931). Linear equations in non-commutative fields. *Ann. Math., 32*, 463-477.

Orlov, Y., Belkoura, L., Richard, J. P., Dambrine, M. (2002). On identifiability of linear time-delay systems. *IEEE T. Aut. Contr., 47*, 1319-1324.

Pohjanpalo. H. (1978). System identifiability based on the power series expansion of the solution. *Math. Biosci., 41*, 21-33.

Richard, J.-P. (2003). Time-delay systems: an overview of some recent advances and open problems. *Automatica, 39*, 1667-1694.

Saccomani, M. P., Audoly, S., Bellu, G., & D'Angiò, L. (2001). A new differential algebra algorithm to test identifiability of nonlinear systems with given initial conditions. *Proceedings of the 40th Conference on Decision and Control*, 3108-3113. Orlando-Florida USA.

Sedaghat, A. R., Sherman, A., & Quon, M. J. (2002). A mathematical model of metabolic insulin signaling pathways. *Am. J. Endocrinol. Metab., 283*, 1084-1101.

Sedoglavic, A. (2002). A probabilistic algorithm to test local algebraic observability in polynomial time. *J. Symb. Comput., 33*, 735-755.

Shampine, L. F. & Thompson S.. Solving DDEs in MATLAB. *Appl. Numer. Math., 37*, 441-448.

Smolen, P., Baxter, D. A., & Byrne, J. H.. (1999). Effects of macromolecular transport and stochastic fluctuations on dynamics of genetic regulatory systems. *Am. J. Cell Physiol., 277*, 777-790.

Sontag, E. D. (1991). I/O equations for nonlinear systems and observation spaces. *Proceedings of the 30th Conference on Decision and Control*, 720-725. Brighton-England.

Srividhya, J., Gopinathan, M. S., & Schnell, S. (2007). The effects of time delays in a phosphorylation-dephosphorylation pathway. *Biophys. Chem., 125*, 286-297.

Timmer, J., Müller, T. G., Swameye, I., Sandra, O., & Klingmüller, U. (2004). Modelling the nonlinear dynamics of cellular signal transduction. *Int. J. Bifurcat. Chaos, 14*, 2069-2079.

Vajda, S., Godfrey, K., & Rabitz, H.. (1989). Similarity transformation approach to identifiability analysis of nonlinear compartmental models. *Math. Biosci., 93*, 217-248.

Verduyn Lunel, S. M. (2001). Parameter identifiability of differential delay equations. *Int. J. Adapt. Control, 15*, 655-678.

Xia, X., Márquez, L. A., Zagalak, P., & Moog, C. H. (2002). Analysis of nonlinear time-delay systems using modules over noncommutative rings. *Automatica, 38*, 1549-1555.

Zhang, J., Xia, X., & Moog, C. H. (2006). Parameter identifiability of nonlinear systems with time-delay. *IEEE T. Aut. Contr., 47*, 371-375.

Paper II

State elimination and identifiability of the delay parameter for nonlinear time-delay systems

Milena Anguelova [a,b,1] Bernt Wennberg [a,b]

[a] *Mathematical Sciences, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden*

[b] *Mathematical Sciences, Göteborg University, SE-412 96 Gothenburg, Sweden*

Abstract

The identifiability of the delay parameter for nonlinear systems with a single constant time-delay is analysed. We show the existence of input-output equations and relate the identifiability of the delay parameter to their form. Explicit criteria based on rank calculations are formulated. The identifiability of the delay parameter is shown not to be directly related to the well-characterized identifiability/observability of the other system parameters/states.

Key words: identifiability; time-delays; nonlinear time delay control systems; state elimination; input-output description.

1 Introduction

Observability and parameter identifiability are important properties of a system where initial state or parameter estimation are concerned. These properties guarantee that the desired quantities can be uniquely determined from the available data.

For nonlinear systems without time delays, these properties are well-characterised, see for instance Hermann and Krener (1977), Pohjanpalo (1978), Vajda, Godfrey and Rabitz (1989), Sontag (1991), Sedoglavic (2002) and the references therein. The characterization of observability and identifiability has now been extended to nonlinear systems with time delays by Xia, Márquez, Zagalak and Moog (2002) and Zhang, Xia and Moog (2006), using an algebraic approach introduced by Moog, Castro-Linares, Velasco-Villa and Márquez-Martínez (2000), and developed in Márquez-Martínez, Moog and Velasco-Villa (2000). In these works the time delays themselves are assumed to be known, or multiples of a unit delay. The identifiability of general unknown time-delays has been analyzed only for linear systems (Nakagiri and Yamamoto, 1995; Verduyn Lunell, 2001; Orlov, Belkoura, Richard and Dambrine, 2002; Belkoura and Orlov, 2002).

In this paper, we use the mathematical setting of Moog et al. (2000), Márquez-Martínez et al. (2000), Xia et al. (2002) and Zhang et al. (2006) to analyze the identifiability of the time-delay parameter for nonlinear control systems with a single unknown constant time delay. It is shown that state elimination produces input-output relations for the system, the form of which decides the identifiability of the time-delay. The value of the delay parameter can be found directly from the input-output equations, if these can be obtained explicitly. We formulate linear-algebraic criteria to check the identifiability of the delay parameter which eliminate the need for an explicit calculation of the input-output relations.

We show that the identifiability of the delay parameter can be a necessary but not sufficient condition for the observability of the state variables (and identifiability of the regular parameters in the system). The already established methods for testing observability and identifiability for nonlinear delay systems alone (Xia et al., 2002; Zhang et al., 2006) cannot be used to determine the identifiability of the time-delay parameter and a prior analysis is necessary for the latter.

The layout of the paper is as follows: the mathematical framework is presented in Section 2. The result on state elimination and existence of input-output relations

Email addresses: milena@math.chalmers.se (Milena Anguelova), wennberg@math.chalmers.se (Bernt Wennberg).

[1] Corresponding author. Tel. +46-31-7723547. Fax +46-31-161973.

is given in Section 3. In Section 4, explicit criteria for the identifiability of the delay parameter are described. These are then applied to examples in Section 5, followed by the conclusion in Section 6.

2 Notation and preliminary definitions

Consider nonlinear time-delay systems of the form:

$$\begin{cases} \dot{x}(t) = f(x(t), x(t-\tau), u, u(t-\tau)) \\ y(t) = h(x(t), x(t-\tau), u, u(t-\tau)) \\ x(t) = \varphi(t), \quad t \in [-\tau, 0] \\ u(t) = u_0(t), \quad t \in [-T, 0] \end{cases}, \qquad (1)$$

where $x \in \mathbb{R}^n$ denotes the state variables, $u \in \mathbb{R}^m$ is the input and $y \in \mathbb{R}^p$ is the output. Any parameters for the system can be written as state variables with time-derivative zero. The unknown constant time-delay is denoted by $\tau \in [0, T)$, $T \in \mathbb{R}$. The entries of f and h are meromorphic in their arguments (quotients of convergent power series with real coefficients) and $\varphi : [-\tau, 0] \to \mathbb{R}^n$ is an unknown continuous function of initial conditions. The set of initial functions for the variables x is denoted by $C := C([-\tau, 0], \mathbb{R}^n)$. A meromorphic function u is called an admissible input if the differential equation above admits a unique solution. The set of all such input functions is denoted by C_U.

One of the objectives of this work is to investigate the property of local identifiability of the delay parameter τ, which we define as follows:

Definition 1 *The delay parameter τ is said to be locally identifiable at $\tau_0 \in (0, T)$ if there exists an open set $W \ni \tau_0$, $W \subset [0, T)$, such that $\forall \tau_1 \in W : \tau_1 \neq \tau_0$, $\forall \varphi_0, \varphi_1 \in C$, there exist $t_0 \geq 0$ and $u \in C_U$ such that $y(t_0, \varphi_1, u, \tau_1) \neq y(t_0, \varphi_0, u, \tau_0)$, where $y(t, \varphi, u, \tau)$ denotes the parameterized output for the initial function φ, the admissible input u and delay τ. The delay parameter τ is said to be locally identifiable if it is locally identifiable for all $\tau_0 \in (0, T)$.*

Following the notations and mathematical setting of Moog et al. (2000), Márquez-Martínez et al. (2000), Xia et al. (2002) and Zhang et al. (2006), let $\mathcal{K}$ be the field of meromorphic functions of a finite number of variables from $\{x(t-k\tau), u(t-k\tau), \dots, u^{(l)}(t-k\tau), \quad k, l \in \mathbb{Z}^+\}$ and let δ denote the time-shift operator, $\delta(\xi(t)) = \xi(t-\tau)$, $\xi(t) \in \mathcal{K}$. Let $\mathcal{K}(\delta]$ denote the set of polynomials of the form

$$a(\delta] = a_0(t) + a_1(t)\delta + \cdots + a_{r_a}(t)\delta^{r_a} \quad , \qquad (2)$$

where $a_j(t) \in \mathcal{K}$. With addition defined as usual and multiplication given by

$$a(\delta]b(\delta] = \sum_{k=0}^{r_a+r_b} \sum_{i+j=k}^{i \leq r_a, j \leq r_b} a_i(t) b_j(t - i\tau)\delta^k \quad , \qquad (3)$$

$\mathcal{K}(\delta]$ is a noncommutative ring, which is Noetherian and a left Ore domain. The latter implies that the rank of a module over $\mathcal{K}(\delta]$ is well-defined (Cohn, 1971). Let $\mathcal{M}$ denote the module $span_{\mathcal{K}(\delta]}\{d\xi : \xi \in \mathcal{K}\}$. The closure of a submodule $\mathcal{N}$ in $\mathcal{M}$ is the submodule $\overline{\mathcal{N}} = \{w \in \mathcal{M} : \exists a(\delta] \in \mathcal{K}(\delta], \quad a(\delta]w \in \mathcal{N}\}$. It is the largest submodule of $\mathcal{M}$ containing $\mathcal{N}$ with rank equal to $rank_{\mathcal{K}(\delta]}\mathcal{N}$ (Xia et al., 2002).

Differentiation of functions $\phi(x(t - i\tau), u(t - j\tau), \dots, u^{(l)}(t - j\tau))$, $0 \leq i, j \leq k$, $l \geq 0$ in $\mathcal{K}$ and one-forms $\omega = \sum_i \kappa_x^i dx(t-i\tau) + \sum_{ij} \nu_i du^{(j)}(t-i\tau)$ in $\mathcal{M}$ is defined in the natural way:

$$\begin{aligned} \dot{\phi} &= \textstyle\sum_{i=0}^{k} \frac{\partial \phi}{\partial x(t-i\tau)} \delta^i f + \\ &\quad + \textstyle\sum_{r=0}^{l} \sum_{j=0}^{k} \frac{\partial \phi}{\partial u^{(r)}(t-j\tau)} u^{(r+1)}(t - j\tau) \\ \dot{\omega} &= \textstyle\sum_i \dot{\kappa}_x^i dx(t - i\tau) + \sum_{ij} \dot{\nu}_i du^{(j)}(t - i\tau) + \\ &\quad + \textstyle\sum_i \kappa_x^i d\delta^i f + \sum_{ij} \nu_i du^{(j+1)}(t - i\tau). \end{aligned} \qquad (4)$$

Define

$$\begin{aligned} \mathcal{X} &= span_{\mathcal{K}(\delta]}\{dx\} \\ \mathcal{Y}_k &= span_{\mathcal{K}(\delta]}\{dy, d\dot{y}, \dots, dy^{(k-1)}\} \\ \mathcal{U} &= span_{\mathcal{K}(\delta]}\{du, d\dot{u}, \dots\}. \end{aligned} \qquad (5)$$

Then $(\mathcal{Y}_k + \mathcal{U}) \cap \mathcal{X} = (\mathcal{Y}_n + \mathcal{U}) \cap \mathcal{X}$ for $k \geq n$ and $rank_{\mathcal{K}(\delta]}(\mathcal{Y}_n + \mathcal{U}) \cap \mathcal{X} \leq n$ (Xia et al., 2002).

3 State elimination

In this section we consider the problem of obtaining an input-output representation from the state-space form of a time-delay control system. This problem has been treated for polynomial systems in Forsman et al. (1994). It will be shown that for a system of the form (1) there always exists, at least locally, a set of input-output delay-differential equations of the form $F(\delta, y, u) = 0$, where:

$$\begin{aligned} F(\delta, y, u) &:= F(\delta, y, \dots, y^{(k)}, u, \dots, u^{(\gamma)}) = \\ &= F(y, \dots, y^{(k)}, \dots, \delta^l(y), \dots, \delta^l(y^{(k)}), \\ &\quad u, \dots u^{(\gamma)}, \dots, \delta^{(\beta)}(u), \dots, \delta^{(\beta)}(u^{(\gamma)})) \end{aligned} \qquad (6)$$

such that any pair $(y(t), u(t))$ which solves (1), also satisfies (6), for t such that all derivatives involved are continuous. The function F is meromorphic in its arguments.

Theorem 1 *Given a system of the form* (1)*, there exist integers* γ*,* β*,* k *and* l *and an open dense subset* V *of* $C \times C_U^{\gamma+1}$*, such that in the neighborhood of any point of* V*, there exists an input-output representation of the system of the form* (6)*.*

Proof: The proof is an adaptation of the proof of Theorem 2.2.1. in Conte, Moog and Perdon (1999) for the analogous result for ODE-systems.

