# Limiting the parameter search space for dynamic models with rational kinetics using semi-definite programming

**Dirk Fey** * **Eric Bullinger** *

* *Systems and Modeling Research Unit, Institute Montefiore, University of Liège, Belgium*
*(email: fey@montefiore.ulg.ac.be, e.bullinger@montefiore.ulg.ac.be)*

**Abstract:** Estimation of kinetic parameters is a key step in modelling, as direct measurements are often expensive, time-consuming or even infeasible. The class of dynamic models in polynomial form is particularly relevant in systems biology and biochemical engineering, as those models naturally arise from modelling biochemical reactions using for instance mass action, Michaelis-Menten or Hill kinetics. Often the parameters are not uniquely identifiable for a given model structure and measurement set. Thus the question of which parameters are consistent or inconsistent with the data arises naturally. Here we present a method capable of proving inconsistency of entire parameter regions with the data. Based on the polynomial representation of the system, we formulate a feasibility problem that can be solved efficiently by semi-definite programming. The feasibility problem allows us to check consistency of entire parameter regions by using upper and lower bounds on the parameters. This drastically limits the search space for subsequent parameter estimation methods. In contrast to similar approaches in the literature, the here presented approach does not require a steady state assumption. Measurements at discrete time points are used, but neither regular sampling intervals, nor a time discretisation of the system is required. Measurement uncertainties are dealt with using upper and lower bounds on the measured states.

Keywords: identification, parameter estimation, model invalidation, sum of squares, semi-definite programming

## 1. INTRODUCTION

Modelling biological systems on the intracellular level has been a research topic for over half a century. For example, Hodgkin and Huxley (1952) explained the neuron function by means of a mathematical model of different ion channels. Hodgkin and Huxley were able to estimate the model parameters from experimental data, a challenging task still today in most biological systems. Nowadays, ever improving developments of experimental techniques provide more and more high quality experimental data, putting the task of identification into reachable scope. However, the straight forward application of systems theoretical methods to biology is impaired by certain particularities of biological systems (Wellstead *et al.*, 2008; Bullinger *et al.*, 2007). Difficulties concern large numbers of parameters, sparse data and reduced sets of possible stimulations.

Biological systems have particular system properties such as for example positivity and monotonicity (Sontag, 2005). Exploiting these properties has the potential of enhancing current identification techniques. For example, Farina *et al.* (2006) exploited the positivity of states and parameters to transform a mass action system into a parameter independent form that facilitated the identifiability analysis. Fey *et al.* (2008) showed that a similar transformation is also possible for systems with certain rational kincetics consisting of products of Hill-like terms and that nonlinear observers can be used to solve the estimation problem. The transformation into the parameter independent form transforms the parameter estimation problem into a state estimation problem. Solving the state estimation problem reliably requires globally convergent observers, which can be constructed either using Lie-algebraic methods or dissipativity arguments (Fey *et al.*, 2009; Fey and Bullinger, 2009). Advantages of this observer based approach are that even time varying parameters can be estimated and that an accurate estimate can be guaranteed by mathematical proof. The disadvantage is that identifiability and observability conditions have to be satisfied. In practice, it is difficult to satisfy these conditions globally without measuring many states and even reaction rates.

Here we present a methodology that allows us to make guaranteed statements about the parameters values, even in situations where identifiability is violated. In line with the spirit of the above mentioned methods, the here presented approach exploits biology specific properties, such as polynomial and/or rational kinetics. The polynomial or rational form allow us to use a sum of squares decomposition to represent the differential equations. The consistency of the differential equations and state measurements is then checked using semi-definite programming. The approach is conceptually based on earlier steady state approaches (Kuepfer *et al.*, 2007) but is extended here to transients, i.e. to include system dynamics. Related

approaches in the literature coping with system dynamics use so called barrier certificates, which are conceptually similar to Lyapunov level functions (Anderson and Papachristodoulou, 2009; El-Samad *et al.*, 2006). The drawback of these barrier based approaches is that the construction of the certifying functions is computationally expensive. Here we simply avoid the construction of the barrier certificates by checking the differential equations pointwise in time.

