Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood

A. Raue[1,*], C. Kreutz[1], T. Maiwald[2], J. Bachmann[3], M. Schilling[3], U. Klingmüller[3] and J. Timmer[1,4]

[1]Physics Institute, University of Freiburg, 79104 Freiburg, Germany, [2]Department of Systems Biology, Harvard Medical School, 02115 Boston, MA, USA, [3]Division of Systems Biology of Signal Transduction, DKFZ-ZMBH Alliance, German Cancer Research Center, 69120 Heidelberg and [4]Freiburg Institute for Advanced Studies, University of Freiburg, 79104 Freiburg, Germany

ABSTRACT

Motivation: Mathematical description of biological reaction networks by differential equations leads to large models whose parameters are calibrated in order to optimally explain experimental data. Often only parts of the model can be observed directly. Given a model that sufficiently describes the measured data, it is important to infer how well model parameters are determined by the amount and quality of experimental data. This knowledge is essential for further investigation of model predictions. For this reason a major topic in modeling is *identifiability analysis*.

Results: We suggest an approach that exploits the *profile likelihood*. It enables to detect *structural non-identifiabilities*, which manifest in functionally related model parameters. Furthermore, *practical non-identifiabilities* are detected, that might arise due to limited amount and quality of experimental data. Last but not least *confidence intervals* can be derived. The results are easy to interpret and can be used for *experimental planning* and for *model reduction*.

Availability: An implementation is freely available for MATLAB and the PottersWheel modeling toolbox at

http://web.me.com/andreas.raue/profile/software.html.

Contact: andreas.raue@me.com

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 INTRODUCTION

Inferring cell biological questions by mathematical modeling of reaction networks became a popular and powerful approach (Kitano, 2005). Tools to build models for complex reaction networks and calibrate model parameters to experimental data are available (Maiwald and Timmer, 2008; Schmidt and Jirstrand, 2006). Statistical tests were established to evaluate, whether a model can explain experimental data sufficiently, as well as to compare the performance of different models or model extensions, e.g. Ghosh and Samanta (2001).

Furthermore, it is usually desired to use an established model for prediction of: model parameters such as rate constants or initial concentrations; model trajectories such as time-courses of experimentally unobserved species concentrations; model behaviour under changed environmental conditions such as altered network structure or different external stimulation. Since the considered models are parametric, these predictions depend intrinsically on the previously calibrated model parameters.

Due to technical limitations, e.g. availability of specific antibodies, biological reaction networks are often only partially observable. This means that not all species incorporated in a model can be measured directly. Given a certain amount and quality of experimental data measured under specific experimental conditions, it is not assured that model parameters can be estimated unambiguously. Frequently, experimental data are insufficient considering the size of the model which results in parameters that are *non-identifiable* (Swameye *et al.*, 2003). Even identifiable parameters can only be determined within *confidence intervals*, which contain the true value of the parameter with a desired probability (Lehmann and Leo, 1983). If model parameters are not well determined also model predictions are not. Consequently, the biological question that should be answered by the model, might not be addressable. Our aim is to develop an approach that enables evaluating which parameters are identifiable, thus infering which model predictions are feasible. Provided that parameters are identifiable, the question that follows is how large their confidence intervals are, which indicates how reliable a model prediction is.

After introducing parameter estimation and discussing how confidence intervals can be derived, different types of identifiability are formulated. A brief overview of existing approaches for identifiability analysis including their assets and drawbacks is given. Subsequently, a novel approach for identifiability analysis by exploiting the *profile likelihood* will be introduced. This approach is able to detect both structural and practically non-identifiable parameters and simultaneously calculates confidence intervals. Since large models are under consideration, the approach needs to be computationally feasible and its output should be

interpretable even if depending on a high-dimensional parameter space. Furthermore, the approach can be used for *experimental planning* to suggest additional measurements that efficiently reduce parameter uncertainties and for *model reduction* to tailor the model complexity to the information content given by the experimental data. Usage and benefit of the approach will be illustrated by applying it to a model of the JAK-STAT signaling pathway, that is calibrated to experimental data.

2 PROBLEM STATEMENT

Given a model $\mathcal{M}$ describing n species concentrations x_i in a reaction network by a system of ordinary differential equations (ODE)

$$\dot{\vec{x}}(t) = f(\vec{x}(t), \vec{u}(t), \vec{p}) \tag{1}$$

$$\vec{y}(t) = g(\vec{x}(t), \vec{s}) + \vec{\epsilon}(t) \tag{2}$$

with *internal model states* $\vec{x}(t)$, an externally given *stimulus* $\vec{u}(t)$, *dynamic parameters* $\vec{p}$, an *m*-dimensional mapping g of the internal model states to the *observables* $\vec{y}(t)$ involving scaling and offset parameters $\vec{s}$. The measurement noise $\vec{\epsilon}(t)$ is assumed to be normally distributed. For partially observed models, the dimension m of observations is smaller than the dimensions n of internal model states. Together with the initial concentrations $\vec{x}(0)$ for Equation (1), we define

$$\theta = \{\vec{p}, \vec{x}(0), \vec{s}\} \tag{3}$$

as set of parameters necessary to fully specify $\mathcal{M}$. For parameters in biological reaction networks, e.g. rate constants or initial concentrations usually $\theta \in \mathbf{R}^+ \setminus \{\mathbf{0}\}$. To avoid the natural lower bound of zero, logarithmic parameter values will be used in the following.