Let f be an r-dimensional vector with entries $f_j \in \mathcal{K}$. Let $\frac{\partial f}{\partial x}$ denote the $r \times n$ matrix with entries

$$\left(\frac{\partial f}{\partial x}\right)_{j,i} = \sum_{\ell} \frac{\partial f_j}{\partial x_i(t-\ell\tau)}\delta^{\ell} \in \mathcal{K}(\delta]\,. \tag{7}$$

Denote by s_1 the least nonnegative integer such that

$$rank_{\mathcal{K}(\delta]} \frac{\partial(h_1, ..., h_1^{(s_1-1)})}{\partial x} = rank_{\mathcal{K}(\delta]} \frac{\partial(h_1, ..., h_1^{(s_1)})}{\partial x}. \tag{8}$$

If $\frac{\partial h_1}{\partial x} \equiv 0$ then we set $s_1 = 0$. Inductively, for $1 < i \leq p$ denote by s_i the least nonnegative integer such that

$$rank_{\mathcal{K}(\delta]} \frac{\partial(h_1, ..., h_1^{(s_1-1)}, \ldots, h_i, \ldots, h_i^{(s_i-1)})}{\partial x} = \\ = rank_{\mathcal{K}(\delta]} \frac{\partial(h_1, ..., h_1^{(s_1-1)}, \ldots, h_i, \ldots, h_i^{(s_i)})}{\partial x}. \tag{9}$$

Let $S = (h_1, ..., h_1^{(s_1-1)}, \ldots, h_p, \ldots, h_p^{(s_p-1)})$, where h_i does not appear if $s_i = 0$. Then

$$rank_{\mathcal{K}(\delta]} \frac{\partial S}{\partial x} = s_1 + \cdots + s_p = K \leq n \quad . \tag{10}$$

If $K < n$, there exist meromorphic functions $g_1(\delta, x)$, $\ldots$, $g_{n-K}(\delta, x)$ such that $rank_{\mathcal{K}(\delta]} \frac{\partial(S, g_1, \ldots, g_{n-K})}{\partial x} = n$.

For simplicity, introduce the notation

$$\begin{cases} \tilde{x}_1 & = h_1 \\ \ldots & \\ \tilde{x}_{s_1} & = h_1^{(s_1-1)} \\ \tilde{x}_{s_1+1} & = h_2 \\ \ldots & \\ \tilde{x}_{s_1+s_2} & = h_2^{(s_2-1)} \\ \ldots & \\ \tilde{x}_{s_1+\cdots+s_p} & = h_p^{(s_p-1)} \\ \tilde{x}_{s_1+\cdots+s_p+1} & = g_1 \\ \ldots & \\ \tilde{x}_n & = g_{n-K} \end{cases} . \tag{11}$$

The one-forms $d\tilde{x}_i$, $i = 1, \ldots, n$ now form a basis of $\mathcal{X}$. The definition of s_i, (9) implies that for each i

$$\frac{\partial h_i^{(s_i)}}{\partial x} \in \overline{span_{\mathcal{K}(\delta]}\Big\{\frac{\partial(h_1, \ldots, h_1^{(s_1-1)}, \ldots, h_i^{(s_i-1)})}{\partial x}\Big\}}. \tag{12}$$

Thus there exist nonzero polynomials $b_i(\delta] \in \mathcal{K}(\delta]$, $i = 1, \ldots, p$ such that

$$b_i(\delta]\frac{\partial h_i^{(s_i)}}{\partial x} \in span_{\mathcal{K}(\delta]}\Big\{\frac{\partial(h_1, \ldots, h_1^{(s_1-1)}, \ldots, h_i^{(s_i-1)})}{\partial x}\Big\}. \tag{13}$$

Therefore $b_i(\delta]dh_i^{(s_i)} + \sum_{r=1}^{m}\sum_{j=0}^{\gamma} c_{j,r}(\delta]du_r^{(j)} \in span_{\mathcal{K}(\delta]}\{d\tilde{x}_1, \ldots, d\tilde{x}_{s_1+\cdots+s_i}\}$ for some $\gamma \geq 0$, where γ is the highest derivative of u appearing in the functions in S, and $c_{j,r}(\delta] \in \mathcal{K}(\delta]$. Hence,

$$b_i(\delta]dh_i^{(s_i)} + \sum_{r=1}^{m}\sum_{j=0}^{\gamma} c_{j,r}(\delta]du_r^{(j)} - \sum_{j=1}^{s_1+\cdots+s_i} a_j(\delta]d\tilde{x}_j = 0 \quad , \tag{14}$$

for some $a_j(\delta] \in \mathcal{K}(\delta]$. Since all functions are assumed meromorphic and we have continuous dependence for the output on the input and initial function, the above equality holds on an open dense set of $C \times C_U^{\gamma+1}$. The left hand side of equation (14), being equal to zero, is a closed one-form on $\mathcal{M}$, and therefore, applying the Poincaré lemma we obtain functions $\xi_i(t) \in \mathcal{K}$ such that $d\xi_i = b_i(\delta]dh_i^{(s_i)} + \sum_{r=1}^{m}\sum_{j=0}^{\gamma} c_{j,r}(\delta]du_r^{(j)} - \sum_{j=1}^{s_1+\cdots+s_i} a_j(\delta]d\tilde{x}_j$ and $\xi_i(\delta, h_i^{(s_i)}, \tilde{x}, u, \ldots, u^{(\gamma)}) = 0$, for each $i = 1, \ldots, p$. The function ξ_i does not depend on $\tilde{x}_j$ for $j > s_1 + \cdots + s_i$, since $d\xi_i = 0$ would then contain terms $d\tilde{x}_j$, $j > s_1 + \cdots + s_i$ which is impossible by (14) due to the variables $d\tilde{x}$ being linearly independent over $\mathcal{K}(\delta]$ by definition. Thus, we have obtained a relation

$$\xi_i(\delta, h_i^{(s_i)}, \tilde{x}_1, \ldots, \tilde{x}_{s_1+\cdots+s_i}, u, \ldots, u^{(\gamma)}) = 0, \tag{15}$$

which together with (11) produces an input-output equation

$$\xi_i(\delta, y_i^{(s_i)}, y_1, \ldots, y_1^{(s_1-1)}, \ldots, y_i^{(s_i-1)}, u, \ldots, u^{(\gamma)}) = 0. \tag{16}$$

This is true for each $i, 1 \leq i \leq p$ resulting in p input-output equations where we can set $k = \max_i s_i \quad i = 1, \ldots, p$, and let l and β correspond to the largest delays $-l\tau$ and $-\beta\tau$ of the output and the input variables, respectively. ■

4 Identifiability of the delay parameter

In this section, we show that the local identifiability of τ in (1) depends on whether it is present in the input-output representation of the system. The latter can be

decided either by the occurrence of a delayed input variable in the time-derivatives of the output functions or by a linear-algebraic criterion involving rank calculation for the set of gradients of the outputs and their time-derivatives over $\mathcal{K}$ and over $\mathcal{K}(\delta]$.

The presence of τ in the input-output representation of the system can be defined formally in terms of δ:

Definition 2 *An input-output equation* $\phi(\delta, y, \dots, y^{(l)}, u, \dots, u^{(k)}) = 0$ *is said to involve* δ *in an essential way if the meromorphic function* $\phi(...)$ *cannot be written as* $c(\delta]\tilde{\phi}(y, \dots, y^{(l)}, u, \dots, u^{(k)})$ *with* $c(\delta] \in \mathcal{K}(\delta]$.

We have the following result:

Theorem 2 *Given a system of the form* (1) *and the set* S *defined in the previous section, there exists an input-output equation* $\phi(\delta, y, \dots, y^{(l)}, u, \dots, u^{(k)}) = 0$, *that involves* δ *in an essential way if and only if (i) or (ii) below are satisfied.*

i. $\frac{\partial h_i^{(j)}(t)}{\partial u_r^{(k)}(t-s\tau)} \neq 0$ *for some* $1 \leq i \leq p$, $0 \leq j \leq s_i$, $s \geq 1$, $1 \leq r \leq m$ *and* $k \geq 0$, *i.e. a delayed input-variable* $u_r^{(k)}$ *occurs in some of the functions in* $\{S, h_1^{(s_1)}, \dots, h_p^{(s_p)}\}$;

ii. $rank_{\mathcal{K}(\delta]} \frac{\partial S}{\partial x} \neq rank_{\mathcal{K}} \frac{\partial (S, h_1^{(s_1)}, \dots, h_p^{(s_p)})}{\partial x}$;

Proof: From the definition of $\mathcal{K}$ and $\mathcal{K}(\delta]$, we have

$$rank_{\mathcal{K}(\delta]} \frac{\partial S}{\partial x} \leq rank_{\mathcal{K}} \frac{\partial S}{\partial x} \leq rank_{\mathcal{K}} \frac{\partial (S, h_1^{(s_1)}, \dots, h_p^{(s_p)})}{\partial x}. \tag{17}$$

Assume first that *i.* is true, i.e. $\frac{\partial h_i^{(j)}(t)}{\partial u_r^{(k)}(t-s\tau)} \neq 0$ for some $1 \leq i \leq p$, $0 \leq j \leq s_i$, $s \geq 1$, $1 \leq r \leq m$ and $k \geq 0$. By differentiation one finds that $\frac{\partial h_i^{(s_i)}(t)}{\partial u_r^{(k+s_i-j)}(t-s\tau)} \neq 0$. Without loss of generality we can assume that $u_r^{(k+s_i-j)}$ is the highest derivative of u_r appearing in $h_i^{(s_i)}$. Then in the equality (14)

$$b_i(\delta] dh_i^{(s_i)} + \sum_{r=1}^{m} \sum_{j=0}^{\gamma} c_{j,r}(\delta] du_r^{(j)} - \sum_{j=1}^{s_1+\dots+s_i} a_j(\delta] d\tilde{x}_j = 0 \quad , \tag{18}$$

$c_{k+s_i-j,r}(\delta]$ is a polynomial of degree greater or equal to $s \geq 1$ in δ. Thus, in the corresponding function $\xi_i(\delta, h_i^{(s_i)}, \tilde{x}, u, \dots, u^{(\gamma)})$, such that $d\xi_i = b_i(\delta] dh_i^{(s_i)} + \sum_{r=1}^{m} \sum_{j=0}^{\gamma} c_{j,r}(\delta] du_r^{(j)} - \sum_{j=1}^{s_1+\dots+s_i} a_j(\delta] d\tilde{x}_j$, the variable $u_r^{(k+s_i-j)}$ appears somewhere with a delay $s\tau$ compared to $h_i^{(s_i)}$ and thus δ cannot be omitted from $\xi_i(\delta, y_i^{(s_i)}, y_1, \dots, y_1^{(s_1-1)}, \dots, y_i^{(s_i-1)}, u, \dots, u^{(\gamma)}) = 0$.

Assume now that *ii.* is true, that is, $rank_{\mathcal{K}(\delta]} \frac{\partial S}{\partial x} < rank_{\mathcal{K}} \frac{\partial (S, h_1^{(s_1)}, \dots, h_p^{(s_p)})}{\partial x}$. If all input-output equations $\phi(\delta, y, \dots, y^{(l)}, u, \dots, u^{(k)}) = 0$ are equivalent to an equation $\tilde{\phi}(y, \dots, y^{(l)}, u, \dots, u^{(k)}) = 0$ without δ, then the p input-output equations in the proof of Theorem 1 are equivalent to equations not containing any delays of the variables:

$$\xi_i(y_i^{(s_i)}, y_1, \dots, y_1^{(s_1-1)}, \dots, y_i^{(s_i-1)}, u, \dot{u}, \dots, u^{(k)}) = 0, \tag{19}$$

for $i = 1, \dots, p$. Therefore, $rank_{\mathcal{K}} \frac{\partial (S, h_1^{(s_1)}, \dots, h_p^{(s_p)})}{\partial x} \leq s_1 + \dots + s_p = rank_{\mathcal{K}(\delta]} \frac{\partial S}{\partial x}$, which contradicts the assumption. This completes the first part of the proof.

Suppose now that there exists at least one meromorphic function ϕ such that $\phi(\delta, y, \dots, y^{(l)}, u, \dots, u^{(k)}) = 0$ for some nonnegative integers k and l, where δ cannot be omitted. We will show that *i.* or *ii.* must be true. Assume the opposite. Then the equality $rank_{\mathcal{K}(\delta]} \frac{\partial S}{\partial x} = rank_{\mathcal{K}} \frac{\partial (S, h_1^{(s_1)}, \dots, h_p^{(s_p)})}{\partial x}$ and (17) imply that $rank_{\mathcal{K}} \frac{\partial S}{\partial x} = rank_{\mathcal{K}} \frac{\partial (S, h_1^{(s_1)}, \dots, h_p^{(s_p)})}{\partial x}$. Hence, for each $i = 1, \dots, p$

$$\frac{\partial h_i^{(s_i)}}{\partial x} \in span_{\mathcal{K}} \Big\{ \frac{\partial (h_1, \dots, h_1^{(s_1-1)}, \dots, h_i^{(s_i-1)})}{\partial x} \Big\} \quad . \tag{20}$$

By assumption, the functions in $S, h_1^{(s_1)}, \dots, h_p^{(s_p)}$ do not contain any delayed input-variables, and hence, for each $i = 1, \dots, p$ we have a relation

$$dh_i^{(s_i)} + \sum_{r=1}^{m} \sum_{j=0}^{\gamma} c_{j,r} du_r^{(j)} - \sum_{j=1}^{s_1+\dots+s_i} a_j d\tilde{x}_j = 0 \quad , \tag{21}$$

where $a_j, c_{j,r} \in \mathcal{K}$. Thus the input-output equations (15) for such a system are of the form

$$\xi_i(h_i^{(s_i)}, h_1, \dots, h_1^{(s_1-1)}, \dots, h_i^{(s_i-1)}, u, \dot{u}, \dots, u^{(\gamma)}) = 0, \tag{22}$$

without any delays in the variables since a delay in a variable would mean that some of the $a_j, c_{j,r}$ were polynomials of degree at least one in δ which would be a contradiction. By the Implicit Function Theorem, there exist meromorphic functions f_i such that

$$h_i^{(s_i)} = \tilde{f}_i(h_1, \dots, h_1^{(s_1-1)}, \dots, h_i^{(s_i-1)}, u, \dots, u^{(\gamma)}) = 0, \tag{23}$$

for $i = 1, \dots, p$. This implies that for each $k_i \geq s_i$, $y_i^{(k_i)}$ can be written locally as a function of $y_1, \dots, y_1^{(s_1-1)}$, $\dots, y_i^{(s_i-1)}, u, \dot{u}, \dots, u^{(\gamma)}$ without any delays in the variables. Applying this to the equality $\phi(\delta, y, \dots, y^{(l)}, u, \dots, u^{(k)}) = 0$ we obtain

$$\tilde{\phi}(\delta, y_1, \dots, y_1^{(s_1-1)}, \dots, y_p^{(s_p-1)}, u, \dots, u^{(\gamma)}) = 0 \tag{24}$$

for some meromorphic $\tilde{\phi}$. This, however, contradicts the linear independence of $dh_1, \ldots, dh_1^{(s_1-1)}, \ldots, dh_p, \ldots, dh_p^{(s_p-1)}$ over $\mathcal{K}(\delta]$ and completes the second part of the proof. ■

Theorem 3 *Given a system of the form* (1), *τ is locally identifiable if and only if there exists an input-output equation $\phi(\delta, y, \ldots, y^{(l)}, u, \ldots, u^{(k)}) = 0$, that involves δ in an essential way. If τ is not locally identifiable, system* (1) *can locally be realised as an ODE-system.*

Proof: Suppose first that there exists an input-output equation $\phi(\delta, y, \ldots, y^{(l)}, u, \ldots, u^{(k)}) = 0$, that involves δ in an essential way. Since y and u are known (meromorphic) functions of time, for a fixed t, ϕ is a function of τ with countably many zeros. Thus, τ is locally identifiable.

