Preprints, 10th IFAC International Symposium on Advanced Control of Chemical Processes Shenyang, Liaoning, China, July 25-27, 2018

Parameter Estimation for Signal Transduction Networks from Experimental Time Series Using Picard Iteration

Friedrich v. Haeseler * **Nadine Rudolph** * **Rolf Findeisen** * **Heinrich J. Huber** *

* *Laboratory for Systems Theory and Automatic Control Otto-von-Guericke-University Magdeburg, Germany*

Abstract: Biological signal transduction models allow to explain and analyze biological cause-effect relationships and to establish and test new hypotheses about biological pathways. Yet their predictive capability crucially depends on the parameters involved. These parameters are usually determined from experimental data. However, due to the appearing nonlinearities, the resulting inverse problem is often ill-posed and difficult to solve. We outline how parameters can be estimated based on Picard iterations. In case of linear parameter dependence and good measurements of the involved entities, the method allows to retrieve good parameter estimates for medium size problems. The proposed method is applied to an IL-6-dependent Jak-STAT3 signalling pathway model. As shown it it is well suited for data generated by life cell imaging where accurate time series are available.

1. INTRODUCTION

Cellular signal transduction pathways are biochemical processes whereby initial biological signals such as receptor stimulation proceed via a cascade of biochemical interactions (between proteins or genes) over time. Eventually they lead to the initiation of a cellular response such as cell growth, differentiation or cell death. Having to integrate several cellular signals, balancing them against the cell viability and being shaped by evolution, these pathways are complex in nature. Furthermore, they contain several regulatory elements such as multiple feedbacks to determine the appropriate cell fate (Milo et al., 2002; López-Caamal et al., 2014). To understand the complexity of such 'cellular integrated circuits', mathematical models, often in form of Ordinary Differential Equation (ODE)-based, have been widely used. These models crucially depend upon several parameters describing biochemical properties such as the speed and equilibrium of biochemical reactions. While being of biochemical nature, the exact values of these parameters often cannot be directly measured. Instead, they have to be inferred indirectly by fitting the model output to experimental time series data.

Many approaches for estimating parameters from measurement data exist, see e.g. (Schaber and Klipp, 2011; Chou and Voit, 2009; Moles et al., 2003; Chou and Voit, 2009; Voit, 2013; Schliemann-Bullinger et al., 2016; Rumschinski et al., 2010). Often parameters are fitted by numerical optimisation (such as gradient descent) to minimise the difference between model predictions and experimental training data. However, such approaches can suffer from the presence of local minima that lead to different solutions than the optimal one, resulting in sub-optimal or spurios parametrisations. Furthermore, they do not guarantee that a unique solution, or a valid solution at all, can be identified. Therefore, several approaches have been performed that, besides providing means for rigorous parameter estimation, investigate structural identifiability. The latter term refers to the theoretical property whether or not a model including its parametrisation can be fully regenerated from the model input, the topology of the dynamical system and the experimental data (Chis et al., 2011; Chou and Voit, 2009). These general methods comprise approaches based on Generalized Mass Action, S-system based methods (Savageau, 1969a,b; Torres and Voit, 2002), and such from control theory (Geffen et al., 2008; Farina et al., 2006; Lillacci and Khammash, 2010; Otter, 1986).

Besides these more fundamental approaches, several methods to estimate parameters values from given experimental data have been developed. Some approaches use direct integration of the underlying ODE system (Docherty et al., 2012), similar as in other fields of parameter estimation inverse problem theory (Huber and Leeb, 1998). As such example, incremental identification estimation methods, whereby parameter subsets are optimised sequentially (Bhatt et al., 2012), have been developed such as the integrated flux parameter estimation (IFPE) method (Liu and Gunawan, 2014). Moreover, the application of an iterative Picard formalism has been suggested (Kunze and Vrscay, 1999; Tanner, 1972) to identify parameters for determining enzyme kinetics and estimating parameters in bioprocessing. We provide an elaboration of these Picard iteration-based approaches, and show their first application in a realistic signal transduction context.

The contribution is structured as follows: Section 2 describes the Picard iteration formalism to determine the kinetic parameters. The biological example, an IL-6-dependent Jak-STAT3 signalling pathway is introduced

[1] NR and RF are members of the International Max Planck Research School (IMPRS) for Advanced Methods in Process and System Engineering, Magdeburg.