2.1 Parameter estimation

The agreement of experimental data with the observables predicted by the model is measured by an objective function, commonly the weighted sum of squared residuals

$$\chi^2(\theta) = \sum_{k=1}^{m} \sum_{l=1}^{d} \left(\frac{y_{kl}^D - y_k(\theta, t_l)}{\sigma_{kl}^D} \right)^2 \tag{4}$$

where y_{kl}^D denotes d data-points for each observable k, measured at time-points t_l. σ_{kl}^D are the corresponding measurement errors and $y_k(\theta, t_l)$ the k-th observable as predicted by parameters θ for time-point t_l. The parameters can be estimated numerically by

$$\hat{\theta} = \operatorname{argmin}\left[\chi^2(\theta)\right]. \tag{5}$$

For normally distributed observational noise $\vec{\epsilon} \sim N(0, \sigma^2)$, this corresponds to the maximum likelihood estimate (MLE) of θ and

$$\chi^2(\theta) = \text{const} - 2 \cdot \log(L(\theta)) \tag{6}$$

where $L(\theta)$ is the likelihood. In the following, χ^2 will be used as placeholder for the likelihood. Furthermore, an appropriate model $\mathcal{M}$ that sufficiently describes the available experimental data is assumed.

2.2 Confidence intervals

A confidence interval $[\sigma_i^-, \sigma_i^+]$ of a parameter estimate $\hat{\theta}_i$ to a confidence level α signifies that the true value θ_i^* is located within this interval with probability α. In the following, *asymptotic* and *finite sample* confidence intervals will be introduced.

2.2.1 Asymptotic confidence intervals Confidence intervals can be derived from the curvature of the likelihood, e.g. the Hessian matrix $\mathbf{H} = \nabla^T \nabla \chi^2|_{\hat{\theta}_i}$. Using the covariance matrix $\mathbf{C} = 2 \cdot \mathbf{H}^{-1}$ of the parameter estimates, asymptotic confidence intervals are given by

$$\sigma_i^{\pm} = \hat{\theta}_i \pm \sqrt{\chi^2(\alpha, df) \cdot \mathbf{C}_{ii}} \tag{7}$$

where $\chi^2(\alpha, df)$ is the α quantile of the χ^2-distribution with df degrees of freedom, as explained in Press *et al.* (1990). The choice of df yields two different types of confidence intervals: $df = 1$ gives *pointwise* confidence intervals that hold individually for each parameter, $df = \#\theta$ being the number of parameters gives *simultaneous* confidence intervals that hold jointly for all parameters.

Asymptotic confidence intervals are a good approximation of the actual uncertainty of $\hat{\theta}_i$, if the amount of experimental data is large compared to $\#\theta$ and/or the measurement noise is small. They are exact if the observables $\vec{y}$ depend linearly on θ. However, even for the simplest reaction network the observables $\vec{y}$ depend non-linearly on θ and the amount and quality of experimental data is often insufficient. Therefore, asymptotic confidence intervals might not be appropriate (Joshi *et al.*, 2006).

2.2.2 Finite sample confidence intervals Confidence intervals can also be derived using a threshold in the likelihood. These so called *likelihood-based confidence intervals* are defined by a confidence region

$$\{\theta \mid \chi^2(\theta) - \chi^2(\hat{\theta}) < \Delta_\alpha\} \quad \text{with} \quad \Delta_\alpha = \chi^2(\alpha, df) \tag{8}$$

whose borders represent confidence intervals (Meeker and Escobar, 1995). The threshold Δ_α is the α quantile of the χ^2-distribution and represents with $df = 1$ and $df = \#\theta$ pointwise, respectively, simultaneous confidence intervals to a confidence level α [see Equation (7)]. Likelihood-based confidence intervals are considered superior to asymptotic confidence intervals for finite samples (Neale and Miller, 1997).

2.3 Identifiability

A parameter θ_i is identifiable, if the confidence interval $[\sigma_i^-, \sigma_i^+]$ of its estimate $\hat{\theta}_i$ is finite. Two phenomena accounting for parameters to be non-identifiable will be discussed here. *Structural non-identifiability* is related to the model structure independent of experimental data which is extensively discussed, e.g. Cobelli and DiStefano III (1980). In contrast, *practical non-identifiability* also takes into account the amount and quality of measured data, that was used for parameter calibration. Practical non-identifiability is less clearly defined in literature, therefore an independent definition will be given.