Assume there is no input-output equation $\phi(\delta, y, \ldots, y^{(l)}, u, \ldots, u^{(k)}) = 0$, that involves δ in an essential way. Then the input-output equations $\xi_i(h_i^{(s_i)}, h_1, \ldots, h_1^{(s_1-1)}, \ldots, h_i^{(s_i-1)}, u, \ldots, u^{(\gamma)}) = 0$, $i = 1, \ldots, p$ do not contain delayed variables. Thus, locally, each $h_i^{(s_i)}$ can be written as a meromorphic function $\tilde{f}_i$ of $h_1, \ldots, h_1^{(s_1-1)}, \ldots, h_i^{(s_i-1)}, u, \ldots, u^{(\gamma)}$, for generic choices of initial conditions $\varphi \in C$. With the notation used in Section 3, we obtain the following ODE-system for the variables $\tilde{x}_i$, $i = 1, \ldots, s_1 + \cdots + s_p$ with the same output as the original system (1) (modulo the output functions h_i for which $s_i = 0$, hence the $\tilde{p}$):

$$\begin{cases} \dot{\tilde{x}}_1 & = \tilde{x}_2 \\ \ldots & \\ \dot{\tilde{x}}_{s_1} & = \tilde{f}_1(\tilde{x}) \\ \dot{\tilde{x}}_{s_1+1} & = \tilde{x}_{s_1+2} \\ \ldots & \\ \dot{\tilde{x}}_{s_1+s_2} & = \tilde{f}_2(\tilde{x}) \\ \ldots & \\ \dot{\tilde{x}}_{s_1+\cdots+s_p} & = \tilde{f}_p(\tilde{x}) \\ \tilde{y}_1 & = \tilde{x}_1 \\ \tilde{y}_2 & = \tilde{x}_{s_1+1} \\ \ldots & \\ \tilde{y}_{\tilde{p}} & = \tilde{x}_{1+s_1+\cdots+s_{p-1}} \end{cases} . \tag{25}$$

We will show that τ in (1) is not locally identifiable by observing the following symmetry involving the delay parameter and the function of initial conditions. Fix the delay parameter to $\tau_0 \in (0, T)$ and the initial conditions to $\varphi_0 \in C$. Let $\tau_1 \in (0, T)$, $\tau_1 \neq \tau_0$ and choose a different continuous function of initial conditions $\varphi_1(t)$ such that the values of $h_k^{(l)}(0)$ are the same (this can be done since $h_k^{(l)}(t)$, $k = 1, \ldots, p$, $l = 0, \ldots, s_i - 1$ do not involve delayed input-variables; keeping $h_k^{(l)}(0)$ invariant amounts to choosing a continuous function $\varphi_1(t)$ which fulfills certain equations for $\varphi_1(0)$ and $\varphi_1(-\tau_1)$). The ODE-system above then has the same initial conditions $\tilde{x}_i(0), i = 1, \ldots, s_1 + \cdots + s_p$ for both sets of τ and φ and thus the same solution which coincides with the output for system (1) (modulo the output functions dependent on the rest, which will also be the same). Since τ_0 is arbitrary, τ is not locally identifiable. ■

Remark: For a system of the form (1), if for some output function h_i each derivative $h_i^{(j)}$, $j = 0, \ldots, n$ contains a state-variable that is delayed compared to the previous derivative, then *ii.* is fulfilled and the delay parameter is identifiable. In such cases the identifiability of the delay can be decided without any rank calculations.

5 Examples

In this section we give simple examples demonstrating the practical calculations and also showing that the identifiability of the delay parameter can be a necessary, but not sufficient condition for the observability of the variables (and/or parameter identifiability of the regular model parameters).

Example 1: The full information case.
Consider a special case of a control system where the output coincides with the undelayed state vector:

$$\begin{cases} \dot{x}(t) = f(x(t), x(t-\tau), u, u(t-\tau)) \\ y(t) = x(t) \end{cases} . \tag{26}$$

We can immediately obtain n input-output equations by taking the first time-derivative of each function in the vector y, $\dot{y} = f(y, y(t-\tau), u, u(t-\tau))$. Clearly, unless the given control system is a system of ODEs, there is at least one input-output equation with delays in the variables and the delay parameter τ is identifiable.

Example 2: In this example x_3 and x_4 are in fact parameters which we have incorporated into the general system form used in this work:

$$\begin{cases} \dot{x}_1 = x_3 x_2(t-\tau) + u \\ \dot{x}_2 = x_4 x_2(t-\tau) \\ \dot{x}_3 = 0 \\ \dot{x}_4 = 0 \\ y = x_1 \end{cases} . \tag{27}$$

We have

$$\begin{aligned} \dot{y} &= x_3\delta x_2 + u \\ \ddot{y} &= x_3x_4\delta^2 x_2 + \dot{u} \\ y^{(3)} &= x_3x_4^2\delta^3 x_2 + \ddot{u} \end{aligned} \tag{28}$$

and

$$\frac{\partial(S,h_1^{(s_1)})}{\partial x} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & x_3\delta & \delta x_2 & 0 \\ 0 & x_3x_4\delta^2 & x_4\delta^2 x_2 & x_3\delta^2 x_2 \\ 0 & x_3x_4^2\delta^3 & x_4^2\delta^3 x_2 & 2x_3x_4\delta^3 x_2 \end{bmatrix}. \tag{29}$$

The above matrix has rank 4 over $\mathcal{K}$ and rank 3 over $\mathcal{K}(\delta]$, and τ is locally identifiable by Theorems 2 and 3. For this simple example, the input-output equation can be obtained explicitly, $(\ddot{y}-\dot{u})(\delta\ddot{y}-\delta\dot{u})-(y^{(3)}-\ddot{u})(\delta\dot{y}-\delta u)=0$. The system as a whole, however, is not weakly observable as the matrix above is not of full-rank over $\mathcal{K}(\delta]$. The value of τ can be obtained by numerically finding the zeros of the meromorphic function $(\ddot{y}(t_0)-\dot{u}(t_0))(\ddot{y}(t_0-\tau)-\dot{u}(t_0-\tau))-(y^{(3)}(t_0)-\ddot{u}(t_0))(\dot{y}(t_0-\tau)-u(t_0-\tau))=0$, where t_0 is a fixed time-point. This delay identification procedure is, of course, sensitive to errors in the computation of the output derivatives.

Example 3: Consider the system

$$\begin{cases} \dot{x}_1 = x_2^2(t-\tau) \\ \dot{x}_2 = x_2 \\ y = x_1 \end{cases} \tag{30}$$

We have

$$\begin{aligned} \dot{y} &= (\delta x_2)^2 \\ \ddot{y} &= 2\delta x_2\delta x_2 = 2(\delta x_2)^2 \end{aligned} \tag{31}$$

and

$$\frac{\partial(S,h_1^{(s_1)})}{\partial x} = \begin{bmatrix} 1 & 0 \\ 0 & 2\delta(x_2)\delta \\ 0 & 4\delta(x_2)\delta \end{bmatrix}. \tag{32}$$

This matrix has rank 2 over both $\mathcal{K}$ and $\mathcal{K}(\delta]$ and τ is not identifiable. This, in turn, means that the state-variable x_2 is not observable. We have the following symmetry involving τ and the function of initial conditions φ_2. For any choice of τ, setting $\varphi_1(t)=c$, $c\in\mathbb{R}$ and $\varphi_2(t)=e^{t+\tau}$, $t\in[-\tau,0]$, leads to the solution $x_1(t)=\frac{e^{2t}}{2}+c-\frac{1}{2}$, $x_2(t)=e^{t+\tau}$, $\forall t\geq 0$. Since $y(t)=x_1(t)$, it is clear that τ cannot be identified from the output.

6 Conclusions

We have analyzed the identifiability of the time-lag parameter in nonlinear delay systems. State elimination yields equations in the inputs and outputs and their derivatives, the form of which decides the identifiability of the delay parameter. For simpler models with few variables and parameters, the input-output equations can be used directly to identify the value of the time-lag from measured data. We have formulated linear-algebraic criteria to check the identifiability of the delay parameter which eliminate the need for an explicit calculation of the input-output relations.

Work is in progress to extend the results of this paper to nonlinear systems with multiple time-delays.

Acknowledgements

This work was supported by the National Research School in Genomics and Bioinformatics, the Swedish Research Council and the Swedish Foundation for Strategic Research via the Gothenburg Mathematical Modeling Center.

References

Belkoura, L., & Orlov, Y.. (2002). Identifiability analysis of linear delay-differential systems. *IMA Journal of Mathematical Control and Information, 19*, 73-81.

Cohn, P. M. (1971). Free rings and their relations. London: Academic Press.

Conte, G., & Perdon, A. M. (1984). The disturbance decoupling problem for systems over a ring. *SIAM Journal on Control and Optimization, 33*, 750-764.

Conte, G., Moog, C. H., & Perdon, A. M. (1999). *Nonlinear control systems: An algebraic setting, Lecture notes in control and information sciences, 242.* London: Springer-Verlag.

Forsman, K., & Habets, L.. (1994). Input-Output Equations and Observability for Polynomial Delay Systems. *In Proceedings of the 33rd IEEE Conference on Decision and Control, December 1994*, 880 - 882.

Hermann, R., & Krener, A. J. (1977). Nonlinear controllability and observability. *IEEE Transactions on Automatic Control, 22*, 728-740.

Márquez-Martínez, L. A., Moog, C. H., & Velasco-Villa, M. (2000). The structure of nonlinear time delay systems. *Kybernetika, 36*, 53-62.

Moog, C. H., Castro-Linares, R., Velasco-Villa, M.,& Márquez-Martínez, L. A. (2000). The disturbance decoupling problem for time-delay nonlinear systems. *IEEE Transactions on Automatic Control, 45*, 305-309.

Nakagiri, S., & Yamamoto, M. (1995). Unique identification of coefficient matrices, time delays and initial functions of functional differential equations. *Journal of Mathematical Systems, Estimation and Control, 5*, 1-22.

Orlov, Y., Belkoura, L., Richard, J. P., Dambrine, M. (2002). On identifiability of linear time-delay systems. *IEEE Transactions on Automatic Control, 47*, 1319-1324.

Pohjanpalo. H. (1978). System identifiability based on the power series expansion of the solution. *Mathematical Biosciences, 41*, 21-33.

Sedoglavic, A. (2002). A probabilistic algorithm to test local algebraic observability in polynomial time. *Journal of Symbolic Computation, 33*, 735-755.

Sontag, E. D. (1991). I/O equations for nonlinear systems and observation spaces. *Proceedings of the 30th Conference on Decision and Control*, 720-725. Brighton-England.

Vajda, S., Godfrey, K., & Rabitz, H.. (1989). Similarity transformation approach to identifiability analysis of nonlinear compartmental models. *Mathematical Biosciences, 93*, 217-248.

Verduyn Lunell, S. M. (2001). Parameter identifiability of differential delay equations. *International Journal of Adaptive Control and Signal Processing, 15*, 655-678.

Xia, X., Márquez, L. A., Zagalak, P., & Moog, C. H. (2002). Analysis of nonlinear time-delay systems using modules over noncommutative rings. *Automatica, 38*, 1549-1555.

Zhang, J., Xia, X., & Moog, C. H. (2006). Parameter identifiability of nonlinear systems with time-delay. *IEEE Transactions on Automatic Control, 47*, 371-375.

Paper III

IDENTIFIABILITY OF THE TIME-LAG PARAMETER IN DELAY SYSTEMS WITH APPLICATIONS TO SYSTEMS BIOLOGY

Milena Anguelova [*,**,1] **Bernt Wennberg** [*,**]

[*] *Mathematical Sciences, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden*
[**] *Mathematical Sciences, Göteborg University, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden*

Abstract:
Time-delay models are used increasingly to describe biological systems. The model parameters, including the time lags, are often unknown and estimated from experimental data. For parameter estimation to be meaningful, the identifiability of the parameters must be established.
The identifiability of the time-delay parameters is not directly related to the well-characterized identifiability of the other model parameters and an independent analysis must be performed. We have recently shown that the identifiability of the delay parameter is determined by the form of the external input-output representation of the system. Explicit criteria have been formulated to analyze the latter without the need for an explicit computation of the input-output equations. We apply the developed identifiability criteria to biological delay systems from the literature, modeling e.g. signal transduction and gene expression, and show that the criteria are often simplified and easily checkable in practice.

Keywords:
Identifiability, Delay differential equations, Time lag, Signal transduction, Gene expression.

1. INTRODUCTION

Time lags are often observed in biological systems and delay differential equations have been used successfully to model a wide range of phenomena in systems biology, including signal transduction (Timmer et al., 2004), metabolic insulin signaling (Sedaghat et al., 2002), phosphorylation-dephosphorylation cycles (Srividhya, 2007), the *lac* operon (Mahaffy and Savev, 1999), the circadian pacemaker (Lema et al., 2000; Smolen et al., 1999), gene expression in cultured mammalian cells (Monk, 2003) and in zebrafish (Lewis, 2003). These delay models usually contain a number of parameters with values that are often unknown and can only be found by parameter estimation from experimental data. A prerequisite for parameter estimation is the property of identifiability which guarantees that the desired quantities can be uniquely determined from the available data.