[2] This work is supported by the BMBF in the frame of the research project InTraSig, grant number 031A300A, and the Center for Dynamical Systems in the frame of the IMPRS.

[3] Authors for correspondence. Email: Heinrich.Huber@ovgu.de, Rolf.Findeisen@ovgu.de

in Section 3. The application of the derived approach to the example is summarized in Section 4 and concluded in Section 5 with an outlook.

2. PROBLEM SET-UP AND SOLUTION METHOD

2.1 Modeling Biological Systems

Cellular signal transduction is the process, whereby an initial cellular signal such a receptor stimulus is transferred by a set of biochemical reaction into a certain cellular response such as cell growth, differentiation and death. To analyse these often intricate processes over time, computational models, based on ODEs are often used. Thereby, the changes of concentrations of biochemical entities, such as proteins, nucleic acids, lipids or metabolites are typically modelled by different state variables (denoted as 'species') $X_1, ..., X_{n_x}$,

$$x = ([x_1], ..., [x_n])^T \in \mathbb{R}^n \tag{1}$$

comprising the concentration of the n species $[x_i]$ within the cell and their variation in time. We further assume concentrations of the model species to be sufficiently large and the cellular content is well-mixed such that spatial gradients and stochastic effects can be neglected. The species are converted by chemical reactions following

$$\alpha_{1j}X_1 + \alpha_{2j}X_2 + ... + \alpha_{nj}X_n \underset{r_j^+}{\overset{r_j^-}{\rightleftharpoons}} \tag{2}$$

$$\beta_{1j}X_1 + \beta_{2j}X_2 + ... + \beta_{nj}X_n. \tag{3}$$

Thereby, α_{ij} and β_{ij} for $i \in \{1, ..., n\}$ are non-negative stoichiometric coefficients that describe the amount of entities i within a reaction $j, \in \{1, ..., n_r\}$. Furthermore, r_j^+ denotes the forward and r_j^- the backward reaction rate for reaction j. For modelling reaction rates, the law of mass action, where the reaction rate is assumed to be proportional to the substrate concentrations, is the simplest and most widely-used assumption (Savageau, 1969a,b; Torres and Voit, 2002), i. e.

$$r_j(t) = p_j^+ \prod_{i=1}^{n_x} x_i^{\alpha_{ij}}(t) - p_j^- \prod_{i=1}^{n_x} x_i^{\beta_{ij}}(t). \tag{4}$$

In Eq. (4), p_j^+ and p_j^- denote the forward and backward reaction constants, whose determination from the state variables over time (1) is the subject of this paper.

With this notation, the biochemical reaction network can be written in a compact form as

$$\dot{x}(t) = Nr(t), \tag{5}$$

where $\dot{x}(t)$ denotes the change of concentration with time. Furthermore, all reaction rates are collected into the vector $r(t) = (r_1(t), ..., r_{n_r}(t))$ and $N \in \mathbb{R}^{n \times n_r}$ denotes the stoichimetric matrix given by

$$N_{ij} = \beta_{ij} - \alpha_{ij}, \quad i \in \{1, ..., n\}, j \in \{1, ..., n_r\}. \tag{6}$$

The overall dynamics for a system without time-variant input and defining a nonlinear vector function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, we have

$$\dot{x}(t) = Nr(t) = f(x(t), p). \tag{7}$$

In Eq. (7), $x(t) \in \mathbb{R}^n$, and $p \in \mathbb{R}^m$ denotes the time-invariant parameter vector collecting all rate constants p_j^+ and p_j^-. Throughout this work, we assume that $x(t)$ belongs to the set $\mathcal{X} \subseteq \mathbb{R}^n$, and p to $\mathcal{P} \subseteq \mathbb{R}^m$.

In general, not all model species can be measured. Therefore, we distinguish between the model states $x(t)$ and model outputs $y(t) \in \mathbb{R}^q$. For simplification, all outputs are aggregated into the nonlinear vector function g, i. e.

$$y(t) = g(x(t), p) \tag{8}$$

where $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^q$.

Summarizing Eq. (7) and Eq. (8), one obtains the system

$$\begin{cases} \dot{x}(t) = f(x(t), p) \\ y(t) = g(x(t), p), \end{cases} \tag{9}$$

which describes the time behavior of the biochemical reaction network depending on the parameters p.