2.3.1 Structural non-identifiability A structural non-identifiability arises from a redundant parameterization in the formal solution of $\vec{y}(t)$, due to an insufficient mapping g of
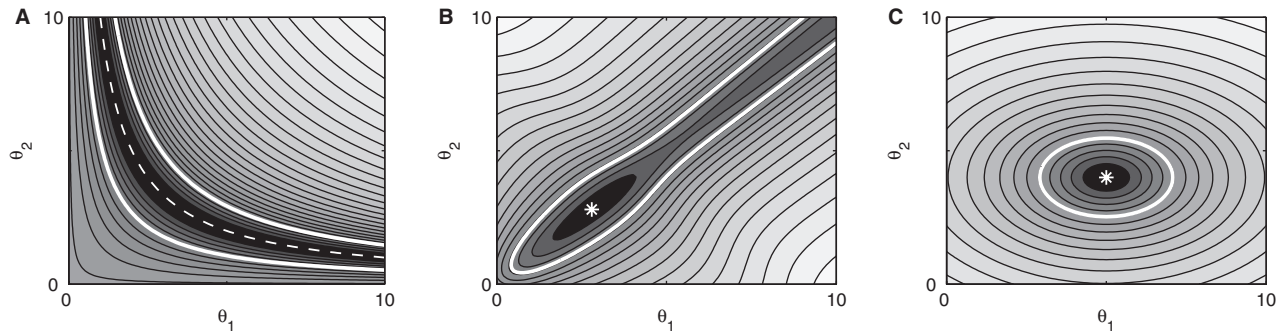
Fig. 1. Contour plots of $\chi^2(\theta)$ for a two-dimensional parameter space, shown in non-logarithmic scale for illustrative reasons. Shades from black to white correspond to low and high values of χ^2, respectively. Thick white lines display likelihood-based confidence regions and white stars the optimal parameter estimates $\hat{\theta}$. Left panel: a structural non-identifiability along the functional relation $h(\theta) = \theta_1 \cdot \theta_2 - 10 = 0$ (dashed line). The likelihood-based confidence region is infinitely extended. Middle panel: a practical non-identifiability. The likelihood-based confidence region is infinitely extended for $\theta_1 \to +\infty$ and $\theta_2 \to +\infty$, lower confidence bounds can be derived. Right panel: both parameters are identifiable.

internal model states $\vec{x}$ to observables $\vec{y}$ in Equation (2). The set of ambiguous parameters $\theta_{\text{sub}} \subset \theta$ may be varied without changing the observables $\vec{y}(t)$, hence keeping $\chi^2(\theta)$ on a constant value. The redundant parameterization manifests as functional relations $\vec{h}$ between the parameters θ_{sub}, representing a manifold with constant χ^2 in parameter space

$$\chi^2(\theta) = \chi^2(\hat{\theta}) \quad \Leftrightarrow \quad \vec{h}(\theta_{\text{sub}}) = 0. \qquad (9)$$

Consequently, the parameter estimates $\hat{\theta}_{\text{sub}}$ and, respectively, the internal model states $\vec{x}(t)$ affected by these parameters are not uniquely identified. Confidence intervals of a structurally non-identifiable parameter $\theta_i \in \theta_{\text{sub}}$ are infinite $[-\infty, +\infty]$ in logarithmic parameter space considered here. Hence, its value cannot be estimated at all. A direct detection of a redundant parameterization in the analytic form of $\vec{y}(t)$ is hampered, because Equation (1) can only be solved analytically in special cases.

$\chi^2(\theta)$ in a two-dimensional parameter space can be visualized as a landscape. A structural non-identifiability results in a perfect flat valley, infinitely extended along the corresponding functional relation, as illustrated in Figure 1, left panel.

Since a structural non-identifiability is independent of the accuracy of available experimental data, it cannot be resolved by a refinement of existing measurements, e.g. by reducing the measurement noise $\vec{\epsilon}(t)$. The only remedy is a qualitatively new measurement which alters the mapping g, e.g. by increasing the number of observed species. A parameter is *structural identifiable*, if a unique minimum of $\chi^2(\theta)$ with respect to θ_i exists.

2.3.2 Practical non-identifiability A parameter that is structurally identifiable may still be practically non-identifiable. This arises frequently if amount and quality of experimental data is insufficient and manifests in a confidence interval that is infinite. Please note that the asymptotic confidence interval of a structural identifiable parameter estimate may be large, but is always finite because $\mathbf{C}_{ii} > 0$ [see Equation (7)]. Therefore, it is not possible to infer practical non-identifiability using asymptotic confidence intervals. We propose a definition that is inspired by likelihood-based confidence intervals [see Equation (8)]:

DEFINITION 1. A parameter estimate $\hat{\theta}_i$ is *practically non-identifiable*, if the likelihood-based confidence region is infinitely

extended in increasing and/or decreasing direction of θ_i, although the likelihood has a unique minimum for this parameter.

This means that the increase in χ^2 stays below the threshold Δ_α for a desired confidence level α in direction of θ_i. Similar to a structural non-identifiability, the flattening out of the likelihood can continue along a functional relation. The confidence interval of a practically non-identifiable parameter is not necessarily extended infinitely to both sides. There can be a finite upper or lower bound of the confidence interval $[\sigma_i^-, \sigma_i^+]$ where either $\sigma_i^- = -\infty$ or $\sigma_i^+ = +\infty$.

In a two-dimensional parameter space, a practical non-identifiability can be visualized as a relatively flat valley, which is infinitely extended. The height distance of the valley bottom to the lowest point $\hat{\theta}$ never excesses Δ_α, as illustrated in Figure 1, middle panel.

Along a practical non-identifiability, the observable $\vec{y}$ change only negligibly remaining compliant with the given measurement accuracy. Nevertheless, model behavior in terms of internal states $\vec{x}$ might vary strongly. Improving the detection of typical dynamical behavior by increasing the amount and quality of measured data and/or the choice of measurement time-points t_{ij} will ultimately resolve a practical non-identifiability, yielding finite likelihood-based confidence intervals (Fig. 1, right panel). Inferring how to decrease confidence intervals most efficiently is the subject of experimental planning, which will be discussed later on.