The present article is organised as follows. Section 2 introduces to dynamic modelling in systems biology and formalises the parameter identification problem. Section 3 presents the methodology in the special case of steady state and polynomial kinetics for clarity reasons. Section 4 extends the methodology to include system dynamics, i.e. use of transient data, and Section 5 further extents to rational kinetics. Finally Section 7 illustrates the methodology on two simple examples.

## 2. BIOLOGICAL MODELLING AND PARAMETER IDENTIFICATION

A common framework for the modelling of biochemical reaction networks are sets of reactions of the following form

$$\alpha_1 \mathrm{S}_1 + \ldots + \alpha_{n_s} \mathrm{S}_{n_s} \rightarrow \beta_1 \mathrm{P}_1 + \ldots + \beta_{n_p} \mathrm{P}_{n_p}, \qquad (1)$$

where $\mathrm{S}_i$ denote substrates that are transformed into the products $\mathrm{P}_i$. The factors $\alpha_i$ and $\beta_i$ denote the stoichiometric coefficients of the reactants. Neglecting spatial and stochastic effects, these reactions are often modelled with systems of ordinary differential equations:

$$\dot{c} = N v(c, p), \qquad (2)$$

where $c \in \mathbb{R}^{n_c}_{\geq 0}$ is the vector of concentrations, $p \in \mathbb{R}^{m}_{>0}$ the parameter vector and $v \in \mathbb{R}^{n_c}_{\geq 0} \times \mathbb{R}^{m}_{>0} \mapsto \mathbb{R}^{n_v}_{\geq 0}$ the vector of the flows. The stoichiometric matrix $N \in \mathbb{R}^{n_c \times n_v}$ depends on the coefficients $\alpha_i$, $\beta_i$ and, possibly on factors compensating different units or volumina. For a more detailed introduction, see for example Klipp *et al.* (2005) or Keener and Sneyd (2001).

In systems biology, the stoichiometry and the reaction rate models are often assumed to be known. Typically the reaction rate models $v(c, p)$ are special polynomial or rational functions such as (generalised) mass action kinetics, Michaelis Menten kinetics, Hill kinetics or enzyme kinetics with inhibition or activation terms (Cornish-Bowden, 2004). In contrast to the form of the kinetics, the specific parameter values $p$ are largely unknown. Thus, the problem can be formulated as follows:

**Given:** The stoichiometric matrix $N$ and the form of the function $v(c, p)$ describing the reaction rates
**Unknown:** The kinetic parameters $p$

These unknown parameters need to be estimated from measurements $c(t_i)$ at certain time points. In principle, the parameter estimation problem can be solved by integrating (2). The difficulty with the integration based approach is twofold. On the one hand, an analytic solution of the integral $\int N v(c, p) dt$ can only be computed in rare special cases, which are generally not relevant in biology. On the other hand, numerical integration requires to specify the unknown parameters $p$ (and initial conditions $c(t = 0)$) beforehand. Most parameter estimation methods deal with the dilemma of specifying the unknown parameters by choosing parameter values a priori, numerically integrating the differential equations and then comparing the result with the data a posteriori(Kuepfer *et al.*, 2007). Advanced methods run a loop in which the parameters are updated after each integration step, for example using an evolutionary strategy or a gradient based approach (Peifer and Timmer, 2007; Moles *et al.*, 2003). But even iterative methods do not resolve the fact that only a single parameter set can be considered at a time. Consequently it is impossible to conclude that the best parameter values were found, or that no good solution exists, i.e. that the considered model is inconsistent with the collected experimental data. In fact, the procedure might simply fail to obtain a good parameter estimate.

The above mentioned drawbacks arise from a point-wise checking of the parameter space, which is highly inefficient for a large number of parameters. Here we circumvent this problem by presenting a methodology that is capable of checking entire parameter regions. Instead of providing a parameter estimate, the proposed methodology provides certificates for different parameter regions by proving their inconsistency with experimental data, thus reducing the entire search space to a comparable small fraction containing good solutions (and the true parameters).

The method is based on a polynomial representation of the system dynamics in terms of sum of squares (Parillo, 2003). Such polynomial representation is not restricted to polynomial kinetics, but also possible for general rational kinetics as they for instance appear in metabolic pathway modelling, as will be shown in Section 5. Based on the sum of squares representation, a relaxed semi-definite program is formulated (Boyd and Vandenberghe, 2004). As a consequence of the relaxation it can not be proven that a parameter is consistent with the data, because a solution of the relaxed problem is not necessarily a solution of the original problem. We can however prove inconsistency of entire parameter regions: If the semi-definite program is *infeasible* for a certain parameter region, then this region does *not* contain parameter values consistent with the data. Checking feasibility of the semi-definite program can be done efficiently using high quality software such as SeDuMi and YALMIP, which also provide a high level programming language for implementing the problem (Sturm, 1999; Löfberg, 2004).