2.2 Determining Parameters based on Picard Iterations

We provide a method to obtain the vector of time invariant parameters p from the model $f(x(t), p)$. For simplicity of presentation, we assume that we know all state values x_k at particular time points k, $k = 1 \dots x_k$. Generalizations are possible, but beyond the scope of this paper.

If $\phi : \mathbb{R} \to \mathbb{R}^n$ is a solution of (9) with initial condition $\phi(0) = x_0$, then ϕ satisfies

$$\phi(t) = x_0 + \int_0^t f(\phi(s), p)\, ds\,. \tag{10}$$

The Picard iteration scheme provides a recursive way to obtain from an approximate solution ψ_0 of ϕ a better, at least local, approximation. To this end, one computes a sequence of approximations defined as

$$\psi_{n+1}(t) = x_0 + \int_0^t f(\psi_n(s), p) ds\,. \tag{11}$$

The sequence of approximations converges towards the solution ϕ on the intervall $[0, T]$. For the technical details, please refer to (Arnold, 1992), chapter 4.

On the other hand, knowing the solution $\phi : \mathbb{R} \to \mathbb{R}^n$ with $\phi(0) = x_0$ enables us to compute the parameters $p \in \mathbb{R}^m$ of the differential equation. To achieve this, we assume that we have k precise values of the states $x_j = \phi(t_j)$ at times t_j, $j = 1, \dots, k$, where k is defined such that $k \times n \geq m$. According to the integral equation we obtain $k \times n$ equations,

$$x_j = x_0 + \int_0^{t_j} f(\phi(s), p)\, ds.$$

These $k \times n$ equations allow, the computation of the unknown parameters $p \in \mathbb{R}^m$, provided that we have chosen the measuring points such that the system of nonlinear equations is solvable for p.

The key idea is now to exploit the method of approximation for parameter determination. We therefore assume that we know the complete states x_0 and x_j at time t_j, $j = 1, \dots, k$. If the time steps t_j are assumed to be sufficiently close to each other, it is the reasonable to posit that an interpolating function, i.e., a function $\varphi : \mathbb{R} \to \mathbb{R}^n$ with

$$\varphi(t_j) = x_j$$

for $j = 0, \dots, k$ provides an approximation of the solution ϕ. Therefore, φ satisfies

$$x_j \approx x_0 + \int_0^{t_j} f(\varphi(s), p)\, ds \tag{12}$$

for $j = 1, \ldots, k$. We then aim to determine parameters $\tilde{p}$ such that (12) becomes an equality. Thus, we obtain

$$\varphi^+(t) = x_0 + \int_0^t f(\varphi(s), \tilde{p})\, ds \tag{13}$$

passing at time t_j through the point x_j. Here, φ^+ is an interpolating function, which under proper circumstances is a better approximation to ϕ than the original initial interpolating function φ. Repeating this procedure now with φ^+ is expected to render better approximations of ϕ and, eventually, the parameters p.

2.3 Linear Parameter Dependency

The formalism of the previous section can be simplified if mass-action kinetics is used. Then, the model $\dot{x} = f(x, p)$ depends linearly on the parameters and in this case the equations (12) become linear equations for the unknown parameters p.

$$\dot{x} = A(x)p \tag{14}$$

where $A : \mathbb{R}^n \to \mathbb{R}^{n\times m}$, $x \mapsto A(x)$ is an $n \times m$-matrix and we get the linear equation system,

$$x_1 = x_0 + \int_0^{t_1} A\left(\varphi(s)\right)\, ds\, p. \tag{15}$$

Since more complex kinetics such as Michaelis-Menten or Hill kinetics are approximations of multi-step mass action reaction kinetics, this assumption does not impede the general applicability of our method to biochemical signal transduction analysis.

2.4 Simple Example

This example illustrates the underlying ideas of the recursive parameter determination, using the simplest case of employing linear interpolation between two time points.

We consider the differential equation

$$\dot{x} = p\, x,$$

where $p \in \mathbb{R}$ has to be determined.

Let $\phi : \mathbb{R} \to \mathbb{R}$ be the solution with $\phi(0) = 1$ and $\phi(1) = 2$. Then $\varphi(t) = t+1$ is an approximation of ϕ and interpolates ϕ at times 0 and 1. Thus, we have the equation

$$2 = 1 + \int_0^1 \tilde{p}\, \varphi(s)\, ds \tag{16}$$

yielding $\tilde{p} = \frac{2}{3}$. Using this parameter and the above interpolation for $\varphi(t)$ and setting it into equation (16) gives for the next approximation

$$\varphi^+(t) = 1 + \int_0^t \frac{2}{3}(1+s)\, ds = 1 + \frac{2}{3}t + \frac{1}{3}t^2,$$

with $\varphi^+(0) = 1$ and $\varphi^+(1) = 2$.