3 EXISTING METHODS

Various methods exist to detect structural non-identifiability by a priori analyzing the system equations (1) and (2), such as the Power Series Expansion (Pohjanpalo, 1978), the Volterra and Generating Power Series Approach (Lecourtier *et al.*, 1987), the Similarity Transform Approach (Vajda *et al.*, 1989b) or differential algebraic methods (e.g. Ljung and Glad, 1994). Unfortunately, these methods become rapidly infeasible with increasing model size (Margaria *et al.*, 2001; White *et al.*, 2001). Practical non-identifiability cannot be detected, since experimental data are disregarded.

Another class of methods aims to detect non-identifiability by flatness of likelihood, using simulated or experimental data. Here, measures of curvature are computed, commonly using a quadratic approximation of χ^2 at the estimated optimum $\hat{\theta}$, e.g. the Hessian

or Fisher information matrix (Jacquez and Greif, 1985; Vajda *et al.*, 1989a; Yao *et al.*, 2003). These methods are appropriate if functional relations $\vec{h}$ between the parameters emerging from structural non-identifiability are linear. This is often not the case for reaction networks modeled by ODE, because observables depend non-linearly on the parameters. Practical non-identifiability cannot be detected, because a quadratic approximation is not able to explain increasing but limited behavior of $\chi^2(\theta)$ as mentioned earlier.

An approach to detect structural non-identifiability by the corresponding functional relations was introduced by Hengl *et al.* (2007). It is able to detect flatness of likelihood for arbitrary models, but it is not intended to detect practical non-identifiability.

Similar to a clear formal definition, an approach for explicit testing of practical non-identifiability is not available to our knowledge. In the following, we introduce a general approach to analyze arbitrary models for structural and practical non-identifiability.

4 APPROACH

The idea of the approach is to explore the parameter space for each parameter in the direction of the least increase in χ^2. For a structurally non-identifiable parameter this means to follow the functional relations $\vec{h}(\theta_{sub}) = 0$. In case of a practically non-identifiable parameter, the aim is to detect directions where the likelihood flattens out.

A useful concept for this task is the *profile likelihood* (PL) χ^2_{PL} (Murphy and van der Vaart, 2000; Venzon and Moolgavkar, 1988). It can be calculated for each parameter individually by

$$\chi^2_{PL}(\theta_i) = \min_{\theta_{j \neq i}} \left[\chi^2(\theta) \right] \qquad (10)$$

meaning re-optimization of $\chi^2(\theta)$ with respect to all parameters $\theta_{j \neq i}$, for each value of parameter θ_i. Hence, the profile likelihood keeps χ^2 as small as possible alongside θ_i. Figure 1 illustrates that the likelihood is explored in the desired way to detect non-identifiabilities. An algorithm to calculate χ^2_{PL} is described in the Supplementary Material.

Structural non-identifiable parameters are characterized by a flat profile likelihood [Equation (9)]. The profile likelihood of a practically non-identifiable parameter has a minimum, but is not excessing a threshold Δ_α for increasing and/or decreasing values of θ_i (see Definition 1). In contrast, the profile likelihood of an identifiable parameter exceeds Δ_α for both increasing and decreasing values of θ_i. The points of passover represent likelihood-based confidence intervals as defined in Equation (8) (Royston, 2007). By following the change of parameters $\theta_{j \neq i}$ along $\chi^2_{PL}(\theta_i)$, the functional relations $\vec{h}(\theta_{sub}) = 0$ corresponding to a structural non-identifiability can be recovered.

Experimental planning To improve certainty of a specific model prediction, it would be valuable to suggest additional measurements that efficiently resolve non-identifiability and narrow the confidence interval of a parameter θ_i affecting this issue. The set of trajectories along the profile likelihood of θ_i reveals spots where the uncertainty of θ_i has the largest impact on the model. Additional measurements at these spots are likely to efficiently reduce this uncertainty. The amplitude of variability of the trajectories at these spots allows to assess the necessary precision of a new measurement to provide adequate data that is able to improve parameter identification.
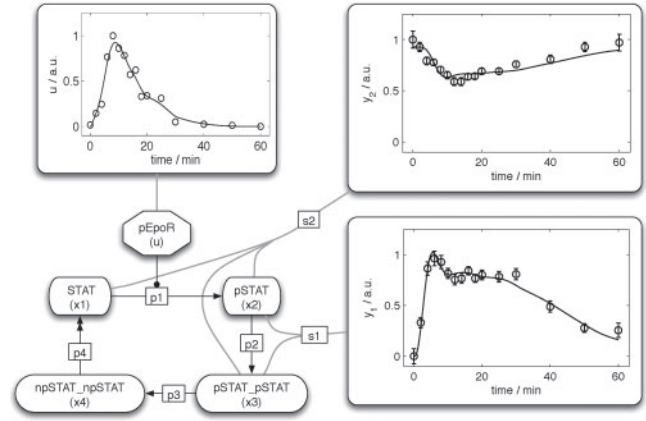

Fig. 2. Network structure of the model and observables calibrated to experimental data. A spline interpolation of pEpoR measurements serves as external stimulus *u*. The double arrowhead denotes a delay reaction modeled by a linear chain approximation.