For systems without time delays there exist several methods to analyze identifiability, both from a differential-geometric (see for instance Hermann and Krener (1977), Pohjanpalo (1978), Vajda et

[1] To whom all correspondence should be addressed. E-mail: milena@math.chalmers.se.

al. (1989), Sontag (1991), Isidori (1995), Chappell et al. (1999) and Saccomani et al. (2001)) and an algebraic approach (Ollivier (1990), Conte et al. (1999), Sedoglavic (2002) and the references therein). Necessary and sufficient conditions have been formulated with checkable criteria for which computer algorithms can be used (Margaria et al., 2001; Saccomani et al., 2001; Sedoglavic, 2002). The characterization of these identifiability criteria has now been extended to nonlinear systems with time delays by Zhang et al. (2006), based on a mathematical framework introduced by Moog et al. (2000), and developed by Márquez-Martínez et al. (2000) and Xia et al. (2002). In these works the time delays themselves are assumed to be known, or multiples of a unit delay. The identifiability of general unknown time-delays has been analyzed for linear systems (Nakagiri et al. (1995), Belkoura et al. (2000), Verduyn Lunell (2001), Orlov et al. (2002) and Belkoura et al. (2002)). Recently, we used the mathematical setting of Moog et al. (2000), Márquez-Martínez et al. (2000), Xia et al. (2002) and Zhang et al. (2006) to analyze the identifiability of time-delay parameters for nonlinear systems with a single unknown constant time delay in Anguelova and Wennberg (2007).

The assumption of a known time delay may not be appropriate for many of the above-mentioned biological systems. In the work of Timmer et al. (2004) on signal transduction, a nonlinear time delay model of the JAK-STAT signalling pathway is shown to give the best agreement with experimental data. In the model, the sojourn time of STAT-5 (a member of STAT, the signal transduction and activator of transcription, family of transcription factors) in the nucleus is modeled by an unknown time delay which is estimated numerically in Timmer et al. (2004). This motivated us to apply the methodology developed in Anguelova and Wennberg (2007) to analyze the identifiability of delay parameter in the JAK-STAT model and the other examples from systems biology given above. It should here be noted that the identifiability of the delay parameters of a system is not directly related to the identifiability of the regular parameters and separate analyses are necessary.

The identifiability of the delay parameter is decided by the form of the external input-output representation of the system. The value of the delay parameter can be found directly from the input-output equations, if they can be obtained explicitly. This is, however, very difficult for highly nonlinear models with many parameters and variables. In the work cited above, we have formulated identifiability criteria based on rank calculations, which eliminate the need for an explicit calculation of the input-output relations. Once the identifiability of the delay parameter is established, its value can be found by parameter estimation for which many methods exist, based e.g. on least squares approximation (see the book by Banks et al. (1989)).

2. MATHEMATICAL FRAMEWORK

In this section we will give a brief and somewhat simplified description of the method to analyze the identifiability of delay parameters for nonlinear systems, developed in Anguelova and Wennberg (2007). We refer to the latter for a more rigorous mathematical treatment.

2.1 Identifying the delay parameter from explicit input-output equations

To introduce the theory, we start by an example of a simple delay model with a corresponding input-output structure for the purpose of parameter estimation. The model has two state variables, x_1 and x_2, two regular parameters k_1 and k_2, and one time-lag parameter, τ. The controlled input variable is denoted by u and the measured data by y. Initial conditions have to be defined for the input, $u_0(t)$ and state variables, $\varphi(t)$, which in the case of delay differential equations are functions, describing the history of the system at least τ units back in time. All parameters and initial conditions for the state variables are unknown, but we assume that τ must take values in the interval $[0, T)$, where T is known.

Example 1:

$$\begin{cases} \dot{x}_1(t) &= k_1 x_2(t-\tau) + u(t) \\ \dot{x}_2(t) &= k_2 x_2(t-\tau) \\ y(t) &= x_1(t) \\ x(t) &= \varphi(t), \quad t \in [-\tau, 0] \\ u(t) &= u_0(t), \quad t \in [-T, 0] \end{cases} \tag{1}$$

By taking time derivatives of the output at a given point in time, we obtain equations for the state variables and parameters:

$$\begin{aligned} \dot{y}(t) &= k_1 x_2(t-\tau) + u(t) \\ \ddot{y}(t) &= k_1 k_2 x_2(t-2\tau) + \dot{u}(t) \\ y^{(3)}(t) &= k_1 k_2^2 x_2(t-3\tau) + \ddot{u}(t) \end{aligned} \tag{2}$$

From the above equations, we can extract an external input-output representation of the system, given by the input-output equation

$$\begin{aligned} &(\ddot{y}(t) - \dot{u}(t))(\ddot{y}(t-\tau) - \dot{u}(t-\tau)) - \\ &-(y^{(3)}(t) - \ddot{u}(t))(\dot{y}(t-\tau) - u(t-\tau)) = 0 \end{aligned} \tag{3}$$

Since the input and output are known functions of time, evaluating the above equation at a chosen point t_0 in time enables us to calculate the time-lag parameter τ, for example, by (numerically) finding the zeros of the function $\xi_{t_0}(\tau) =$

$(\ddot{y}(t_0) - \dot{u}(t_0))(\ddot{y}(t_0-\tau)) - \dot{u}(t_0-\tau) - (y^{(3)}(t_0) - \ddot{u}(t_0))(\dot{y}(t_0-\tau) - u(t_0-\tau))$. As an illustration, we have used the `dde23.m` differential equation solver in Matlab (Shampine and Thompson , 2001) to simulate an output for the above system by choosing $k_1 = -2, k_2 = -3, \varphi_1(t) = t+1, \varphi_2(t) = t^2 + 1, u(t) = t$ and $\tau = 1$ and plotted the function $\xi_6(\tau) = (\ddot{y}(6) - \dot{u}(6))(\ddot{y}(6-\tau) - \dot{u}(6-\tau) - (y^{(3)}(6) - \ddot{u}(6))(\dot{y}(6-\tau) - u(6-\tau))$ for τ in the interval $[0, 2]$. As expected, this function takes the value zero for $\tau = 1$.

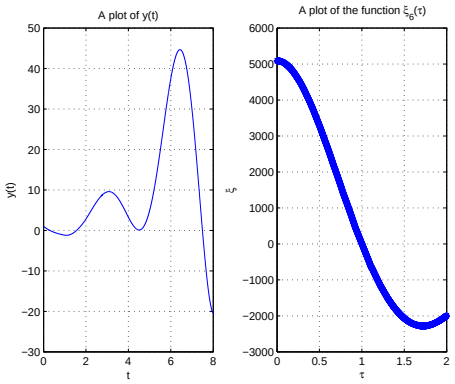


Fig. 1. *The output $y(t)$ and the function $\xi_6(\tau)$.*

On the other hand, it can be the case that the input-output equations of a system do not contain τ at all. Then there are infinitely many values that produce the same output and τ cannot be identified from the available data. This is demonstrated by the following example:

Example 2:

$$\begin{cases} \dot{x}_1 &= x_2^2(t-\tau) \\ \dot{x}_2 &= x_2 \\ y &= x_1 \\ x(t) &= \varphi(t), \quad t \in [-\tau, 0] \end{cases} \tag{4}$$

Calculating time-derivatives of the output function as above, we obtain

$$\begin{aligned} \dot{y} &= x_2^2(t-\tau) \\ \ddot{y} &= 2(x_2(t-\tau))^2 \end{aligned} \tag{5}$$

An output equation of lowest degree (of derivation) for the above system is $\ddot{y}(t) - 2\dot{y}(t) = 0$, which does not involve delays of the variables. We can see that τ is not identifiable for this example, by observing the following symmetry involving the functions of initial conditions φ and τ. For any choice of τ, setting

$$\begin{cases} \varphi_1(t) &= c \\ \varphi_2(t) &= e^{t+\tau} \end{cases} \quad , \quad t \in [-\tau, 0] \quad , \tag{6}$$

where c is a constant, leads to the solution

$$\begin{cases} x_1(t) &= \dfrac{e^{2t}}{2} + c - \dfrac{1}{2} \\ x_2(t) &= e^{t+\tau} \end{cases} \quad , \quad \forall t \geq 0 \quad . \tag{7}$$

Since $y(t) = x_1(t)$, it is clear that τ cannot be identified from the output.

2.2 Linear-algebraic identifiability criteria

The above way of deciding the identifiability of the delay parameter is, of course, difficult for more complicated models with many variables and parameters, when the explicit input-output equations cannot be obtained. We have therefore formulated identifiability criteria based on linear algebra calculations which can be performed also for more complex models of the form:

$$\begin{cases} \dot{x}(t) &= f(x(t), x(t-\tau), u, u(t-\tau)) \\ y(t) &= h(x(t), x(t-\tau), u, u(t-\tau)) \\ x(t) &= \varphi(t), \quad t \in [-\tau, 0] \\ u(t) &= u_0(t), \quad t \in [-T, 0] \end{cases} \quad , \tag{8}$$

where $x \in \mathbb{R}^n$ denotes the state variables, $u \in \mathbb{R}^m$ is the input and $y \in \mathbb{R}^p$ is the output. The unknown constant time-delay is denoted by $\tau \in [0, T), \quad T \in \mathbb{R}$. The entries of f and h are meromorphic in their arguments and $\varphi : [-\tau, 0] \to \mathbb{R}^n$ is an unknown continuous function of initial conditions. The set of initial functions for the variables x is denoted by $C := C([-\tau, 0], \mathbb{R}^n)$. A meromorphic function $u(t)$ is called an admissible input if the differential equation above admits a unique solution. The set of all such input functions is denoted by C_U. Observe that the above system form also allows for model parameters, which can simply be considered as state variables with time-derivative zero.

Following the notations and mathematical setting of Moog et al. (2000), Márquez-Martínez et al. (2000), Xia et al. (2002) and Zhang et al. (2006), let $\mathcal{K}$ be the field of meromorphic functions of a finite number of variables from $\{x(t-k\tau), u(t-k\tau), \ldots, u^{(l)}(t-k\tau), \quad k, l \in \mathbb{Z}^+\}$ and let δ denote the time-shift operator, $\delta(\xi(t)) = \xi(t-\tau)$, $\xi(t) \in \mathcal{K}$. Let $\mathcal{K}(\delta]$ denote the set of polynomials of the form

$$a(\delta] = a_0(t) + a_1(t)\delta + \cdots + a_{r_a}(t)\delta^{r_a} \quad , \tag{9}$$

where $a_j(t) \in \mathcal{K}$. With addition defined as usual and multiplication given by

$$a(\delta]b(\delta] = \sum_{k=0}^{r_a+r_b} \sum_{i+j=k}^{i \leq r_a, j \leq r_b} a_i(t) b_j(t-i\tau)\delta^k \quad , \tag{10}$$

$\mathcal{K}(\delta]$ is a noncommutative ring, which is Noetherian and a left Ore domain. The latter implies that the rank of a module over $\mathcal{K}(\delta]$ is well-defined (Cohn, 1971) and that row elimination can be performed in a matrix consisting of elements from $\mathcal{K}(\delta]$. This will be done for the matrix formed by the gradients of the time-derivatives of the output functions h_i.

Let $\mathcal{M}$ denote the module $span_{\mathcal{K}(\delta]}\{d\xi : \xi \in \mathcal{K}\}$. Differentiation of functions $\phi(x(t-i\tau), u(t-j\tau), \ldots, u^{(l)}(t-j\tau))$, $0 \leq i, j \leq k$, $l \geq 0$ in $\mathcal{K}$ and one-forms $\omega = \sum_i \kappa_x^i dx(t-i\tau) + \sum_{ij} \nu_i du^{(j)}(t-i\tau)$ in $\mathcal{M}$ is defined in the natural way:

$$\begin{aligned}\dot{\phi} &= \sum_{i=0}^{k} \frac{\partial \phi}{\partial x(t-i\tau)}\delta^i f + \\ &+ \sum_{r=0}^{l}\sum_{j=0}^{k} \frac{\partial \phi}{\partial u^{(r)}(t-j\tau)} u^{(r+1)}(t-j\tau) \\ \dot{\omega} &= \sum_i \dot{\kappa}_x^i dx(t-i\tau) + \sum_{ij} \dot{\nu}_i du^{(j)}(t-i\tau) + \\ &+ \sum_i \kappa_x^i d\delta^i f + \sum_{ij} \nu_i du^{(j+1)}(t-i\tau).\end{aligned} \tag{11}$$

The following definitions are necessary before we can state the identifiability criteria. Let f be an r-dimensional vector with entries $f_j \in \mathcal{K}$. Let $\frac{\partial f}{\partial x}$ denote the $r \times n$ matrix with entries

$$\left(\frac{\partial f}{\partial x}\right)_{j,i} = \sum_{\ell} \frac{\partial f_j}{\partial x_i(t-\ell\tau)}\delta^{\ell} \in \mathcal{K}(\delta]. \tag{12}$$

Denote by s_1 the least nonnegative integer such that

$$rank_{\mathcal{K}(\delta]} \frac{\partial(h_1, ..., h_1^{(s_1-1)})}{\partial x} = rank_{\mathcal{K}(\delta]} \frac{\partial(h_1, ..., h_1^{(s_1)})}{\partial x}. \tag{13}$$

If $\frac{\partial h_1}{\partial x} \equiv 0$ then we define $s_1 = 0$. Inductively, for $1 < i \leq p$ denote by s_i the least nonnegative integer such that

$$\begin{aligned} &rank_{\mathcal{K}(\delta]} \frac{\partial(h_1, ..., h_1^{(s_1-1)}, \ldots, h_i, \ldots, h_i^{(s_i-1)})}{\partial x} = \\ &= \quad rank_{\mathcal{K}(\delta]} \frac{\partial(h_1, ..., h_1^{(s_1)}, \ldots, h_i, \ldots, h_i^{(s_i)})}{\partial x} \end{aligned} \tag{14}$$

Let

$$S = (h_1, ..., h_1^{(s_1-1)}, \ldots, h_p, \ldots, h_p^{(s_p-1)}) \quad ,$$

where h_i does not appear if $s_i = 0$. Then

$$rank_{\mathcal{K}(\delta]} \frac{\partial S}{\partial x} = s_1 + \cdots + s_p = K \leq n \quad . \tag{15}$$

We can now state our identifiability criteria (Corollary 1 in Anguelova and Wennberg (2007)):

Corollary 1. Given a system of the form (8) and the set S defined above, τ is locally identifiable if and only if at least one of the following is true:

i. $\frac{\partial h_i^{(j)}(t)}{\partial u_r^{(k)}(t-s\tau)} \neq 0$ for some $1 \leq i \leq p$, $0 \leq j \leq s_i$, $s \geq 1$, $1 \leq r \leq m$ and $k \geq 0$, i.e. a delayed input-variable $u_r^{(k)}$ occurs in some of the functions in $\{S, h_1^{(s_1)}, \ldots, h_p^{(s_p)}\}$;

ii. $rank_{\mathcal{K}(\delta]} \frac{\partial S}{\partial x} \neq rank_{\mathcal{K}} \frac{\partial(S, h_1^{(s_1)}, ..., h_p^{(s_p)})}{\partial x}$.