In turn, when φ^+ is substituted by φ in Equation (16), we can calculate the next iteration for the parameter $\tilde{p}$. We thereby obtain $\frac{9}{13} = 0.692308$ which is close to the true value $\log(2) = 0.693147$ that is derived from exact solution of above example with the mentioned restrictions for the values at $t = 1$ and 2. Continuing with this iteration scheme leads to a convergent sequence of interpolating functions with limit $\phi(t) = 2^t$ together with a convergent sequence of parameters with limit $\log(2)$.

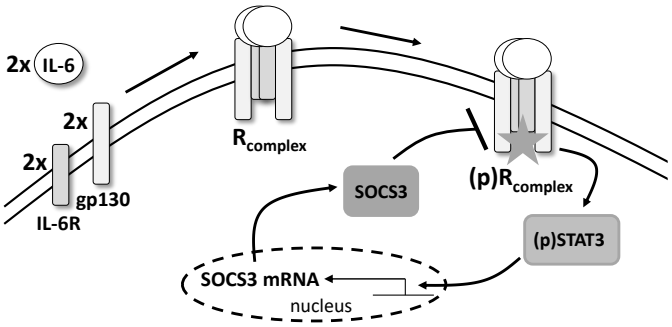

Fig. 1. Schematic representation of IL-6-dependent receptor complex assembly, Jak-STAT3 pathway activation and activation of target genes.

3. BIOLOGICAL EXAMPLE

To illustrate the Picard iteration approach, we consider in the following the activation of the Jak-STAT3 pathway, a major pathway in cellular signalling. We thereby model IL-6-dependent receptor complex assembly and signalling.

3.1 IL-6-dependent Jak-STAT3 Signalling

IL-6 is known as a key regulator of inflammatory processes. Dysregulation of IL-6 function leads to numerous pathologies such as Rheumatoid Arthritis, Crohn's disease and Multiple Sclerosis (Scheller et al., 2006). Computational analysis of inflammatory signalling has proven to be useful for studying underlying disease mechanisms (Dittrich et al., 2012; Veltman et al., 2017).

The mechanisms by which IL-6 initiates signal transduction are given in Fig. 1 and have been reviewed in (Heinrich et al., 2003). Briefly, IL-6 first forms a complex with the receptor subunit glycoprotein 80 (IL-6R). Two of such complexes dimerise and bind to another two adaptor proteins of type glycoprotein 130 (gp130), resulting in a hexameric receptor complex ($R_{complex}$). Next, tyrosine kinases of the Jak protein family, which are constitutively bound to the intracellular domain of gp130, become activated after receptor complex assembly and phosphorylate gp130 ((p)$R_{complex}$). Then, Signal Transducer and Activator of Transcription 3 proteins (STAT3) are recruited to the phosphorylated gp130 complex. Thereby, STAT3 proteins are in turn phosphorylated by Jak proteins, leading to the formation of active STAT3 dimers. Phosphorylated STAT3 (pSTAT3) dimers are assumed to translocate into the nucleus and act as transcription factors which induce several target genes. These genes include the negative regulator of the Jak-STAT3 pathway, the Suppressors of Cytokine Signalling 3 (SOCS3) (Endo et al., 1997) (modelled here as mRNAs and proteins), eventually leading to a termination of Jak-STAT3 signalling.

3.2 Modeling Assumptions and Equations

The dynamic processes of IL-6-induced Jak-STAT signalling as well as inhibition via SOCS3 can be modelled by the following set of differential equations:

$$\begin{aligned}
\frac{dx_1}{dt} &= p_1 x_7(t)u - p_2 x_1(t) - 2p_3 x_2(t)^2 x_1(t)^2 + 2p_4 x_8(t)\\
\frac{dx_2}{dt} &= 2p_4 x_8(t) - 2p_3 x_2(t)^2 x_1(t)^2\\
\frac{dx_3}{dt} &= \frac{p_5 x_8(t)}{1 + p_{13} x_6(t)} - p_6 x_3(t)\\
\frac{dx_4}{dt} &= p_7 y_3(t) x_9(t) - p_8 x_4(t)\\
\frac{dx_5}{dt} &= p_9 x_4(t) - p_{10} x_5(t)\\
\frac{dx_6}{dt} &= p_{11} x_5(t) - p_{12} x_6(t).
\end{aligned} \tag{17}$$