The impact of new measurements can be evaluated by Monte Carlo simulations. To this aim, the described analysis of the profile likelihood is repeated, taking into account additional simulated data. The resulting change of the profile likelihood and correspondingly the resolution of non-identifiability and the narrowing of the likelihood-based confidence intervals allow to justify the effort of new measurements to gain a more confident model prediction.

Model reduction The approach can be used for model reduction by considering a threshold Δ_α with $df = 1$ [see Equation (8)]. Assume that a parameter θ_i is practically non-indentifiable for decreasing parameter value. Consider a reduced model $\mathcal{M}^*$ with simplified kinetics concerning θ_i, e.g. for mass action kinetics by removing the corresponding reaction. In this case, the threshold Δ_α corresponds to a likelihood ratio test of the reduced model $\mathcal{M}^*$ against the original model $\mathcal{M}$ to a significance level $1 - \alpha$. Falling below this threshold, the profile likelihood indicates that it is not possible to dismiss $\mathcal{M}^*$ in favor of $\mathcal{M}$, based on the available experimental data.

5 APPLICATION

To illustrate usage and benefit of the approach, it was applied to a model of the JAK-STAT signaling pathway inspired by Swameye *et al.* (2003), which is calibrated to the experimental data available at http://webber.physik.uni-freiburg.de/~jeti/PNAS_Swameye_Data/ (dataset 1). The model represents the STAT signaling cascade including nuclear shuttling upon stimulation with Erythropoietin: phosphorylation of cytoplasmatic STAT (x_1) triggered by active Erythropoietin receptor pEpoR (u); homo-dimerization of pSTAT (x_2); import of the pSTAT_pSTAT complex (x_3) into the nucleus; dissociation and dephosphorylation of npSTAT_npSTAT (x_4) and export to cytoplasm (Fig. 2). A spline interpolation of pEpoR measurements serves as external stimulation *u* for STAT phosphorylation. In terms of ODE the model reads as

$$\begin{aligned}
\dot{x}_1 &= -p_1 \cdot x_1 \cdot u + 2 \cdot p_4 \cdot x_4^\tau & y_1 &= s_1 \cdot (x_2 + 2 \cdot x_3) \\
\dot{x}_2 &= +p_1 \cdot x_1 \cdot u - p_2 \cdot x_2^2 & y_2 &= s_2 \cdot (x_1 + x_2 + 2 \cdot x_3) \\
\dot{x}_3 &= +\tfrac{1}{2} \cdot p_2 \cdot x_2^2 - p_3 \cdot x_3 \\
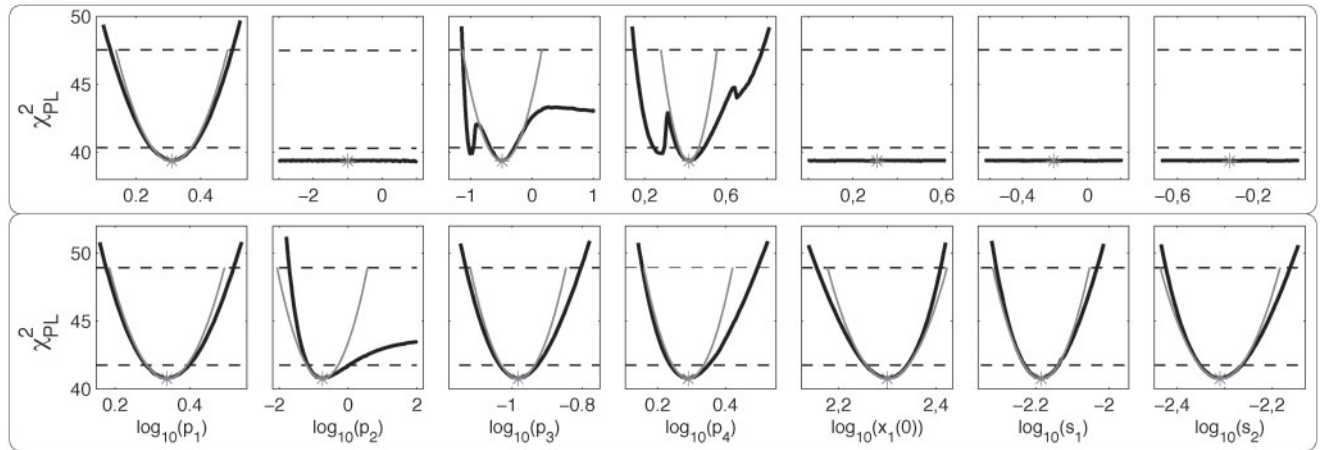\dot{x}_4 &= +p_3 \cdot x_3 - p_4 \cdot x_4^\tau
\end{aligned}$$

Fig. 3. Black lines display profile likelihood versus parameter. The results for the original dataset are shown in the upper panel. The lower panel shows the alteration of the profile likelihood, after taking into account the additional data. Calibrated parameter values $\hat{\theta}$ are displayed by gray stars, thresholds for simultaneous and pointwise 1-σ confidence intervals by upper, respectively, lower dashed lines. Gray parabolas indicate the quadratic approximation used for asymptotic confidence intervals. They are very flat for the structurally non-identifiable parameters. Discontinuities in the profile likelihood of p_3 and p_4 in the upper panel stem from local minima that govern the profile likelihood in remote regions. Parameter values are given in orders of magnitude.