## 3. STEADY STATE ANALYSIS FOR SYSTEMS WITH POLYNOMIAL KINETICS

The following considerations address the problem of whether certain parameter regions consistent with measurements of the species concentrations in steady state (Kuepfer *et al.*, 2007). This section is restricted to systems with polynomial kinetics in steady state in order to facilitate the communication of the main ideas. Afterwards the methodology will be extended in the later sections.

The problem can be formulated as follows

$$\begin{aligned} \mathbf{P1}: \quad & \text{find} && c, p \\ & \text{s.t.} && N v(c, p) = 0 \\ & && \bar{c}_i - \epsilon_i \leq c_i \leq \bar{c}_i + \epsilon_i \quad i = 1, \ldots, m \\ & && p_{j,\min} \leq p_j \leq p_{j,\max} \quad j = 1, \ldots, m \end{aligned}$$

where $\bar{c}_i$ is the measured concentration of species $i$, $\epsilon_i$ the corresponding measurement uncertainty and $p_{j,\min}$, $p_{j,\max}$ lower and upper bounds on the parameter defining the parameter region to analyse.

### *3.1 Expressing the righthand-side of the ordinary differential equations as sum of squares*

Often, the reaction rates are linear (affine) in the parameters and polynomial of degree $d$ in the concentrations. i.e. $v(c,p) \in \mathbb{R}[c,p]^{d,1}$. Let

$$\xi^T = \begin{bmatrix} 1 & p_1 & \dots & c_1 & \dots & c_1c_2 & \dots & c_n^d \end{bmatrix} \tag{3}$$

be the vector of monomials constituting a basis for $\mathbb{R}[c,p]^{d,1}$, then the reaction rate can be written as a sum of squares (SOS)

$$v_j(c,p) = \xi V_j \xi^T, \qquad j = 1,\dots,m,$$

where $V_j$ is a symmetric matrix. Therewith a SOS-representation for the righthand-side of the ordinary differential equations (2) is given by

$$\sum_j^m N_{i,j} v_j(c,p) = \xi R_i \xi^T, \qquad i = 1,\dots,n,$$

where the symmetric matrices $R_i$ are given by

$$R_i = \sum N_{ij} V_j, \qquad i = 1,\dots,n. \tag{4}$$

Thus the inequality constraint in Problem $P1$ can be expressed in terms of the monomial basis vector $\xi$ as follows:

$$\xi^T R_i \xi = 0 \qquad i = 1,\dots,n, \tag{5}$$

where the symmetric matrices $R_i$ are constructed according to Eq. (4) and $\xi$ a monomial basis vector as defined in (3).

### *3.2 Measurements and parameter regions as constraints*

Let $\mu(p,c)$ be the mapping of the parameters $p$ and states $c$ to basic vector of monomials, i.e.

$$\xi = \begin{bmatrix} 1 \\ \mu(p,c) \end{bmatrix}.$$

By construction, $\mu$ is monotone and therefore assumes its minimum and maximum for $(p_{\min}, c_{\min})$ and $(p_{\max}, c_{\max})$ respectively. The inequalities $c_{i,\min} \le c_i \le c_{i,\max}$ are thus covered by

$$B\xi \ge 0, \qquad \text{with } B = \begin{bmatrix} \vdots & \vdots \\ -\mu_j(p_{\max}, c_{\max}) & e_j^T \\ \mu_j(p_{\min}, c_{\min}) & -e_j^T \\ \vdots & \vdots \end{bmatrix},$$

where $e_j$ is the is the unit vector with $k$ the only nonzero entry, i.e.

$$e_{j,k} = \begin{cases} 1 & \text{if } k = j \\ 0 & \text{otherwise} \end{cases}.$$