If τ is not locally identifiable, system (8) can locally be realised as an ODE-system.

Put simply, the local identifiability of τ depends on the presence or absence of τ in the input-output equations for the system (that such equations exist is not trivial, see Anguelova and Wennberg (2007)). Whether τ is present in the i-o equations is, in its turn, decided either by the occurrence of a delayed input variable in the time-derivatives of the output functions or by a linear-algebraic criterion involving rank calculation for the set of gradients of the outputs and their time-derivatives over $\mathcal{K}$ and over $\mathcal{K}(\delta]$ (that is, row elimination without and with δ).

To illustrate the above identifiability criteria, we return to the examples from the previous subsection.

Example 1, revisited:

As no delayed input variables occur in the output derivatives above, we will check whether the second criterium, $rank_{\mathcal{K}(\delta]} \frac{\partial S}{\partial x,k} \neq rank_{\mathcal{K}} \frac{\partial(S, h_1^{(s_1)})}{\partial x,k}$ is fulfilled:

$$\begin{pmatrix} dy \\ d\dot{y} - du \\ d\ddot{y} - d\dot{u} \\ dy^{(3)} - d\ddot{u} \end{pmatrix} = \tag{16}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & k_1\delta & \delta x_2 & 0 \\ 0 & k_1k_2\delta^2 & k_2\delta^2 x_2 & k_1\delta^2 x_2 \\ 0 & k_1k_2^2\delta^3 & k_2^2\delta^3 x_2 & 2k_1k_2\delta^3 x_2 \end{pmatrix} \begin{pmatrix} dx_1 \\ dx_2 \\ dk_1 \\ dk_2 \end{pmatrix} \tag{17}$$

The above matrix has rank 4 over $\mathcal{K}$ and rank 3 over $\mathcal{K}(\delta]$. Thus, τ is locally identifiable even though the system as a whole is not weakly observable as the matrix above is not of full-rank over $\mathcal{K}(\delta]$ (according to the definitions in Xia et al. (2002) and Zhang et al. (2006) with a fixed τ, k_1 and x_2 are not identifiable/observable).

Example 2, revisited:

Performing the same analysis for this example, we obtain:

$$\begin{pmatrix} dy \\ d\dot{y} \\ d\ddot{y} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2\delta(x_2)\delta \\ 0 & 4\delta(x_2)\delta \end{pmatrix} \begin{pmatrix} dx_1 \\ dx_2 \end{pmatrix} \quad . \tag{18}$$

Clearly, the above matrix has rank 2 over both $\mathcal{K}$ and $\mathcal{K}(\delta]$. If τ is known, the system is weakly observable. However, τ is not identifiable.

3. APPLICATION TO EXAMPLES OF BIOLOGICAL SYSTEMS

Example 3: This example is of the biological model describing the JAK-STAT signalling pathway by Timmer et al. (2004) that was referred to in the introduction. The analysis of the identifiability of the delay parameter reduces to the observation of a delayed input-variable in the third time-derivative of one of the outputs. The parameter τ

is identified numerically in Timmer et al. (2004). The delay-system is given by:

$$\begin{cases} \dot{x_1} &= -k_1 x_1 u/k_7 + 2k_4 x_3(t-\tau) \\ \dot{x_2} &= k_1 x_1 u/k_7 - k_2 x_2^2 \\ \dot{x_3} &= -k_3 x_3 + 0.5 k_2 x_2^2 \\ \dot{x_4} &= k_3 x_3 - k_4 x_3(t-\tau) \\ y_1 &= k_5(x_2 + 2x_3) \\ y_2 &= k_6(x_1 + x_2 + 2x_3) \quad , \end{cases} \tag{19}$$

where we have denoted the external input $k_7 EpoR_A$ in Timmer et al. (2004) by u in accordance with the notation in this paper. We have

$$\begin{aligned} \dot{y}_2 &= 2k_6(k_4\delta x_3 - k_3 x_3) \qquad (20) \\ \ddot{y}_2 &= 2k_6(-k_3 k_4 \delta x_3 + 0.5 k_2 k_4 (\delta x_2)^2 + k_3^2 x_3 - \\ &\quad -0.5 k_2 k_3 x_2^2) \\ y_2^{(3)} &= k_6(-2k_3 k_2 x_2 k_1 x_1 u + 2k_3 k_2^2 x_2^3 k_7 - \\ &\quad -2k_3^3 k_7 x_3 + k_3^2 k_7 k_2 x_2^2 + 2k_4 k_2 \delta x_2 k_1 \delta x_1 \delta u - \\ &\quad -2k_4 k_2^2 \delta x_2^3 k_7 + 2k_4 k_3^2 k_7 \delta x_3 - \\ &\quad -k_4 k_3 k_7 k_2 \delta x_2^2)/k_7 \quad . \end{aligned}$$

We can now see that δu appears in $y_2^{(3)}$ which is enough to conclude that τ is identifiable according to *i.* in Corollary 1 (obviously $dy, d\dot{y}_2$ and $d\ddot{y}_2$ are linearly independent as they contain different parameters).

Example 4: In this example we analyze a gene expression model for Hes1 by Monk (2003). This time-delay model has two state variables P and M and six parameters $p = (\alpha_m, P_0, n, \mu_m, \alpha_p, \mu_p)$.

$$\begin{cases} \dot{M} &= \dfrac{\alpha_m}{1 + (P(t-\tau)/P_0)^n} - \mu_m M \\ \dot{P} &= \alpha_p M - \mu_p P \\ y_1 &= M \\ y_2 &= P \quad , \end{cases} \tag{21}$$

We obtain the equations

$$\dot{y_1} = \frac{\alpha_m}{1 + (\delta P/P_0)^n} - \mu_m M \tag{22}$$

$$\ddot{y_1} = h_1^{(2)}(\delta P, \delta M, P, M, \dots) \tag{23}$$

$$y_1^{(3)} = h_1^{(3)}(\delta^2 P, \delta P, \delta M, P, M, \dots) \tag{24}$$

$$y_1^{(4)} = h_1^{(3)}(\delta^2 M, \delta^2 P, \delta P, \delta M, P, M, \dots) \tag{25}$$

$$\dots$$

For each $j \geq 1$, either $\frac{\partial h_1^{(j)}}{\partial M}$ or $\frac{\partial h_1^{(j)}}{\partial P}$ is a polynomial in δ of degree higher than for $j-1$. This means that $rank_{\mathcal{K}} \frac{\partial (S, h_1, \dots, h_1^{(8)})}{\partial x} = 9$, which is then greater than $rank_{\mathcal{K}(\delta)} \frac{\partial S}{\partial x, p}$ as the latter is limited by the number of variables and parameters, eight in this case. This implies that τ is identifiable by Corollary 1.

It can be concluded from Example 4 that if each derivative of the output functions $h_j^{(j)}$, $j = 0, \dots, n$ contains a state-variable that is delayed compared to the previous function, then the delay parameter is identifiable. In such cases the identifiability of the delay can be decided without any rank calculations. This applies for example to the delay model of the circadian pacemaker by Lema et al. (2000), the model of the reduced phosphorylation-dephosphorylation network in Srividhya (2007) and the discrete-delay models of genetic regulatory systems in Smolen et al. (1999).

4. CONCLUSIONS

A recently developed method to analyze the identifiability of the time-lag parameter in nonlinear delay systems has been applied to biological models from the literature.

The identifiability of the time-delay parameter is decided by the form of the external input-output representation of the system. For simpler models with few variables and parameters, the input-output equations can be used directly to identify the time-lag from measured data. For more complex models, the formulated linear-algebraic identifiability criteria can be applied, which often reduce to a simple book-keeping of the delays in the input- and state-variables in each of the output function derivatives.

Work is in progress to extend the method to nonlinear systems with multiple time-delays.

5. ACKNOWLEDGEMENTS

Financial support from the National Research School in Genomics and Bioinformatics, Göteborg University, Sweden, is gratefully acknowledged.

Bernt Wennberg was partially supported by the Swedish Research Council, and by the Swedish Foundation for Strategic Research via the Gothenburg Mathematical Modelling Center.

REFERENCES

Anguelova, M & Wennberg, B. (2007). State elimination and identifiability of the delay parameter for nonlinear time-delay systems. *Provisionally accepted.*

Banks, H., & Kunisch, K.. (1989). *Estimation Techniques for Distributed Parameter Systems.* Boston: Birkhauser.

Belkoura, L., Dambrine, M., & Richard, J. P. (2000). Identifiability of delay systems: a distribution approach. *In Proceedings of 14th MTNS 2000, Mathematical Theory of Networks and Systems.* France-Perpignan.

Belkoura, L., & Orlov, Y.. (2002). Identifiability analysis of linear delay-differential systems. *IMA J. Math. Control I., 19*, 73-81.

Chappell, M. J., Evans, N. D., Godfrey, K. R., & Chapman, M. J.. (1999). Structural identifiability of controlled state space systems: a quasi-automated methodology for generating identifiable reparameterisations of unidentifiable systems. *IEE Colloquium on Symbolic Computation for Control (Ref. No. 1999/088)*, 8/1-12.

Cohn, P. M. (1971). Free rings and their relations. London: Academic Press.

Conte, G., & Perdon, A. M. (1984). The disturbance decoupling problem for systems over a ring. *SIAM J. Control Optim., 33*, 750-764.

Conte, G., Moog, C. H., & Perdon, A. M. (1999).*Nonlinear control systems: An algebraic setting, Lecture notes in control and information sciences, 242.* London: Springer-Verlag.

Hermann, R., & Krener, A. J. (1977). Nonlinear controllability and observability. *IEEE T. Aut. Contr., 22*, 728-740.

Horbelt, W., Timmer, J., & Voss, H. U.. (2002). Parameter estimation in nonlinear delayed feedback systems from noisy data. *Phys. Lett. A., 299*, 513-521.

Isidori, A. (1995) *Nonlinear control systems.* Springer-Verlag.

Lema, M. A., Golombek, D. A., & Echave, J.. (2000). Delay model of the circadian pacemaker. *J. Theor. Biol., 204*, 565-573.

Lewis, J. Autoinhibition with transcriptional delay: a simple mechanism for the zebrafish somitogenesis oscillator. (2003). *Curr. Biol., 13*, 1398-1408.

J. M. Mahaffy, J. M., & Savev, E. S.. (1999). Stability analysis for a mathematical model of the *lac* operon. *Q. Appl. Math., LVII*, 37-53.

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

Márquez-Martínez, L. A., Moog, C. H., & Velasco-Villa, M. (2000). The structure of nonlinear time delay systems. *Kybernetika, 36*, 53-62.

Shampine, L. F. & Thompson S.. Solving DDEs in MATLAB. *Appl. Numer. Math., 37*, 441-448.

Monk, N. A. M. (2003). Oscillatory expression of Hes1, p53, and NF-κB driven by transcriptional time delays. *Curr. Biol., 13*, 1409-1413.

Moog, C. H., Castro-Linares, R., Velasco-Villa, M.,& Márquez-Martínez, L. A. (2000). The disturbance decoupling problem for time-delay nonlinear systems. *IEEE T. Aut. Contr., 45*, 305-309.

Nakagiri, S., & Yamamoto, M. (1995). Unique identification of coefficient matrices, time delays and initial functions of functional differential equations. *Journal of Mathematical Systems, Estimation and Control, 5*, 1-22.

Ollivier, F. (1990). *Le problème de l'identifiabilité structurelle globale: approche théorique, méthodes effectives et bornes de complexité.* PhD thesis, École polytechnique.

Orlov, Y., Belkoura, L., Richard, J. P., Dambrine, M. (2002). *IEEE T. Aut. Contr., 47*, 1319-1324.

Pohjanpalo. H. (1978). System identifiability based on the power series expansion of the solution. *Math. Biosci., 41*, 21-33.

Saccomani, M. P., Audoly, S., Bellu, G., & D'Angiò, L. (2001). A new differential algebra algorithm to test identifiability of nonlinear systems with given initial conditions. *Proceedings of the 40th Conference on Decision and Control*, 3108-3113. Orlando-Florida USA.

Sedaghat, A. R., Sherman, A., & Quon, M. J. (2002). A mathematical model of metabolic insulin signaling pathways. *Am. J. Endocrinol. Metab., 283*, 1084-1101.

Sedoglavic, A. (2002). A probabilistic algorithm to test local algebraic observability in polynomial time. *J. Symb. Comput., 33*, 735-755.

Smolen, P., Baxter, D. A., & Byrne, J. H.. (1999). Effects of macromolecular transport and stochastic fluctuations on dynamics of genetic regulatory systems. *Am. J. Cell Physiol., 277*, 777-790.

Sontag, E. D. (1991). I/O equations for nonlinear systems and observation spaces. *Proceedings of the 30th Conference on Decision and Control*, 720-725. Brighton-England.