where the superscript τ denotes a delay reaction implemented by a linear chain approximation, yielding an effective delay $\tau = 10/p_4$ with 10 intermediate steps (MacDonald, 1976). Experimentally observable quantities are phosphorylated STAT in cytoplasm (y_1) and total STAT in cytoplasm (y_2), both measured in arbitrary units by quantitative western blotting. Two compartments are considered in the model: the cytoplasm and the nucleus with $1400\mu m^3$ and $450\mu m^3$. Species concentrations are modeled in nanometer. Besides the dynamic parameters p_1 to p_4 and scaling parameters s_1 and s_2, the initial concentration $x_1(0)$ belongs to the parameters θ that need to be calibrated by the experimental data. The initial values of species x_2 to x_4 are assumed to be equal to zero. The model is implemented and calibrated using the PottersWheel fitting toolbox (Maiwald and Timmer, 2008), resulting trajectories of observables are shown in Figure 2.

Calculating the profile likelihood takes 54 ± 18 s per parameter, using an implementation of the approach embedded in the Potters-Wheel fitting toolbox (1 GHz CPU, 2 GB RAM), which is described in the Supplementary Material. The resulting plots of profile likelihood versus parameter reveal four structurally non-identifiable parameters p_2, $x_1(0)$, s_1, s_2 by their flat profile likelihood, see Figure 3, upper panel.

Structural non-identifiability The functional relations $\vec{h}$ connecting these structurally non-identifiable parameters can be recovered from the changes of the remaining parameters, while calculating the profile likelihood of the structurally non-identifiable parameters, as shown in Figure 4. Using, for example, $\chi^2_{PL}(x_1(0))$, the manifold can be characterized as

$$h_1(\theta_{sub}) = p_2 \cdot x_1(0) - \text{const} = 0$$
$$h_2(\theta_{sub}) = s_1 \cdot x_1(0) - \text{const} = 0 \qquad (11)$$
$$h_3(\theta_{sub}) = s_2 \cdot x_1(0) - \text{const} = 0$$

which is compliant with analytical considerations given in Timmer *et al.* (2004). Recovering the functional relations unambiguously from the change of parameters along the profile likelihood is only
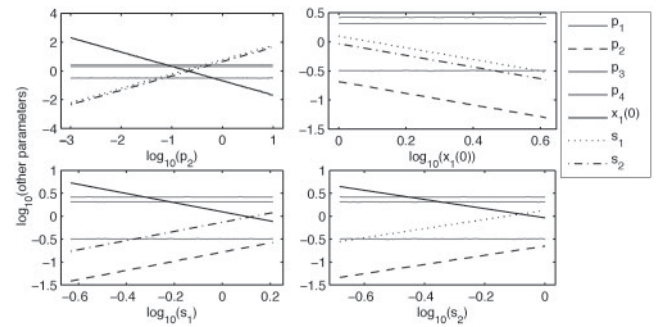

Fig. 4. While exploiting the profile likelihood for each of the structurally non-identifiable parameters, the parameters connected by this structural non-identifiability change accordingly (black lines). Thereby the functional relations between p_2, $x_1(0)$, s_1, s_2 can be characterized. Parameters that are not involved in the structural non-identifiability are unaffected (gray horizontal lines).

possible if the corresponding manifold is one-dimensional. This is because one parameter is fixed at a time while computing the profile likelihood. To recover functional relations that correspond to manifolds with dimension larger than one, a further analysis of the functionally related subsets is necessary, e.g. the approach by Hengl *et al.* (2007).

The variability of the internal model states $\vec{x}$ imposed by this structural non-identifiability can be analyzed by plotting the trajectories for parameter values along the profile likelihoods of the structural non-identifiable parameters (Fig. 5). All internal model states can only be identified up to a common factor. This implies that the structural non-identifiability is derived from the fact that no information about absolut concentration is included in the experimental data.

Practical non-identifiability Thresholds $\Delta_{\alpha=0.68}$ for both pointwise and simultaneous 1$-\sigma$ confidence intervals are displayed
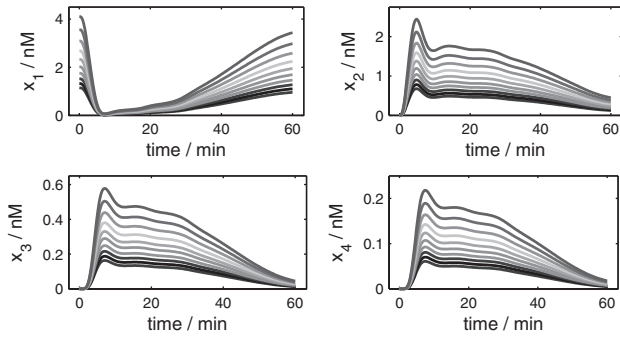
Fig. 5. Trajectories of internal model states $\vec{x}$ for parameter values along the structural non-identifiability. All internal model states can only be identified up to a common factor.
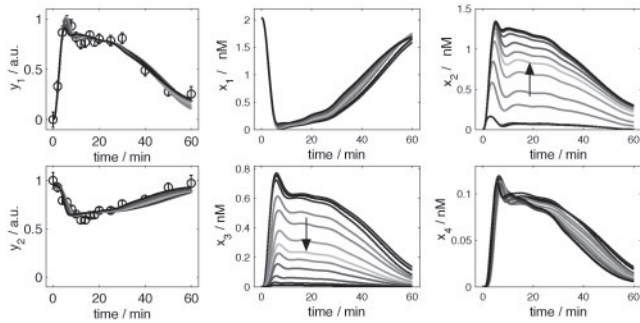

Fig. 6. Trajectories of observables $\vec{y}$ and internal model states $\vec{x}$ along the profile likelihood of the practically non-identifiable parameter p_3. Please note, that species x_2 and x_3 change antipodal.

in Figure 3. In the following, we will consider the more reliable higher threshold yielding simultaneous confidence intervals.