### *3.3 Relaxation to a semi-definite program*

A relaxation of the original problem is now found by defining $X = \xi\xi^T$ (Parillo, 2003). The resulting non-convex constraint $\operatorname{rank}(X) = 1$ is dropped in the semi-definite program. Several other convex constraints arising from he definition of $X$ such as $X_{11} = 1$ and $X \succeq 0$ are still used.

$$\begin{aligned} \textbf{RP1} \quad \text{find} \quad & X \in \mathbb{S} \\ \text{s.t.} \quad & \operatorname{tr}(Q_i X) = 0 \qquad i = 1,\dots,n \\ & \operatorname{tr}(e_1 e_1^T X) = 1 \\ & BXe_1 \ge 0 \\ & BXB^T \ge 0 \\ & BXe_1 \succeq 0, \end{aligned}$$

where $e_1 = [1,\ 0,\ \dots]$.

## 4. EXTENSION TO TRANSIENTS

Kuepfer *et al.* (2007) considered steady state analysis, i.e. $\dot{c} = 0$. This section extends the methodology to include dynamics, i.e. $\dot{c} \neq 0$. The approach requires the knowledge of the time derivative of the states $\dot{c}$. In principle this can be achieved by measuring $c(t)$ at least two proximate time points and calculating $\dot{c} = \frac{c(t_2)-c(t_2)}{t_2-t_1}$. In practice, curve-fitting techniques in which basic functions $\psi_i(t)$ with well known derivatives $\dot{\psi}_i(t)$ are fitted to the measurements, are preferable because of better accuracy and measurement noise reduction.

The problem can be formulated as follows

$$\begin{aligned} \textbf{P2}: \quad \text{find} \quad & c, p \\ \text{s.t.} \quad & Nv(c,p) - \dot{c} = 0 \\ & c_{i,\min} \le c_i \le c_{i,\max} \quad i = 1,\dots,n \\ & p_{j,\min} \le p_j \le p_{j,\max} \quad j = 1,\dots,m \\ & \dot{c}_{i,\min} \le c_i \le \dot{c}_{i,\max} \quad i = 1,\dots,n \end{aligned}$$

where $c_{i,\min}$, $c_{i,\max}$, $\dot{c}_{i,\min}$, $\dot{c}_{i,\max}$ are lower and upper bounds on the measurements of the species concentrations and their derivatives respectively, and $p_{j,\min}$, $p_{j,\max}$ are lower and upper bounds defining the parameter region which is to analyse. Similarly to Section 3 this section finds a relaxation of the problem to a semi-definite program using sum of squares.

### *4.1 Representing the ordinary differential equations as sum of squares*

Recall that Section 3 established a sum of squares representation for the righthand-side of the differential equations

$$\dot{c}_i = \xi^T R_i \xi. \tag{6}$$

Using the extended version of the basis vector of monomials

$$\zeta^T = \begin{bmatrix} \xi^T & \dot{c} \end{bmatrix},$$

Eq. (6) is equivalent to

$$\zeta^T Q_i \zeta = 0 \quad \text{with } Q_i = \begin{bmatrix} R_i & S_i \\ S_i^T & 0 \end{bmatrix}, \quad i = 1,\dots,n, \tag{7}$$

and with the matrix $R_i$ as in Eq. (4) and with the matrix $S_i$ of dimension $\dim\xi \times n$ defined elementwise, i.e. the element $k,l$ of $S_i$ is

$$S_{i_{k,l}} = \begin{cases} -1/2 & \text{if } k = 1, l = i \\ 0 & \text{otherwise} \end{cases}.$$

Therewith the equality in Problem $P2$ is expressed by

$$\zeta^T Q_i \zeta = 0, \qquad i = 1,\dots,n. \tag{8}$$

*4.2 Measurements and parameter regions as constraints*

Setting up the constraints is virtually the same as in Section 3. The only difference is that the mapping $\mu$ has to be extended to include $\dot{c}$, i.e.

$$(p, c, \dot{c}) \mapsto \zeta : \qquad \zeta = \begin{bmatrix} 1 \\ \mu(p, c) \\ \dot{c} \end{bmatrix}.$$

The inequalities are the set up similarly to Section 3, with the extended mapping

$$\nu(p, c, \dot{c}) = \begin{bmatrix} \mu(p, c) \\ \dot{c} \end{bmatrix}.$$

leading to the matrix $B$

$$B = \begin{bmatrix} \vdots & \vdots \\ -\nu_j(p_{\max}, c_{\max}, \dot{c}_{\max}) & e_j^T \\ \nu_j(p_{\min}, c_{\min}, \dot{c}_{\min}) & -e_j^T \\ \vdots & \vdots \end{bmatrix}, \tag{9}$$

covering the with $B\zeta \geq 0$ the inequality constraints in Problem $P2$.