In (17), u denotes the input IL-6 which is assumed to be time-invariant, hence considering a constant receptor stimulus. The dynamical variables of the system (17) $x_1(t)$, $x_2(t)$, $x_3(t)$, $x_4(t)$, $x_5(t)$, $x_6(t)$ denote IL-6~IL-6R, gp130, (p)R_{complex}, (p)STAT3, SOCS3 mRNA and SOCS3 protein, respectively. We have further introduced $x_7(t)$, $x_8(t)$ and $x_9(t)$ which denote the biochemical entities IL-6R, R_{complex} and STAT3, respectively. These are dependent variables, which can be obtained from biological conservation laws as follows:

$$\begin{aligned}
x_2^{\text{total}} &= x_2(t) + 2x_3(t) + 2x_8(t) \;:=\; \alpha\\
x_7^{\text{total}} &= x_1(t) + 2x_3(t) + x_7(t) + 2x_8(t) \;:=\; \beta\\
x_4^{\text{total}} &= x_9(t) + x_4(t) \;:=\; \gamma,
\end{aligned} \tag{18}$$

where we have introducted greek letters for simplifying the further notation. Thereby, x_2^{total} is set to 16.8 nM, x_4^{total} to 83 nM, and x_7^{total} to 2.1 nM according to quantitative biochemistry measurements (F. Schaper, Magdeburg, private communication). The constant IL-6 stimulus u is assumed to be 0.17 nM, which is a typical concentration for IL-6 stimulation experiments (Dittrich et al., 2012). The initial conditions were set to $x(0) = (0, 16.8, 0, 0, 0, 0)$, reflecting the fact that, while gp130 is at maximum, all other entities are not available at time point $t = 0$ and formed after IL-6 stimulation. Notably, the (nonzero) initial conditions for the dependent variables $x_7(t)$ (IL-6R), $x_8(t)$ (R_{complex}), $x_9(t)$ (STAT3) can be derived from the conservation laws (18). We finally remark that we have not explicitely modelled Jak kinases as we assumed them to be an integral part of gp130.

3.3 Application of the Picard Iteration

To exemplify our method, we cast our biological model (17) into the compact notation

$$\dot{x} = f_c(x, p)$$

where $f_c : \mathbb{R}^6 \times \mathbb{R}^{13} \to \mathbb{R}^6$ and is defined as

$$f_c(x,p) = \begin{pmatrix}
p_1\, u\,(-\alpha + \beta - x_1 + x_2) + \\
p_4(\alpha - x_2 - 2\,x_3) - \\
p_3\, x_1^2\, x_2^2 - p_2\, x_1\,, \\
p_4\,(\alpha - x_2 - 2x_3) - p_3\, x_1^2\, x_2^2\,, \\
\dfrac{p_5\left(\frac{\alpha - x_2}{2} - x_3\right)}{1 + p_{13} x_6} - p_6\, x_3\,, \\
x_7\, x_3(\gamma - x_4) - p_8\, x_5\,, \\
p_9\, x_4 - p_{10}\, x_5\,, \\
p_{11}\, x_5 - p_{12}\, x_6
\end{pmatrix}, \tag{19}$$

and where we have separated the rows by commata for clarity reasons. As we see, the parameter p_{13}, is included in a rational term denoting the SOCS3 feedback and hence prevents the linear dependency of f_c on its parameters. For providing a straightforward exemplification of our method, we will therefore first set $p_{13} = 0$ and later calculate this parameter by steady state assumptions. With this choice, our model $\dot{x} = f(x, p)$ depends linearly on the parameters, and we can hence write

$$\dot{x} = A(x)p\,, \tag{20}$$

where $A(x)$ is a 6×12-matrix with block structure

$$A(x) = \begin{pmatrix} A_1(x) & 0 \\ 0 & A_2(x) \end{pmatrix},$$

and with $A_1(x)$ and $A_2(x)$ being 3×6-matrices given as

$$A_1(x) = \begin{pmatrix}
u\,(-\alpha + \beta - x_1 + x_2) & -x_1 & -x_1^2\, x_2^2 & & & \\
0 & 0 & -x_1^2\, x_2^2 & & & \\
0 & 0 & 0 & & & \\
\alpha - x_2 - 2\,x_3 & 0 & 0 & & & \\
\alpha - x_2 - 2\,x_3 & 0 & 0 & & & \\
0 & \dfrac{\alpha - x_2}{2} - x_3 & -x_3 & & &
\end{pmatrix} \tag{21}$$