The profile likelihood reveals that parameter p_3 is practically non-identifiable for increasing parameter values. This indicates that the amount and quality of the experimental data provided does not contain enough information to yield an upper limit for the rate of nuclear import p_3. The variability of the predicted model observables $\vec{y}$ along the profile likelihood of p_3 remains consistent with the measurement errors, as shown in Figure 6. Nevertheless, a lower confidence bound can be derived.

Confidence intervals Table 1 compares finite sample confidence intervals derived from the profile likelihood and asymptotic confidence intervals derived from the Hessian matrix. For identifiable parameters the discrepancies are small. Working in logarithmic parameter space linearizes the functional relations given in Equation (11). In this case, asymptotic approximation leads to large but nevertheless finite confidence intervals for the structurally non-identifiable parameters. The largest discrepancy occurs for practically non-identifiable parameters, where asymptotic confidence intervals are significantly smaller than likelihood-based confidence intervals. Asymptotic confidence intervals are also indicated in Figure 3 by gray parabolas.

Experimental planning To resolve the structural non-identifiability between parameters p_2, $x_1(0)$, s_1, s_2, a measurement of absolut concentration is necessary as mentioned earlier. Figure 5 shows trajectories for parameters along this non-identifiability.

Table 1. Likelihood-based confidence intervals $\sigma^{\pm,\mathrm{PL}}$ derived from the profile likelihood are compared with asymptotic confidence intervals $\sigma^{\pm,\mathrm{Hess}}$ derived from the Hessian matrix [see Equation (7)]

Name	$\hat{\theta}_i$	Non-identifiability	$\sigma^{-,\mathrm{PL}}$	$\sigma^{+,\mathrm{PL}}$	$\sigma^{-,\mathrm{Hess}}$	$\sigma^{+,\mathrm{Hess}}$
p_1	$+0.31$		$+0.12$	$+0.50$	$+0.14$	$+0.48$
p_2	-1.00	Structural	$-\infty$	$+\infty$	-33	$+31$
p_3	-0.49	Practical	-1.14	$+\infty$	-1.14	$+0.15$
p_4	$+0.42$		$+0.15$	$+0.78$	$+0.28$	$+0.56$
$x_1(0)$	$+0.31$	Structural	$-\infty$	$+\infty$	-31	$+32$
s_1	-0.21	Structural	$-\infty$	$+\infty$	-32	$+31$
s_2	-0.34	Structural	$-\infty$	$+\infty$	-32	$+31$

Values are given in orders of magnitude and correspond to $1-\sigma$ simultaneous confidence intervals.

Spots of largest variability suggest where and when a measurement of a species most efficiently determines theses parameters: x_1 at times $t=0$ or $t>50$; x_2 to x_4 at times $5<t<30$.

The trajectories of the internal model states $\vec{x}$ along the profile likelihood of the practically non-identifiable parameter p_3 shown in Figure 6 comprises large antipodal variability of species x_2 and x_3, revealing that the experimental setup is inappropriate to estimate this parameter with confidence. Therefore, an additional measurement to discriminate phosphorylated STAT species x_2 and x_3 is suggested, e.g. the fraction of dimerized pSTAT relative to total phosphorylated STAT in cytoplasm $x_3/(x_2+x_3)$ between 5 and 30 min. If no further quantities than y_1 and y_2 can be measured directly, a refined measurement of phosphorylated STAT in cytoplasm (y_1) at times $t>50$ or of total STAT in cytoplasm (y_2) at times $t>30$ where largest variability of the observables occurs are the best options.

To evaluate the impact of additional measurements on identifiability and confidence intervals, we assume hypothetic measurements yielding an initial concentration of unphosphorylated STAT in cytoplasm $x_1(0)=200\pm20\,\mathrm{nM}$ and a fraction of $x_3/(x_2+x_3)=0.90\pm0.05$ at time $t=20\,\mathrm{min}$. The recalculated profile likelihood reveals, that parameters p_2, $x_1(0)$, s_1, s_2 become structurally identifiable (Fig. 3, lower panel), by measuring only one of them. This accentuates the benefit of knowing the functional relations between structurally non-identifiable parameters, as given in Equation (11). Parameter p_3 becomes practically identifiable, only previously structurally non-identifiable parameter p_2 remains practically non-identifiable.