Srividhya, J., Gopinathan, M. S., & Schnell, S. (2007). The effects of time delays in a phosphorylation-dephosphorylation pathway. *Biophys. Chem., 125*, 286-297.

Timmer, J., Müller, T. G., Swameye, I., Sandra, O., & Klingmüller, U. (2004). Modelling the nonlinear dynamics of cellular signal transduction. *Int. J. Bifurcat. Chaos, 14*, 2069-2079.

Vajda, S., Godfrey, K., & Rabitz, H.. (1989). Similarity transformation approach to identifiability analysis of nonlinear compartmental models. *Math. Biosci., 93*, 217-248.

Verduyn Lunell, S. M. (2001). Parameter identifiability of differential delay equations. *Int. J. Adapt. Control, 15*, 655-678.

Xia, X., Márquez, L. A., Zagalak, P., & Moog, C. H. (2002). Analysis of nonlinear time-delay systems using modules over noncommutative rings. *Automatica, 38*, 1549-1555.

Zhang, J., Xia, X., & Moog, C. H. (2006). Parameter identifiability of nonlinear systems with time-delay. *IEEE T. Aut. Contr., 47*, 371-375.

Paper IV

STATE ELIMINATION AND IDENTIFIABILITY OF DELAY PARAMETERS FOR NONLINEAR SYSTEMS WITH MULTIPLE TIME-DELAYS

Milena Anguelova [*,**,1] **Bernt Wennberg** [*,**]

[*] *Mathematical Sciences, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden*
[**] *Mathematical Sciences, Göteborg University, SE-412 96 Gothenburg, Sweden*

Abstract: This paper considers the identifiability of time-lag parameters in nonlinear delay differential equations. Their identifiability is not directly related to the well-characterized identifiability/observability of the other model parameters/state variables and an independent analysis must be performed. We show the existence of input-output relations for nonlinear delay systems with multiple constant time-delays and relate the identifiability of the delay parameters to their form. The values of the time lags can be found directly from the input-output equations, if these can be obtained explicitly. Linear-algebraic criteria are formulated to decide the identifiability of the delay parameters which eliminate the need for explicit computation of the input-output relations.

Keywords: identifiability, delay differential equations, time lags, state elimination.

1. INTRODUCTION

Observability and parameter identifiability are important properties of a system where initial state or parameter estimation are concerned. These properties guarantee that the desired quantities can be uniquely determined from the available data.

For nonlinear systems without time delays, these properties are well-characterised, see for instance (Hermann and Krener, 1977; Pohjanpalo, 1978; Vajda et al., 1989; Sedoglavic, 2002) and the references therein. The characterization of observability and identifiability has now been extended to nonlinear systems with time delays by Xia et al. (2002) and Zhang et al. (2006), using an algebraic approach introduced by Moog et al. (2000), and developed by Márquez-Martínez et al. (2000). In these works the time delays themselves are assumed to be known, or multiples of a unit delay. The identifiability of general unknown time-delays has been analyzed only for linear systems (Nakagiri et al., 1995; Verduyn Lunell, 2001; Orlov et al., 2002; Belkoura et al., 2002). Recently, we used the mathematical setting of (Moog et al., 2000; Márquez-Martínez et al., 2000) and (Xia et al., 2002) to analyze the identifiability of the time-delay parameter for nonlinear systems with a single unknown constant time delay in (Anguelova and Wennberg, 2007).

In this paper we analyze the identifiability of the time-delay parameters for nonlinear control systems with several unknown constant time delays. It is shown that state elimination produces input-output relations for the system, the form of which decides the identifiability of the delay parameters. The values of the delay parameters can be found

[1] To whom all correspondence should be addressed. E-mail: milena@math.chalmers.se.

directly from the input-output equations, if these can be obtained explicitly. We formulate linear-algebraic criteria to check the identifiability of the delay parameters which eliminate the need for an explicit calculation of the input-output relations. We observe that the identifiability of the delay parameters can be a necessary but not sufficient condition for the observability of the state variables (and identifiability of the regular parameters in the system). The already established methods for testing weak observability and identifiability for nonlinear delay systems alone (Xia et al., 2002; Zhang et al., 2006) cannot be used to determine the identifiability of the time-delay parameters and a prior analysis is necessary for the latter.

2. NOTATION AND PRELIMINARY DEFINITIONS

Consider nonlinear time-delay systems of the form:

$$\begin{cases} \dot{x}(t) = f(x(t), x(t-\boldsymbol{\tau}), u, u(t-\boldsymbol{\tau})) \\ y(t) = h(x(t), x(t-\boldsymbol{\tau})) \\ x(t) = \varphi(t), \quad t \in [-max_i \tau_i, 0] \\ u(t) = u_0(t), \quad t \in [-T, 0] \end{cases}, \quad (1)$$

where $x \in \mathbb{R}^n$ denotes the state variables, $u \in \mathbb{R}^m$ is the input and $y \in \mathbb{R}^p$ is the output. The unknown constant time-delays are denoted by the vector $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_\ell)$, $\tau_i \in [0, T)$, $\quad T \in \mathbb{R}$. The expression $x(t-\boldsymbol{\tau})$ means the set of variables $\{x(t-\tau_1), \ldots, x(t-\tau_\ell)\}$, and analogously for u. The entries of f and h are meromorphic in their arguments and $\varphi : [-max_i(\tau_i), 0] \to \mathbb{R}^n$ is an unknown continuous function of initial conditions. The set of initial functions for the variables x is denoted by $C := C([-max_i(\tau_i), 0], \mathbb{R}^n)$. A meromorphic input function $u(t)$ is called an admissible input if the differential equation above admits a unique solution. The set of all such input functions is denoted by C_U.

One of the objectives of this work is to investigate the property of local identifiability of the delay parameters τ_i. Intuitively, τ_i are identifiable if any two sets can be distingushed by the system's input-output behaviour. A formal definition is as follows:

Definition 1. The delay parameters $\boldsymbol{\tau}$ are said to be locally identifiable at $\boldsymbol{\tau_0} \in (0, T)^\ell$ if there exists an open set $W \ni \boldsymbol{\tau_0}$, $W \subset [0, T)^\ell$, such that $\forall \boldsymbol{\tau_1} \in W : \boldsymbol{\tau_1} \neq \boldsymbol{\tau_0}$, $\forall \varphi_0, \varphi_1 \in C$, there exist $t \geq 0$ and $u \in C_U$ such that $y(t, \varphi_1, u, \boldsymbol{\tau_1}) \neq y(t, \varphi_0, u, \boldsymbol{\tau_0})$, where $y(t, \varphi, u, \boldsymbol{\tau})$ denotes the parameterized output for the initial function φ, the admissible input u and delays $\boldsymbol{\tau}$.

Following the notations and algebraic setting of Moog et al. (2000); Márquez-Martínez et al. (2000); Xia et al. (2002) and Zhang et al. (2006), let $\mathcal{K}$ be the field of meromorphic functions of a finite number of variables from $\{x(t-\boldsymbol{i\tau}), u(t-\boldsymbol{i\tau}), \ldots, u^{(l)}(t-\boldsymbol{i\tau}), \quad \boldsymbol{i} = (i_1, \ldots, i_\ell), i_j, l \in \mathbb{Z}^+\}$, where we have denoted $i_1\tau_1 + \cdots + i_\ell\tau_\ell$ by $\boldsymbol{i\tau}$. Let $\mathcal{K}(\boldsymbol{\delta})$ denote the set of polynomials in $\delta_1, \ldots, \delta_\ell$ with coefficients from $\mathcal{K}$. This set is a noncommutative ring, where addition is defined as usual, while multiplication is given by

$$a(\boldsymbol{\delta})b(\boldsymbol{\delta}) = \sum_{\boldsymbol{k}=0}^{\boldsymbol{r_a}+\boldsymbol{r_b}} \sum_{\boldsymbol{i}+\boldsymbol{j}=\boldsymbol{k}}^{\boldsymbol{i}\leq\boldsymbol{r_a}, \boldsymbol{j}\leq\boldsymbol{r_b}} a_i(t) b_j(t-\boldsymbol{i\tau})\boldsymbol{\delta^k} \quad , \quad (2)$$

where we have denoted $\delta_1^{k_1} \ldots \delta_\ell^{k_\ell}$ by $\boldsymbol{\delta^k}$. The ring $\mathcal{K}(\boldsymbol{\delta})$ is Noetherian and a left Ore domain. The latter implies that the rank of a free module over $\mathcal{K}(\boldsymbol{\delta})$ is well-defined (Cohn, 1971). Let $\mathcal{M}$ denote the module $span_{\mathcal{K}(\delta)}\{d\xi : \xi \in \mathcal{K}\}$. The closure of a submodule $\mathcal{N}$ in $\mathcal{M}$ is the submodule $\overline{\mathcal{N}} = \{w \in \mathcal{M} : \exists a(\boldsymbol{\delta}] = a(\delta_1, \ldots, \delta_\ell] \in \mathcal{K}(\boldsymbol{\delta}],\ a(\boldsymbol{\delta}]w \in \mathcal{N}\}$. $\overline{\mathcal{N}}$ is the largest submodule of $\mathcal{M}$ containing $\mathcal{N}$ and having a rank equal to $rank_{\mathcal{K}(\boldsymbol{\delta})}\mathcal{N}$ (Xia et al., 2002).

Differentiation of functions $\phi(x(t-\boldsymbol{i\tau}), u(t-\boldsymbol{j\tau}), \ldots, u^{(l)}(t-\boldsymbol{j\tau}))$ in $\mathcal{K}$ and one-forms $\omega = \sum_{\boldsymbol{i}} \kappa_x^{\boldsymbol{i}} dx(t-\boldsymbol{i\tau}) + \sum_{\boldsymbol{j},r} \nu_{\boldsymbol{j}} du^{(r)}(t-\boldsymbol{j\tau})$ in $\mathcal{M}$ is defined in the natural way (Xia et al., 2002; Zhang et al., 2006):

$$\dot{\phi} = \sum_{\boldsymbol{i}} \frac{\partial \phi}{\partial x(t-\boldsymbol{i\tau})} \boldsymbol{\delta^i} f + \sum_{\boldsymbol{j},r} \frac{\partial \phi}{\partial u^{(r)}(t-\boldsymbol{j\tau})} u^{(r+1)}(t-\boldsymbol{j\tau}) \quad (3)$$

$$\dot{\omega} = \sum_{\boldsymbol{i}} \dot{\kappa}_x^{\boldsymbol{i}} dx(t-\boldsymbol{i\tau}) + \sum_{\boldsymbol{j},r} \dot{\nu}_{\boldsymbol{j}} du^{(r)}(t-\boldsymbol{j\tau}) + \sum_{\boldsymbol{i}} \kappa_x^{\boldsymbol{i}} d\boldsymbol{\delta^i} f + \sum_{\boldsymbol{j},r} \nu_{\boldsymbol{j}} du^{(r+1)}(t-\boldsymbol{j\tau}) \quad (4)$$

Define

$$\mathcal{X} = span_{\mathcal{K}(\boldsymbol{\delta})}\{dx\} \quad (5)$$

$$\mathcal{Y}_k = span_{\mathcal{K}(\boldsymbol{\delta})}\{dy, d\dot{y}, \ldots, dy^{(k-1)}\} \quad (6)$$

$$\mathcal{U} = span_{\mathcal{K}(\boldsymbol{\delta})}\{du, d\dot{u}, \ldots\} \quad . \quad (7)$$

Then $(\mathcal{Y}_k + \mathcal{U}) \cap \mathcal{X} = (\mathcal{Y}_n + \mathcal{U}) \cap \mathcal{X}$ for $k \geq n$ and $rank_{\mathcal{K}(\boldsymbol{\delta})}(\mathcal{Y}_n + \mathcal{U}) \cap \mathcal{X} \leq n$ (Xia et al., 2002).

3. STATE ELIMINATION

In this section we consider the problem of obtaining an input-output representation from the state-space form of a time-delay control system. This problem has been treated for polynomial systems in Forsman et al. (1994). We will show that for a system of the form (1) there always exists, at least locally, a set of input-output delay-differential equations of the form:

$$F(\boldsymbol{\delta}, y, \dots, y^{(k)}, u, \dots, u^{(J)}) := \\ = F(y(t - \boldsymbol{i_0\tau}), \dots, y^{(k)}(t - \boldsymbol{i_k\tau}), \\ u(t - \boldsymbol{j_0\tau}), \dots, u^{(J)}(t - \boldsymbol{j_l\tau})) = 0\,, \tag{8}$$

such that any pair $(y(t), u(t))$ which solves (1), also satisfies (8), for t such that all derivatives involved are continuous. The function F is meromorphic in its arguments.

Theorem 1. Given a system of the form (1), there exists an integer $J \geq 0$ and an open dense subset V of $C \times C_U^{J+1}$, such that in the neighborhood of any point of V, there exists an input-output representation of the system of the form (8).

Proof: The proof is an adaptation of the proof of Theorem 2.2.1. in Conte et al. (1999) for the analogous result for ODE-systems.

Let f be an r-dimensional vector with entries $f_j \in \mathcal{K}$. Let $\frac{\partial f}{\partial x}$ denote the $r \times n$ matrix with entries

$$\left(\frac{\partial f}{\partial x}\right)_{j,i} = \sum_{\boldsymbol{k}} \frac{\partial f_j}{\partial x_i(t - \boldsymbol{k\tau})} \boldsymbol{\delta^k} \in \mathcal{K}(\boldsymbol{\delta}]\,. \tag{9}$$

Denote by s_1 the least nonnegative integer such that

$$rank_{\mathcal{K}(\delta]} \frac{\partial(h_1, ..., h_1^{(s_1-1)})}{\partial x} = rank_{\mathcal{K}(\boldsymbol{\delta}]} \frac{\partial(h_1, ..., h_1^{(s_1)})}{\partial x}\,. \tag{10}$$

If $\frac{\partial h_1}{\partial x} \equiv 0$ then we define $s_1 = 0$. Inductively, for $1 < i \leq p$ denote by s_i the least nonnegative integer such that

$$rank_{\mathcal{K}(\boldsymbol{\delta}]} \frac{\partial(h_1, ..., h_1^{(s_1-1)}, \dots, h_i^{(s_i-1)})}{\partial x} = \\ = rank_{\mathcal{K}(\boldsymbol{\delta}]} \frac{\partial(h_1, ..., h_1^{(s_1-1)}, \dots, h_i^{(s_i)})}{\partial x}\,. \tag{11}$$

Let $S = (h_1, ..., h_1^{(s_1-1)}, \dots, h_p^{(s_p-1)})$, where h_i does not appear if $s_i = 0$. Then

$$rank_{\mathcal{K}(\boldsymbol{\delta}]} \frac{\partial S}{\partial x} = s_1 + \dots + s_p = K \leq n\,. \tag{12}$$

If $K < n$, there exist meromorphic functions $g_1(\boldsymbol{\delta}, x), \dots, g_{n-K}(\boldsymbol{\delta}, x)$ such that $rank_{\mathcal{K}(\boldsymbol{\delta}]} \frac{\partial(S, g_1, \dots, g_{n-K})}{\partial x} = n$.