and

$$A_2(x) = \begin{pmatrix}
(\gamma - x_4)\, x_3 & -x_4 & 0 & 0 & 0 & 0 \\
0 & 0 & x_4 & -x_5 & 0 & 0 \\
0 & 0 & 0 & 0 & x_5 & -x_6
\end{pmatrix}. \tag{22}$$

As a consequence, for the determination of the 12 unknown parameters we need the initial state $x_0 \in \mathbb{R}^6$ and two more states x_1 and $x_2 \in \mathbb{R}^6$ at times t_1 and t_2, respectively. Assuming these two states, the Equations (12) become linear equations and read as

$$\begin{aligned}
x_1 \;&=\; x_0 + \int_0^{t_1} A\left(\varphi(s)\right)\, ds\, p\\
&\text{and}\\
x_2 \;&=\; x_0 + \int_0^{t_2} A\left(\varphi(s)\right)\, ds\, p,
\end{aligned} \tag{23}$$

where φ is an interpolating function of our state vector x passing through x_1 and x_2 at t_1 and t_2, respectively. Note that we have 12 equations and 12 unknown parameters.

4. IMPLEMENTATION AND RESULTS

To numerically exemplify the method, we first calculated the direct problem (17) giving us the temporal evolvement of the state variables by assuming a set of biologically reasonable parameters as given in Table 2. This is first performed for the case without SOCS3 negative feedback

Table 1. Outline: Picard iteration method

- We assume the ODE (20) with linear dependency in the parameters p.
- We take a set of k experimental points of all n state variables
- We assume a first approximation of the temporal behaviour of the solution by interpolating with a polynomial function of k^{th} order according the example in subsection 2.4
- We plug the first interpolating function into the Picard integral and integrate the polynomials
- We solve the resulting equation for $k \times n$ equations and m parameters (23)
- Finally, we construct the next interpolating function over the x_j, $j = 0, \ldots, k$ at time t_j and repeat for the procedure

($p_{13} = 0$) leading to linear dependency on the parameters. Using the initial condition of $x(0)$=(0, 16.8, 0, 0, 0, 0) and the input u being fixed to 0.17 nM, we calculated the dynamical behaviour for $x_j(t)$ at a step size of 0.04 min. We then calculated the inverse problem for this situation following the protocol in Table 1. We thereby used the time series values of x_1 at $t_1 = 1$ (min) and x_2 at $t_2 = 2$ (min) that determines the first interpolating function of the solution ϕ. We subsequently addressed the problem including the negative feedback, where we recalculated the direct probem with parameters above, yet including now the negative SOCS3 feedback and setting $p_{13} = 0.029$. Using the newly generated set of state variables at large time points (i. e. at the steady state), we can recover p_{13} using the steady state condition where the temporal changes in the system (17) are set to zero. In particular, the third component of Equation (19) becomes

$$0 = \frac{p_5 \left(\frac{\alpha - x_2}{2} - x_3\right)}{1 + p_{13} x_6} - p_6 \, x_3$$

which yields

$$p_{13} = \left[\frac{p_5 \cdot \left(0.5 \cdot (x_2^{\text{total}} - x_2(\infty)) - x_3(\infty)\right)}{p_6 \cdot x_3(\infty)} - 1 \right] \cdot \frac{1}{x_6(\infty)} \tag{24}$$

Thereby, we have used the values for x_2, x_3, x_6 at steady state conditions $x_2(\infty)$, $x_3(\infty)$, $x_6(\infty)$, taken their numerical approximation at late simulation times, $t \approx 90$ min.

Analytical and numerical calculations were performed in Mathematica 11 (Wolfram Research, Champain, Illinois). Results of numeric calculations are depicted in Table 2. There, recalculated parameters were retrieved with an accuracy of 90% for smaller and 95% for larger parameters.