*4.3 Relaxation to a semi-definite program*

Based on Eq. (8) and (9) the relaxed problem in the transient case is the following semi-definite program

**RP2** find $Z \in \mathbb{S}$

s.t. $\operatorname{tr}(Q_i Z) = 0 \qquad i = 1, \ldots, n$

$\operatorname{tr}(e_1 e_1^T Z) = 1$

$BZe_1 \geq 0$

$BZB^T \geq 0$

$BZe_1 \succeq 0,$

where the $Q_i$ are defined in Eq. (7), $B$ is defined in (9) and $Z \in \mathbb{S}$ is the relaxed version of $\zeta\zeta^T$ corresponding to $X$ in Section 3.

## 5. EXTENSION TO RATIONAL KINETICS

The presented methodology is easily extended to systems with rational kinetics. All that is required for the relaxation to a semi-definite program as presented in the previous sections is a sum of squares representation of the system of ordinary differential equations

$$\dot{c} = Nv(c, p), \qquad \text{with } v_j(c, p) = \frac{r(c, p)}{q(c, p)}, \tag{10}$$

where $r, q \in \mathbb{R}[c, p]$ are polynomial functions.

A polynomial representation of System (10) is easily obtained by multiplying left- and righthand-side of each equation with all denominators of $v(c, p)$, yielding

$$\sum_j N_{i,j} p_j(c(t_l), p) = \prod_{k \neq j} q_k(c(t_l), p) - \dot{c}_i(t_l) \prod_k q_k(c(t_l), p).$$

Defining $z^T = \begin{bmatrix} p^T & c^T & \dot{c}^T \end{bmatrix}$ this can be expressed as

$$0 = f_i(z), \qquad i = 1, \ldots, n$$

where $f \in \mathbb{R}[z]^{2d}$ is a multivariate polynomial of degree $2d < \infty$. Let $\zeta$ be a basis of $\mathbb{R}[z]^d$ consisting of monomials, then there exists a symmetric matrix $Q \in \mathbb{S}^{d \times d}$ such that

$$f_i(z) = \zeta^T Q_i \zeta, \qquad i = 1, \ldots, n.$$

In practice the polynomial representations is often sparse, i.e. the matrices $Q_i$ have common zero-rows and -columns. Then these zero-rows and -columns in $Q_i$ as well as the corresponding monomials in the basis $\zeta$ should be removed (Waldherr *et al.*, 2008). Let $\tilde{Q}_i$ be these reduced matrices and let $\tilde{\zeta}$ be the corresponding reduced basis, then $Q_i$ is to be used in the semi-definite programm $RP2$, and the matrices $B$ have to be constructed according to the monotone mapping $(c, p, \dot{c}) \mapsto \tilde{\zeta}$ as described in Section 4.2.

## 6. RELATION TO BARRIER CERTIFICATES

Similar approaches in the literature use the so called barrier certificates to certify parameter regions inconsistent with data and model (Anderson and Papachristodoulou, 2009). Similar to a Lyapunov function, a barrier certificate $B(x, p, t)$ is a function depending on the states, parameters and time. Given two measurements $x_1, x_2$ at different time points, the idea is to find a real-valued function that for some parameter region $P \ni p$ has a higher value at the second time point compared to the first time point $B(x_2, p, t_2) - B(x_1, p, t_1) > 0$ but at the same time is nonincreasing along the trajectory $\frac{\partial B}{\partial x} f + \frac{\partial B}{\partial t} \leq 0$. Such construction obviously creates a contradiction and it can be concluded that the parameter region $P$ is inconsistent with the measurement.

The advantage of a barrier certificate is that it incorporates the righthandside of the differential equations between the measurement points. Generally this may allow to find tighter bounds on the feasible parameter regions compared to the simple method proposed here (Sections 3 to 5) that only uses the differential equations at the measured time points.