For simplicity, introduce the notation

$$\begin{cases} \tilde{x}_1 & = h_1 \\ \dots & \\ \tilde{x}_{s_1} & = h_1^{(s_1-1)} \\ \tilde{x}_{s_1+1} & = h_2 \\ \dots & \\ \tilde{x}_{s_1+s_2} & = h_2^{(s_2-1)} \\ \dots & \\ \tilde{x}_{s_1+\dots+s_p} & = h_p^{(s_p-1)} \\ \tilde{x}_{s_1+\dots+s_p+1} & = g_1 \\ \dots & \\ \tilde{x}_n & = g_{n-K} \end{cases}\,. \tag{13}$$

The one-forms $d\tilde{x}_i, \quad i = 1, \dots, n$ now form a basis of $\mathcal{X}$. The definition of s_i (equation (11)) implies that $\frac{\partial h_i^{(s_i)}}{\partial x}$ is in

$$\overline{span_{\mathcal{K}(\boldsymbol{\delta}]} \left\{ \frac{\partial(h_1, \dots, h_1^{(s_1-1)}, \dots, h_i^{(s_i-1)})}{\partial x} \right\}}\,. \tag{14}$$

Thus there exist nonzero polynomials $b_i(\boldsymbol{\delta}] \in \mathcal{K}(\boldsymbol{\delta}]$, $i = 1, \dots, p$ such that $b_i(\boldsymbol{\delta}] \frac{\partial h_i^{(s_i)}}{\partial x}$ is in $span_{\mathcal{K}(\boldsymbol{\delta}]} \left\{ \frac{\partial(h_1, \dots, h_1^{(s_1-1)}, \dots, h_i^{(s_i-1)})}{\partial x} \right\}$. Therefore, $b_i(\boldsymbol{\delta}) dh_i^{(s_i)} + \sum_{r=1}^{m} \sum_{j=0}^{J} c_{j,r}(\boldsymbol{\delta}) du_r^{(j)} \in span_{\mathcal{K}(\boldsymbol{\delta}]}\{d\tilde{x}_1, \dots, d\tilde{x}_{s_1+\dots+s_i}\}$, for some $J \geq 0$, where J is the highest derivative of u appearing in the functions in S and $c_{j,r}(\boldsymbol{\delta}] \in \mathcal{K}(\boldsymbol{\delta}]$. Hence,

$$b_i(\boldsymbol{\delta}] dh_i^{(s_i)} + \sum_{r=1}^{m} \sum_{j=0}^{\gamma} c_{j,r}(\boldsymbol{\delta}] du_r^{(j)} - \\ - \sum_{j=1}^{s_1+\dots+s_i} a_j(\boldsymbol{\delta}] d\tilde{x}_j = 0\,, \tag{15}$$

for some $a_j(\boldsymbol{\delta}] \in \mathcal{K}(\boldsymbol{\delta}]$. Since all functions are assumed meromorphic and we have continuous dependence for the output on the input and initial function, the above equality holds on an open dense set of $C \times C_U^{J+1}$. The left hand side of equation (15), being equal to zero, is a closed one-form on $\mathcal{M}$. Applying the Poincaré lemma, we obtain functions $\xi_i(t) \in \mathcal{K}$ such that $d\xi_i = b_i(\boldsymbol{\delta}] dh_i^{(s_i)} + \sum_{r=1}^{m} \sum_{j=0}^{J} c_{j,r}(\boldsymbol{\delta}] du_r^{(j)} - \sum_{j=1}^{s_1+\dots+s_i} a_j(\boldsymbol{\delta}] d\tilde{x}_j$ and $\xi_i(\boldsymbol{\delta}, h_i^{(s_i)}, \tilde{x}, u, \dots, u^{(J)}) = 0$, for each $i = 1, \dots, p$. The function ξ_i does not depend on $\tilde{x}_j$, $j > s_1 + \dots + s_i$, since $d\xi_i = 0$ would then contain terms $d\tilde{x}_j$, $j > s_1 + \dots + s_i$ which is impossible by (15) due to the variables $d\tilde{x}$ being linearly independent over $\mathcal{K}(\boldsymbol{\delta}]$ by definition. Thus, we have obtained a relation

$$\xi_i(\boldsymbol{\delta}, h_i^{(s_i)}, \tilde{x}_1, \dots, \tilde{x}_{s_1+\dots+s_i}, u, \dots, u^{(J)}) = 0\,, \tag{16}$$

which together with (13) produces an input-output equation

$$\xi_i(\boldsymbol{\delta}, y_i^{(s_i)}, y_1, \dots, y_1^{(s_1-1)}, \dots, y_i^{(s_i-1)}, \\ u, \dots, u^{(J)}) = 0. \tag{17}$$

This is true for each $i, 1 \leq i \leq p$ resulting in p input-output equations of the form (8). ■

4. IDENTIFIABILITY OF THE DELAY PARAMETERS

This section deals with the problem of the identifiability of the time-delay parameters $\tau_i, \quad i = 1, \dots, \ell$ in (1). Linear-algebraic criteria are formulated, based on the linear form of the input-output equations (15) from the previous section.

The equations (15) give

$$a_i(\boldsymbol{\delta}]dy_i^{(s_i)} + \sum_{r=1}^{m}\sum_{j=0}^{\gamma} c_{j,r}(\boldsymbol{\delta}]du_r^{(j)} = \\ = \sum_{l=1}^{i}\sum_{j=0}^{s_l-1} a_{i,l,j}(\boldsymbol{\delta}]dy_l^{(j)} \quad , \qquad (18)$$

With no loss of generality, we assume that the polynomials $a_i(\boldsymbol{\delta})$ on the left-hand side of the above equations are irreducible. Let $a_i(\boldsymbol{\delta}] = \sum_{\boldsymbol{k}} a_{i_{\boldsymbol{k}}}\boldsymbol{\delta}^{\boldsymbol{k}}$, $c_{j,r}(\boldsymbol{\delta}] = \sum_{\boldsymbol{k}} c_{j,r_{\boldsymbol{k}}}\boldsymbol{\delta}^{\boldsymbol{k}}$ and $a_{i,l,j}(\boldsymbol{\delta}] = \sum_{\boldsymbol{k}} a_{i,l,j_{\boldsymbol{k}}}\boldsymbol{\delta}^{\boldsymbol{k}}$. Denote all the different monomials $\boldsymbol{\delta}^{\boldsymbol{k}}$ appearing above by $\Delta_{i1},\dots,\Delta_{iq}$ and the (integer) linear combinations of $\tau_1,\dots,\tau_\ell$ that they represent by $T_{i1},\dots,T_{iq}$. If some of the terms in $a_i(\boldsymbol{\delta}]$, $c_{j,r}(\boldsymbol{\delta}]$ or $a_{i,l,j}(\boldsymbol{\delta}]$ is a polynomial in $\boldsymbol{\delta}$ of degree zero, that is, the input-output equations contain undelayed variables, then we set Δ_{i0} equal to δ_0, where δ_0 denotes the identity operator, and the corresponding T_{i0} is zero. It is the combinations $T_{i0},\dots,T_{iq}$ of $\tau_1,\dots,\tau_\ell$ that determine the local identifiability of the latter.

Let Δ_{i_0} be the monomial $\boldsymbol{\delta}^{\boldsymbol{k}}$ in $a_i(\boldsymbol{\delta}]$ with smallest index $\boldsymbol{k}$ (ordered after $k_1,\dots,k_\ell$) - it is either equal to Δ_{i0} or is among the $\Delta_{i1},\dots,\Delta_{iq}$. The input-output equations corresponding to (18) can locally be written

$$y_i^{(s_i)} = \tilde{f}_i(\Delta_{i_0}^{-1}\Delta_{i1},\dots,\Delta_{i_0}^{-1}\Delta_{iq}, y_1,\dots,y_1^{(s_1-1)}, \\ \dots, y_i,\dots,y_i^{(s_i-1)}, y_i^{(s_i)}, u,\dots,u^{(\gamma)}) \qquad (19)$$

or

$$y_i{}^{(s_i)}(t) = \tilde{f}_i(y_1(t),\dots,y_i^{(s_i-1)}(t), u(t),\dots,u^{(\gamma)}(t) \\ y_1^{(s_1-1)}(t-T_{i1}+T_{i_0}),\dots,y_i^{(s_i)}(t-T_{i1}+T_{i_0}), \\ u(t-T_{i1}+T_{i_0}),\dots,u^{(\gamma)}(t-T_{i1}+T_{i_0}), \\ \dots, \\ y_1^{(s_1-1)}(t-T_{iq}+T_{i_0}),\dots,y_i^{(s_i)}(t-T_{iq}+T_{i_0}), \\ u(t-T_{iq}+T_{i_0}),\dots,u^{(\gamma)}(t-T_{iq}+T_{i_0})). \qquad (20)$$

Let

$$(T_{11}-T_{1_0},\dots,T_{1q}-T_{1_0},\dots,T_{p1}-T_{p_0}, \\ \dots,T_{pq}-T_{p_0})^{tr} = M(\tau_1,\dots,\tau_\ell)^{tr} \quad , \qquad (21)$$

where M is a $(1q+\dots+pq)\times\ell$ integer matrix and $\boldsymbol{v}^{tr}$ denotes the transpose of $\boldsymbol{v}$. We can now formulate the identifiability criteria for $\tau_1,\dots,\tau_\ell$ in the following proposition:

Proposition 1. If M is defined as above (21) and $y_i^{(s_i+j)}(t)$ is not identically equal to zero for any $0 \le j < iq-1$, $i=\{1,\dots,p\}$, then $\tau_1,\dots,\tau_\ell$ are locally identifiable generically, if and only if $rank(M)=\ell$.

Proof: Let first $rank(M)<\ell$, that is, there are infinitely many τ_i:s which give the same linear combinations $\tilde{\boldsymbol{\tau}} := \{T_{i1}-T_{i_0},\dots,T_{iq}-T_{i_0}\}$, $i=1,\dots,p$ and the input-output equations (20) cannot be used to identify the τ_i:s. We treat the case $\ell \ge 2$ and refer to Anguelova and Wennberg (2007) for the case of a single delay. If $rank(M)=0$, the proof is analogous to the proof in Anguelova and Wennberg (2007) and is therefore left out.

We will show that $\tau_1,\dots,\tau_\ell$ are not locally identifiable generically by showing that (1) can be represented locally as a neutral system with time lags $\tilde{\boldsymbol{\tau}}$. This system is obtained from the input-output equations (20) using the notation from Section 3

$$\begin{cases} \dot{\tilde{x}}_1 &= \tilde{x}_2 \\ \dots & \\ \dot{\tilde{x}}_{s_1-1} &= \tilde{x}_{s_1} \\ \dot{\tilde{x}}_{s_1} &= \tilde{f}_1(\dot{\tilde{x}}_{s_1}(t-\tilde{\boldsymbol{\tau}}),\tilde{x},\tilde{x}(t-\tilde{\boldsymbol{\tau}}),u,\dots, \\ & \quad u^{(\gamma)},u(t-\tilde{\boldsymbol{\tau}}),\dots,u^{(\gamma)}(t-\tilde{\boldsymbol{\tau}})) \\ \dot{\tilde{x}}_{s_1+1} &= \tilde{x}_{s_1+2} \\ \dots & \\ \dot{\tilde{x}}_{s_1+s_2-1} &= \tilde{x}_{s_1+s_2} \\ \dot{\tilde{x}}_{s_1+s_2} &= \tilde{f}_2(\dot{\tilde{x}}_{s_1+s_2}(t-\tilde{\boldsymbol{\tau}}),\tilde{x},\tilde{x}(t-\tilde{\boldsymbol{\tau}}),u, \\ & \quad \dots,u^{(\gamma)},u(t-\tilde{\boldsymbol{\tau}}),\dots,u^{(\gamma)}(t-\tilde{\boldsymbol{\tau}})) \\ \dots & \\ \dot{\tilde{x}}_{s_1+\dots+s_p} &= \tilde{f}_p(\dot{\tilde{x}}_{s_1+\dots+s_p}(t-\tilde{\boldsymbol{\tau}}),\tilde{x},\tilde{x}(t-\tilde{\boldsymbol{\tau}}),u, \\ & \quad \dots,u^{(\gamma)},u(t-\tilde{\boldsymbol{\tau}}),\dots,u^{(\gamma)}(t-\tilde{\boldsymbol{\tau}})) \\ \tilde{y}_1 &= \tilde{x}_1 \\ \tilde{y}_2 &= \tilde{x}_{s_1+1} \\ \dots & \\ \tilde{y}_{\tilde{p}} &= \tilde{x}_{1+s_1+\dots+s_{p-1}} \\ \tilde{x}(t) &= \tilde{\varphi}(t), \quad t\in[t_0-max_j\tilde{\tau}_j,t_0] \end{cases} . \qquad (22)$$

Any pair $(y(t),u(t))$ which solves the original system (1) also satisfies the above neutral system for $t_0 \ge max_i\tilde{\tau}_j$ such that all derivatives are continuous. Thus, the input-output behaviour of the system does not distinguish the infinitely many τ_i:s which give the same linear combinations $\tilde{\boldsymbol{\tau}}$ and $\tau_1,\dots,\tau_\ell$ are not locally identifiable generically.