5. DISCUSSION AND CONCLUSIONS

We devised here a method to retrieve ODE model parameters from experimental time series based on Picard iteration. The method was able to recover correct parameters for a medium sized inverse problem, the Jak-STAT3 activation pathway, provided accurate time series of all involved entities and assuming linear dependency of the model parameters. To guarantee linear dependency, we first applied our method to a subproblem where feedback was neglected and then demonstrated how the feedback parameter can be retrieved using steady state conditions. We note that, neglecting this feedback and determining it in a second step requires also to generate experimental data where this feedback is broken. This can be achieved either by chemical inhibitors or by point mutations that block the inhibitor or render it useless through missense mutation. We further note that restriction to linear parameters is a feature of our current implementation solving a Gaussian system rather than a feature of the Picard implementation and expect future implementations to address the issue of non-linearity in an appropriate fashion.

Table 2. Assumed kinetic parameter (17) and those recalculated from the proposed inversion method. The asterisk denotes the calculation via steady state conditions.

Parameter	**Assumed**	**Calculated**
p_1	0.122	0.1218
p_2	0.04	0.0388
p_3	3.59	3.5920
p_4	0.05	0.0484
p_5	0.08	0.0803
p_6	0.08	0.0864
p_7	0.16	0.156
p_8	0.09	0.010
p_9	0.03	0.026
p_{10}	0.02	0.021
p_{11}	0.03	0.029
p_{12}	0.01	0.008
p_{13}	0.029	0.028*

For our analysis, we posit that the biological process can be described by a given topology of ODEs, that the entire set of biological intermediaries (such as truncated proteins, multimers, etc.) is detectable and that the initial conditions are given. Indeed, while measuring these biological intermediaries can be cumbersome, it is expected that they will be in future measurable through improvements in protein labelling, immunoprecipitation of multimeric complexes and usage of time-dependent expression techniques (PCR, RNA-Seq).

Several limitations of the methods apply. First, convergence of the Picard iteration requires good initial approximation in the function space. More importantly, our method was sensitive against increase of the time interval of the measurements that were taken into account for generating the Gaussian system. This is especially envisaged to cause issues for stiff systems where certain variables contain less information about temporal changes than others. This drawback may argue for combining our methods with such of experimental design (such as analysing the Fisher Information Matrix) such that the most informative time points for measurements can be taken into account. To this end, we note that our method does not require to use the same time points for all observables. To this end, our numerical experiments allow us to determine the size of the time interval in order to compute the parameters within a certain accuracy, hence providing a hint for the number and distance of measurements in a specific experiment.

Finally, the method requires that all states are detectable. Despite these limitations, however, the method would be suitable for experiments such as single cell fluorescent

imaging (Connolly et al., 2016) where high frequency time lapse-data are available for a large quantity of observables. We further envisage the mehod to be extendible to the inclusion of experimental errors and for disentangling time-scales for fastly and slowly changing state variables. An application to non-linear parameters to cover more sophisticated kinetics such as Michaelis-Menten or Hill kinetics is in preparation.

6. ACKNOWLEDGEMENT

We thank Dr. Eric Bullinger for suggestions.

REFERENCES

Arnold, V.I. (1992). *Ordinary Differential Equations*. Springer, 3rd edition.

Bhatt, N., Kerimoglu, N., Amrhein, M., Marquardt, W., and Bonvin, D. (2012). Incremental identification of reaction systemsA comparison between rate-based and extent-based approaches. *Chemical Engineering Science*, 83, 24–38.

Chis, O.T., Banga, J.R., and Balsa-Canto, E. (2011). Structural identifiability of systems biology models: a critical comparison of methods. *PloS One*, 6(11), e27755.

Chou, I.C. and Voit, E.O. (2009). Recent developments in parameter estimation and structure identification of biochemical and genomic systems. *Mathematical Biosciences*, 219(2), 57–83.

Connolly, N.M., D'Orsi, B., Monsefi, N., Huber, H.J., and Prehn, J.H. (2016). Computational analysis of AMPK-mediated neuroprotection suggests acute excitotoxic bioenergetics and glucose dynamics are regulated by a minimal set of critical reactions. *PloS One*, 11(2), e0148326.

Dittrich, A., Quaiser, T., Khouri, C., Görtz, D., Mönnigmann, M., and Schaper, F. (2012). Model-driven experimental analysis of the function of SHP-2 in IL-6-induced Jak/STAT signaling. *Molecular BioSystems*, 8(8), 2119–2134.