The drawback of a barrier certificate is that it requires the construction of a polynomial function, which is computationally demanding for complicated systems. The number of decision variables required for this construction increases polynomially with the number of states. In contrast, the here presented method (Sections 3 to 5) does not require the construction of a certificate function and the number of decision variables increases only linearly with the states.

## 7. EXAMPLE

The proposed methodology is illustrated using two simple toy examples.

*Simple toy model*

We chose this toy example such that an analytical solution to the parameter estimation problem can be found easily. This allowed us to validate the proposed approach by comparing to the analytical solution.

Consider the following reaction

$$S + E \rightleftharpoons C + E.$$

Using the law of mass action the concentrations are described by the following system of ordinary differential equations

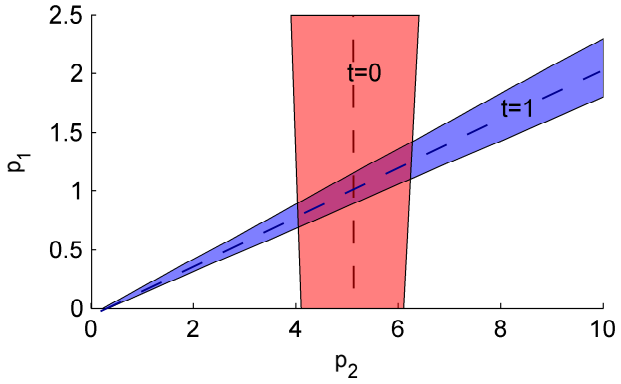



Fig. 1. Consistent parameter regions for measurements at two different points in time. Red: for the transient at t=0. Blue: in steady state (t=1). The dashed lines indicate the parameter dependencies for perfect measurements. The true parameter value that generated the data is the intersection of both dashed lines. The cones represent the admissible parameter values for 10% measurement uncertainty. Only the intersection is consistent with both measurements.

$$\frac{d}{dt}\begin{bmatrix} c \\ e \\ s \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 0 \\ -1 & 1 \end{bmatrix} \mathbf{v}(\mathbf{c}, p), \qquad \begin{aligned} v_1 &= k_1 ce \\ v_2 &= p_2 c. \end{aligned}$$

Using that $e$ is constant ($\dot{e} = 0$) we can set $p_1 = k_1 e$ as a constant parameter. Further using the conservation law $s(t) + c(t) = S_0$, the system is described by a single ordinary differential equation for the concentration of $C$

$$\dot{c} = [1 \ -1]\, \mathbf{v}(c, p), \qquad \begin{aligned} v_1 &= p_1 S_0 - p_1 c \\ v_2 &= p_2 c \end{aligned}.$$

A monomial basis vector for this system is

$$\zeta = [1 \ p_1 \ p_2 \ S_0 \ c \ \dot{c}]\,,$$

which allows the application of the method as described in Section 4. However, due to the simplicity of the system we can also analyse the problem analytically as follows. By setting $\dot{c} = 0$ we can see that in steady state the parameter $p_1$ depends linearly on the parameter $p_2$ with $p_1 = \frac{\bar{c}}{S_0 - \bar{c}} p_2$ (see Figure 1). In the case of perfect measurements, this dependency is a one dimensional line in the parameter space. In the case of uncertain measurements (i.e. upper and lower bounds on the measured concentration), this dependency results in a two dimensional cone in the parameter space (see Figure 1 for an example with 10% measurement accuracy). In the dynamic case, the concentration changes over time depending on the initial condition $c(t = 0) \neq \bar{c}$. When $\dot{c}$ is measured or estimated from measurement of $c$ at several time points $t_i$, the cone of admissible parameter regions changes over time. Figure 1 illustrates this time dependency for $t = 0$ and $t = 1$. Since the parameters have to be consistent at all time points, only the intersection contains admissible parameter values, and in fact also the true parameter values.

*Lotka-Volterra model*

The famous LotkaVolterra equations describe the population dynamics of two interacting species, with one being the prey, the other being a predator:

$$\begin{aligned} \dot{c}_1 &= c_1(p_1 - c_2), \\ \dot{c}_2 &= c_2(-p_2 + c_1), \end{aligned}$$

where $c_1$ and $c_2$ are the concentration (e.g. number of animals per habitat) of prey and predator respectively, and $p_1$ and $p_2$ are parameters describing the birth rate of the prey and the mortality rate of the predator (Lotka, 1925; Volterra, 1926).