Suppose now that $rank(M)=\ell$. We first observe that we then must have $q \ge \ell$ and we thus have at least ℓ different Δ_{iq}:s. Consider those of the equations (20), for which $iq \ge 1$. Evaluated at a fixed time point $t_0 \ge T$, (20) gives an equation for $T_{i1},\dots,T_{iq}$:

$$y_i^{(s_i)}(t_0) = \xi(y_1(t_0),\dots,y_i^{(s_i-1)}(t_0), \\ u(t_0),\dots,u^{(\gamma)}(t_0), \\ y_1^{(s_1-1)}(t_0-T_{i1}+T_{i_0}),\dots,y_i^{(s_i)}(t_0-T_{i1}+T_{i_0}), \\ u(t_0-T_{i1}+T_{i_0}),\dots,u^{(\gamma)}(t_0-T_{i1}+T_{i_0}), \\ \dots, \\ y_1^{(s_1-1)}(t_0-T_{iq}+T_{i_0}),\dots,y_i^{(s_i)}(t_0-T_{iq}+T_{i_0})), \\ u(t_0-T_{iq}+T_{i_0}),\dots,u^{(\gamma)}(t_0-T_{iq}+T_{i_0})). \qquad (23)$$

If the time-point t_0 is chosen large enough to ensure the existence of all time-derivatives involved

(which can be achieved by choosing for example $t_0 \geq (max_i s_i - 1)T$), then differentiating (23) with respect to time gives new equations for $T_{i1}, \ldots, T_{iq}$ which are independent, since $dy_i^{(j)}, j \geq 0$ are linearly independent over $\mathcal{K}$ due to $iq \geq 1$:

$$\begin{array}{l} y_i^{(s_i+j)}(t_0) = \dfrac{d^j}{dt^j}\xi(y_1(t), \ldots, y_i^{(s_i-1)}(t), \\ u(t), \ldots, u^{(\gamma+j)}(t), \\ y_1(t - T_{i1} + T_{i_0}), \ldots, y_i^{(s_i-1)}(t - T_{i1} + T_{i_0}, \\ u(t - T_{i1} + T_{i_0}), \ldots, u^{(\gamma+j)}(t - T_{i1} + T_{i_0}), \ldots, \\ y_1(t - T_{iq} + T_{i_0}), \ldots, y_i^{(s_i-1)}(t - T_{iq} + T_{i_0}), \\ u(t - T_{iq} + T_{i_0}), \ldots, u^{(\gamma+j)}(t - T_{iq} + T_{i_0}))_{|t_0}. \end{array} \tag{24}$$

Unless $y_i^{(s_i+j)}(t)$ is identically zero for some $0 \leq j < q-1$, the first q of these equations identify $T_{i1} - T_{i_0}, \ldots, T_{iq} - T_{i_0}$ locally (the rest can in some cases be used to analyze global identifiability). Since $rank(M) = \ell$, $\tau_i, i = 1, \ldots, \ell$ are generically defined uniquely by the locally identifiable linear combinations $T_{i1} - T_{i_0}, \ldots, T_{iq} - T_{i_0}$ (an obvious exception is the case of commensurate time-delays). Thus, all τ_i, $i = 1, \ldots, \ell$ are generically locally identifiable, which completes the proof. ■

5. EXAMPLES

In this section we illustrate the theory by simple examples and show that weak observability (and/or parameter identifiability for the regular model parameters) does not necessarily imply identifiability of the delay parameters, or vice versa. Thus, the already established methods for testing observability and identifiability for nonlinear delay systems (Xia et al., 2002; Zhang et al., 2006) cannot be used to determine the identifiability of the delay parameters.

Example 1:

$$\begin{cases} \dot{x}_1(t) &= x_2^2(t-\tau_1) + u(t) \\ \dot{x}_2(t) &= x_1(t-\tau_2)x_2(t) \\ y(t) &= x_1(t) \\ x(t) &= \varphi(t), \quad t \in [-\tau, 0] \end{cases} \tag{25}$$

We have

$$\begin{array}{l} \dot{y} = (\delta_1(x_2))^2 + u \\ \ddot{y} = 2\delta_1(x_2)\delta_1\delta_2(x_1)\delta_1(x_2) + \dot{u} = \\ \quad = 2(\delta_1(x_2))^2\delta_1\delta_2(x_1) + \dot{u} \end{array} \tag{26}$$

and

$$\begin{bmatrix} dy \\ d\dot{y} - du \\ d\ddot{y} - d\dot{u} \end{bmatrix} = \frac{\partial(S, h_1^{(s_1)})}{\partial x} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} = \\ = \begin{bmatrix} 1 & 0 \\ 0 & 2\delta_1(x_2)\delta_1 \\ 2(\delta_1(x_2))^2\delta_1\delta_2 & 4\delta_1\delta_2(x_1)\delta_1(x_2)\delta_1 \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix}. \tag{27}$$

Clearly, the matrix $\frac{\partial(S, h_1^{(s_1)})}{\partial x}$ has rank 2 over $\mathcal{K}(\boldsymbol{\delta}]$ and so the system is weakly observable according to the definition in Xia et al. (2002), if τ_1 and τ_2 are known. However, τ_1 and τ_2 are not identifiable. The input-output equation in linear form (eq. (18)) is:

$$\begin{array}{l} d\ddot{y} - d\dot{u} + 2\delta_1\delta_2(x_1)du = \\ = 2(\delta_1(x_2))^2\delta_1\delta_2 dy + 2\delta_1\delta_2(x_1)d\dot{y}, \end{array} \tag{28}$$

and we see that there are two monomials, Δ_0 and $\Delta_1 = \delta_1\delta_2$ with corresponding combinations $T_0 = 0$ and $T_1 = \tau_1 + \tau_2$ of the two time-delays. Thus, $M = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$ with rank 1, and the time lags are not identifiable.

Following the the first part of the proof of Corollary 1, we can use the change of variables $\tilde{x}_1 = y = x_1$, $\tilde{x}_2 = \dot{y} = (\delta_1(x_2))^2 + u$ to rewrite the system as

$$\begin{cases} \dot{\tilde{x}}_1(t) &= \tilde{x}_2(t) \\ \dot{\tilde{x}}_2(t) &= 2(\tilde{x}_2(t) - u(t))\tilde{x}_1(t - T_1) + \dot{u}(t) \\ y(t) &= \tilde{x}_1(t) \end{cases}. \tag{29}$$

Example 2:

$$\begin{cases} \dot{x}_1(t) &= -x_2(t-\tau_1) \\ \dot{x}_2(t) &= x_1(t-\tau_2) \\ y_1(t) &= x_1(t) \\ y_2(t) &= x_2(t-\tau_2) \\ x &= \varphi(t), \quad t \in [-T, 0] \end{cases} \tag{30}$$

We have

$$\begin{array}{l} \dot{y_1} = -\delta_1 x_2 \\ \ddot{y}_1 = -\delta_1\delta_2 x_1 \end{array} \tag{31}$$

and $\frac{\partial(S, h_1^{(s_1)}, h_2^{(s_2)})}{\partial x} = \begin{bmatrix} 1 & 0 \\ 0 & \delta_1 \\ \delta_1\delta_2 & 0 \\ 0 & \delta_2 \end{bmatrix}$. The input-output equations in linear form are

$$\begin{array}{ll} d\ddot{y}_1 &= -\delta_1\delta_2 dy_1 \\ \delta_1 dy_2 &= -\delta_2 d\dot{y}_1 \Leftrightarrow dy_2 = -\delta_1^{-1}\delta_2 d\dot{y}_1 \end{array}. \tag{32}$$

We thus have Δ_{10}, $\Delta_{11} = \delta_1\delta_2$, $\Delta_{21} = \delta_1$ and $\Delta_{22} = \delta_2$ and thus, $T_{10} = T_{1_0} = 0$, $T_{11} = \tau_1 + \tau_2$, $T_{2_0} = T_{21} = \tau_1$, $T_{22} = \tau_2$ and $M = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ -1 & 1 \end{bmatrix}$, which is of rank 2. Thus τ_1 and τ_2 are identifiable.

For this simple example, we can actually calculate the values of the two time lags from the explicit input-output equations $\ddot{y}_1 = -\delta_1\delta_2 y_1$, $y_2 = -\delta_1^{-1}\delta_2\dot{y}_1$. To illustrate, we simulate outputs for this system with MATLAB's dde23 function (Shampine and Thompson, 2001), by setting $\tau_1 = 1$, $\tau_2 = \sqrt{2}$, $\varphi_1(t) = e^t$ and $\varphi_2(t) = t + 1$. We then plot $\mu_1(T_{11} - T_{10})|_{t_0} := \ddot{y}_1(t_0) + y_1(t_0 - T_{11} + T_{10})$ and $\mu_2(T_{22} - T_{21})|_{t_0} := y_2(t_0) + \dot{y}_1(t_0 - T_{22} + T_{21})$ for $t_0 = 4$. As expected, the functions are zero for $T_{11} - T_{10} = 1 + \sqrt{(2)}$ and $T_{22} - T_{21} =$

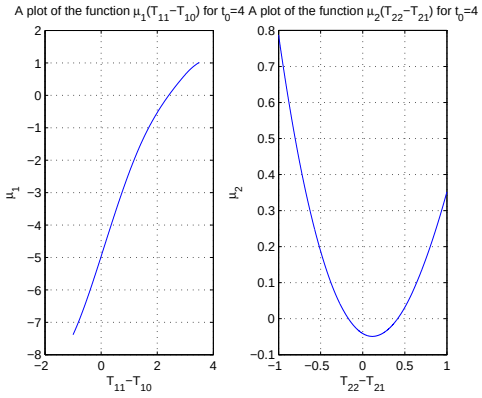


Fig. 1. *The functions* $\mu_1(T_{11}-T_{10})|_4$ *and* $\mu_2(T_{22}-T_{21})|_4$.

$\sqrt{(2)}-1$ and locally these are the only roots, see Fig. 1. Globally, there are also other roots for $\mu_2(T_{22}-T_{21})$ that can be seen in Fig. 1. They can be discarded in this case by plotting $\dot{\mu}_2(T_{22}-T_{21})|_{t_0} := \frac{d}{dt}(y_2(t_0)+\dot{y}_1(t_0-T_{22}+T_{21}))|_{t_0}$ (and other subsequent time-derivatives) since $T_{22}-T_{21}$ must be a root for this, too.

6. CONCLUSIONS

We have analyzed the identifiability of the time-lag parameters in nonlinear systems with multiple constant time delays.

State elimination is shown to yield an external input-output representation of the system, the form of which decides the identifiability of the delay parameters. For simpler models with few variables and parameters, the input-output equations can be used directly to identify the values of the time-lags from measured data.

We have formulated linear-algebraic criteria to check the identifiability of the delay parameters which eliminate the need for an explicit calculation of the input-output relations.

7. ACKNOWLEDGEMENTS

This work was supported by the National Research School in Genomics and Bioinformatics, the Swedish Research Council, and the Swedish Foundation for Strategic Research via the Gothenburg Mathematical Modeling Center.

REFERENCES

Anguelova, M & Wennberg, B. State elimination and identifiability of the delay parameter for nonlinear time-delay systems. *Provisionally accepted.*

Belkoura, L., & Orlov, Y.. (2002). Identifiability analysis of linear delay-differential systems. *IMA J. Math. Control I., 19*, 73-81.

Cohn, P. M. (1971). Free rings and their relations. London: Academic Press.

Conte, G., & Perdon, A. M. (1984). The disturbance decoupling problem for systems over a ring. *SIAM J. Control Optim., 33*, 750-764.

Conte, G., Moog, C. H., & Perdon, A. M. (1999).*Nonlinear control systems: An algebraic setting, Lecture notes in control and information sciences, 242.* London: Springer-Verlag.

Forsman, K., & Habets, L.. (1994). Input-Output Equations and Observability for Polynomial Delay Systems. *In Proceedings of the 33rd IEEE Conference on Decision and Control, December 1994*, 880 - 882.

Hermann, R., & Krener, A. J. (1977). Nonlinear controllability and observability. *IEEE T. Aut. Contr., 22*, 728-740.

Márquez-Martínez, L. A., Moog, C. H., & Velasco-Villa, M. (2000). The structure of nonlinear time delay systems. *Kybernetika, 36*, 53-62.

Moog, C. H., Castro-Linares, R., Velasco-Villa, M.,& Márquez-Martínez, L. A. (2000). The disturbance decoupling problem for time-delay nonlinear systems. *IEEE T. Aut. Contr., 45*, 305-309.

Nakagiri, S., & Yamamoto, M. (1995). Unique identification of coefficient matrices, time delays and initial functions of functional differential equations. *Journal of Mathematical Systems, Estimation and Control, 5*, 1-22.

Orlov, Y., Belkoura, L., Richard, J. P., Dambrine, M. (2002). On identifiability of linear time-delay systems. *IEEE T. Aut. Contr., 47*, 1319-1324.

Pohjanpalo. H. (1978). System identifiability based on the power series expansion of the solution. *Math. Biosci., 41*, 21-33.

Sedoglavic, A. (2002). A probabilistic algorithm to test local algebraic observability in polynomial time. *J. Symb. Comput., 33*, 735-755.

Shampine, L. F. & Thompson S.. Solving DDEs in MATLAB. *Appl. Numer. Math., 37*, 441-448.

Vajda, S., Godfrey, K., & Rabitz, H.. (1989). Similarity transformation approach to identifiability analysis of nonlinear compartmental models. *Math. Biosci., 93*, 217-248.

Verduyn Lunell, S. M. (2001). Parameter identifiability of differential delay equations. *Int. J. Adapt. Control, 15*, 655-678.

Xia, X., Márquez, L. A., Zagalak, P., & Moog, C. H. (2002). Analysis of nonlinear time-delay systems using modules over noncommutative rings. *Automatica, 38*, 1549-1555.

Zhang, J., Xia, X., & Moog, C. H. (2006). Parameter identifiability of nonlinear systems with time-delay. *IEEE T. Aut. Contr., 47*, 371-375.