Docherty, P.D., Chase, J.G., and David, T. (2012). Characterisation of the iterative integral parameter identification method. *Medical & Biological Engineering & Computing*, 50(2), 127–134.

Endo, T.A., Masuhara, M., Yokouchi, M., Suzuki, R., Sakamoto, H., Mitsui, K., Matsumoto, A., Tanimura, S., Ohtsubo, M., Misawa, H., et al. (1997). A new protein containing an SH2 domain that inhibits JAK kinases. *Nature*, 387(6636), 921–924.

Farina, M., Findeisen, R., Bullinger, E., Bittanti, S., Allgower, F., and Wellstead, P. (2006). Results towards identifiability properties of biochemical reaction networks. In *Proc. of the IEEE Conference on Decision and Control*, 2104–2109.

Geffen, D., Findeisen, R., Schliemann, M., Allgower, F., and Guay, M. (2008). Observability based parameter identifiability for biochemical reaction networks. In *Proc. of the IEEE American Control Conference*, 2130–2135.

Heinrich, P., Behrmann, I., Haan, S., Hermanns, H., Muller-Newen, G., and Schaper, F. (2003). Principles of Interleukin (IL)-6-type cytokine signalling and its regulation. *Journal of Biochemistry*, 374, 1–20.

Huber, H. and Leeb, H. (1998). An exact inversion method for the determination of spinorbit potentials from scattering data. *The European Physical Journal A - Hadrons and Nuclei*, 1(2), 221–228.

Kunze, H.E. and Vrscay, E.R. (1999). Solving inverse problems for ordinary differential equations using the picard contraction mapping. *Inverse Problems*, 15(3), 745.

Lillacci, G. and Khammash, M. (2010). Parameter estimation and model selection in computational biology. *PLoS Computational Biology*, 6(3), e1000696.

Liu, Y. and Gunawan, R. (2014). Parameter estimation of dynamic biological network models using integrated fluxes. *BMC Systems Biology*, 8(1), 127.

López-Caamal, F., Middleton, R.H., and Huber, H.J. (2014). Equilibria and stability of a class of positive feedback loops. *Journal of Mathematical Biology*, 68(3), 609–645.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Network motifs: simple building blocks of complex networks. *Science*, 298(5594), 824–827.

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

Otter, P.W. (1986). Dynamic structural systems under indirect observation: Identifiability and estimation aspects from a system theoretic perspective. *Psychometrika*, 51(3), 415–428.

Rumschinski, P., Borchers, S., Bosio, S., Weismantel, R., and Findeisen, R. (2010). Set-based dynamical parameter estimation and model invalidation for biochemical reaction networks. *BMC Systems Biology*, 4(1), 69.

Savageau, M.A. (1969a). Biochemical systems analysis: I. some mathematical properties of the rate law for the component enzymatic reactions. *Journal of Theoretical Biology*, 25(3), 365–369.

Savageau, M.A. (1969b). Biochemical systems analysis: II. The steady-state solutions for an n-pool system using a power-law approximation. *Journal of Theoretical Biology*, 25(3), 370–379.

Schaber, J. and Klipp, E. (2011). Model-based inference of biochemical parameters and dynamic properties of microbial signal transduction networks. *Current Opinion in Biotechnology*, 22(1), 109–116.

Scheller, J., Ohnesorge, N., and Rose-John, S. (2006). Interleukin-6 trans-signalling in chronic inflammation and cancer. *Scandinavian Journal of Immunology*, 63(5), 321–329.

Schliemann-Bullinger, M., Fey, D., Bastogne, T., Findeisen, R., Scheurich, P., and Bullinger, E. (2016). The experimental side of parameter estimation. In *Uncertainty in Biology*, 127–154. Springer.

Tanner, R. (1972). Estimating kinetic rate constants using orthogonal polynomials and picard's iteration method. *Industrial and Engineering Chemistry Fundamentals*, 11(1), 1–8.

Torres, N.V. and Voit, E.O. (2002). *Pathway analysis and optimization in metabolic engineering*. Cambridge University Press.

Veltman, D., Laeremans, T., Passante, E., and Huber, H.J. (2017). Signal transduction analysis of the NLRP3-inflammasome pathway after cellular damage and its paracrine regulation. *Journal of Theoretical Biology*, 415, 125–136.

Voit, E.O. (2013). Biochemical systems theory: a review. *ISRN Biomathematics*, 2013.