The model is used to illustrate that the proposed methodology is applicable to systems with periodic (non-steady-state) trajectories. The model was chosen because it is well studied and results obtained by the proposed methodology can thus be verified easily. The semi-definite program was constructed according to Section 4 with the monomial basis vector

$$\zeta = [1 \ p_1 \ p_2 \ S_0 \ c_1 \ c_2 \ c_1 c_2 \ \dot{c}_1 \ \dot{c}_2]\,.$$

Feasibility checking on a grid of parameter regions yielded the following results.

- Measuring $c_1, c_2, \dot{c}_1, \dot{c}_2$ resulted in small oval-shaped parameter regions. For measurements at different time points, these regions were slightly different, such that the admissible parameter region could be reduced slightly by taking the intersection.
- Not measuring $\dot{c}_1$ or $\dot{c}_2$ rendered $p_1$ or $p_2$ respectively unidentifiable. The admissible parameter region was a narrow rectangle parallel to the unidentifiable parameter axis.
- Not measuring $c_1$ yielded nonlinearly curved admissible parameter regions that changed dramatically over time (Figure 2). Taking measurements at different time points reduced the region of admissible parameter values dramatically by taking the intersection.

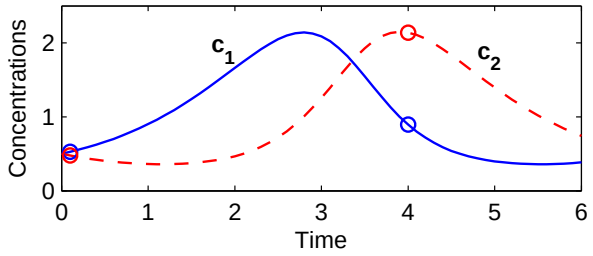
The two examples nicely illustrate that the methodology is applicable to measurements of transients, and that combining measurements at two or more time points significantly reduces the region of admissible parameters.

## 8. CONCLUSIONS

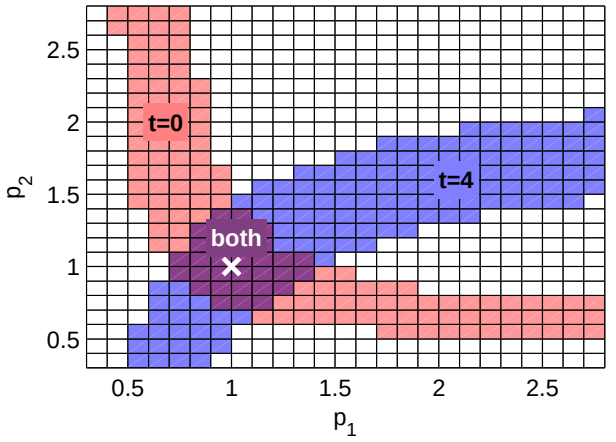
We presented a methodology that is capable of proving inconsistency of entire parameter regions with the data. Based on the polynomial representation of the system, we formulate a feasibility problem that can be solved efficiently by semi-definite programming. The feasibility problem allows us to check consistency of entire parameter regions by using different upper and lower bounds on the parameters. This drastically limits the search space for subsequent parameter estimation methods. In contrast to similar approaches in the literature, the here presented approach does not require a steady state assumption (Kuepfer *et al.*, 2007), nor a discretisation of the system (Borchers *et al.*, 2009). In fact, the presented examples show that using dynamic information by measuring the transient at different time points significantly reduce the regions of admissible parameter regions. Further, measurement uncertainties can be included in the feasibility problem naturally by using upper and lower bounds and regular sampling times are not required.

## ACKNOWLEDGEMENTS

This paper presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Opti-




(a) Trajectories that generated the in-silico measurement data. Circles indicate the time points where the feasibility problem was checked.




(b) Consistent parameter regions for measurements at two different points in time. Only the intersection is consistent with both measurements. Results were obtained with 10% measurement uncertainty and $c_1$ was not measured.

Fig. 2. Simulation results of the Lotka-Volterra example.

mization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office. The scientific responsibility rests with its authors

## REFERENCES


Anderson, J. and Papachristodoulou, A. (2009). On validation and invalidation of biological models.. *BMC Bioinformatics* **10**, 132.

Borchers, S., P. Rumschinski, S. Bosio, R. Weismantel and Findeisen., R. (2009). Model invalidation and system identification of biochemical reaction networks. In: *15th IFAC Symposium on System Identification, Saint Malo, France*.

Boyd, S. and Vandenberghe, L. (2004). *Convex optimization*. Cambridge University Press.

Bullinger, E., R. Findeisen, D. Kalamatianos and Wellstead, P. (2007). System and control theory furthers the understanding of biological signal transduction. In: *Biology and Control Theory: Current Challenges*. pp. 123–135. Number 357 In: *Lecture Notes in Control and Information Sciences*. Springer Berlin / Heidelberg.

Cornish-Bowden, A. (2004). *Fundamentals of Enzyme Kinetics*. third ed.. Portland Press.

El-Samad, H., S. Prajna, A. Papachristodoulou, J. Doyle and Khammash, M. (2006). Advanced methods and algorithms for biological networks analysis. *Proceedings of the IEEE* **94**(4), 832–853.

Farina, M., R. Findeisen, E. Bullinger, S. Bittanti, F. Allgöwer and Wellstead, P. (2006). Results towards identifiability properties of biochemical reaction networks. In: *45th IEEE Conference on Decision and Control, San Diego, USA*. IEEE. pp. 2104–2109.

Fey, D., R. Findeisen and Bullinger, E. (2008). Parameter estimation in kinetic reaction models using nonlinear observers is facilitated by model extensions. In: *17th IFAC World Congress, Seoul, Korea*. pp. 313–318.

Fey, D. and Bullinger, E. (2009). A dissipative approach to the identification of biochemical reaction networks. In: *15th IFAC Symposium on System Identification, Saint Malo, France*. pp. 1259–1264.

Fey, D., R. Findeisen and Bullinger, E. (2009). Identification of biochemical reaction networks using a parameter-free coordinate system. In: *Control Theory and Systems Biology* (P. A. Iglesias and B. Ingalls, Eds.). pp. 297–316. MIT press.

Hodgkin, A. and Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol* **117**(4), 500–544.

Keener, J. and Sneyd, J. (2001). *Mathematical Physiology*. Vol. 8 of *Interdisciplinary Applied Mathematics*. second ed.. Springer-Verlag. New York.

Klipp, E., R. Herwig, A. Kowald, C. Wierling and Lehrach, H. (2005). *Systems Biology in Practice: Concepts, Implementation and Application.*. Wiley-VCH. Weinheim.

Kuepfer, L., U. Sauer and Parrilo, P. A. (2007). Efficient classification of complete parameter regions based on semidefinite programming.. *BMC Bioinformatics* **8**, 12.

Löfberg, J. (2004). Yalmip : A toolbox for modeling and optimization in MATLAB. In: *Proceedings of the CACSD Conference*. Taipei, Taiwan.

Lotka, A. J. (1925). *Elements of physical biology*. Baltimore: Williams & Wilkins Co.

Moles, C. G., P. Mendes and Banga, J. R. (2003). Parameter estimation in biochemical pathways: a comparison of global optimization methods.. *Genome Res* **13**(11), 2467–2474.

Parillo, P. A. (2003). Semidefinite programming relaxations for semialgebraic problems. *Math Prog* **96**(2, Ser B), 293–320.

Peifer, M. and Timmer, J. (2007). Parameter estimation in ordinary differential equations for biochemical processes using the method of multiple shooting. *IET Systems Biology* **1**(2), 78–88.

Sontag, E. (2005). Molecular systems biology and control. *European J Control* **11**(4–5), 396–435.

Sturm, J. F. (1999). Using sedumi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software* **11–12**, 625–653.

Volterra, V. (1926). Variazioni e fluttuazioni del numero d'individui in specie animali conviventi.. *Mem. R. Accad. Naz. dei Lincei* **2**, Ser VI.

Waldherr, S., R. Findeisen and Allgöwer, F. (2008). Global sensitivity analysis of biochemical reaction networks via semidefinite programming. In: *17th IFAC World Congress, Seoul, Korea*. pp. 9701–9706.

Wellstead, P., E. Bullinger, D. Kalamatianos, O. Mason and Verwoerd, M. (2008). The role of control and system theory in systems biology. *Annual Reviews in Control* **32**(1), 33–47.