6 DISCUSSION

Exploiting the profile likelihood is a powerful approach to infer parameter uncertainties in a high-dimensional parameter space. Since it is a systematic and directed exploration, it has less computational cost than sampling parameter space randomly, which gets intractable for high dimensions. The profile likelihood can be calculated for each parameter separately. Thereby it is possible to restrict the analysis to the parameters relevant for the biological question. Moreover, this allows to perfectly parallelize the approach, which is a major benefit for its scalability. An analysis of the runtime of the approach for a test case model is shown in the Supplementary Material. The approach can be applied to any parameter estimation problem, where a likelihood or a similar objective criterion is

available, e.g. partial differential equations (PDE) or stochastic differential equations (SDE).

The approach results in easily interpretable plots of profile likelihood versus parameter. It can be automated, but an explicit advantage is that the output might be evaluated visually. This gives insight into a complex and high-dimensional parameter space. Structural non-identifiabilities originating from incomplete observation of the internal model states can be detected. Arising from limited amount and quality of experimental data, also practical non-identifiabilities can be inferred. Bridging the gap between identifiability and confidence intervals, the profile likelihood allows to derive likelihood-based confidence intervals for each parameter. Functional relations between parameters occurring due to non-identifiabilities can be recovered. The results of the approach can be used on the one hand to design new experiments that efficiently resolve non-identifiability and narrow confidence intervals and on the other hand to facilitate model reduction. Thus, identifiability analysis ensures that the model complexity is tailored to the information content given by the experimental data. Whether a model that is not well determined by the experimental data, should be reduced or additional data should be measured depends on the biological issue to be addressed.

The approach was applied to a model of the JAK-STAT signaling pathway. Non-identifiable parameters were detected, revealing limitations in the experimental setup. Additional measurements that efficiently improve parameter identification were suggested and validated.

ACKNOWLEDGEMENTS

REFERENCES

Cobelli,C. and DiStefano III,J. (1980) Parameter and structural identifiability concepts and ambiguities: a critical review and analysis. *Am. J. Physiol. Regul. Integr. Comp. Physiol.*, **239**, 7–24.

Ghosh,J. and Samanta,T. (2001) Model selection–an overview. *Curr. Sci.*, **80**, 1135–1144.

Hengl,S. *et al.* (2007) Data-based identifiability analysis of non-linear dynamical models. *Bioinformatics*, **23**, 2612.

Jacquez,J. and Greif,P. (1985) Numerical parameter identifiability and estimability: integrating identifiability, estimability, and optimal sampling design. *Math. Biosci.*, **77**, 201–228.

Joshi,M. *et al.* (2006) Exploiting the bootstrap method for quantifying parameter confidence intervals in dynamical systems. *Metab. Eng.*, **8**, 447–455.

Kitano,H. (2005) International alliances for quantitative modeling in systems biology. *Mol. Syst. Biol.*, **1**, 1–2.

Lecourtier,Y. *et al.* (1987) Volterra and generating power series approaches to identifiability testing. In *Identifiability of Parametric Models*. Pergamon Press, Oxford.

Lehmann,E. and Leo,E. (1983) *Theory of point estimation*. Wiley, Hoboken, New Jersey.

Ljung,L. and Glad,T. (1994) On global identifiability for arbitrary model parametrizations. *Automatica*, **30**, 265–265.

MacDonald,N. (1976) Time delay in simple chemostat models. *Biotechnol. Bioeng.*, **18**, 805–812.

Maiwald,T. and Timmer,J. (2008) Dynamical modeling and multi-experiment fitting with PottersWheel. *Bioinformatics*, **24**, 2037–2043.

Margaria,G. *et al.* (2001) Differential algebra methods for the study of the structural identifiability of rational function state-space models in the biosciences. *Math. Biosci.*, **174**, 1–26.

Meeker,W. and Escobar,L. (1995) Teaching about approximate confidence regions based on maximum likelihood estimation. *Am. Stat.*, **49**, 48–53.

Murphy,S. and van der Vaart,A. (2000) On profile likelihood. *J. Am. Stat. Assoc.*, **95**, 449–485.

Neale,M. and Miller,M. (1997) The use of likelihood-based confidence intervals in genetic models. *Behav. Genet.*, **27**, 113–120.

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

Press,W. *et al.* (1990) *Numerical Recipes: FORTRAN*. Cambridge University Press, Cambridge.

Royston,P. (2007) Profile likelihood for estimation and confidence intervals. *Stata J.*, **7**, 376.

Schmidt,H. and Jirstrand,M. (2006) Systems biology toolbox for matlab: a computational platform for research in systems biology. *Bioinformatics*, **22**, 514–515.

Swameye,I. *et al.* (2003) Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by databased modeling. *Proc. Natl Acad. Sci. USA*, **100**, 1028–1033.

Timmer,J. *et al.* (2004) Modeling the nonlinear dynamics of cellular signal transduction. *Int. J. Bifurcat. Chaos*, **14**, 2069–2079.

Vajda,S. *et al.* (1989a) Qualitative and quantitative identifiability analysis of nonlinear chemical kinetic models. *Chem. Eng. Commun.*, **83**, 191–219.

Vajda,S. *et al.* (1989b) Similarity transformation approach to identifiability analysis of nonlinear compartmental models. *Math. Biosci.*, **93**, 217–248.

Venzon,D. and Moolgavkar,S. (1988) A method for computing profile-likelihood-based confidence intervals. *Appl. Stat.*, **37**, 87–94.

White,L. *et al.* (2001) The structural identifiability and parameter estimation of a multispecies model for the transmission of mastitis in dairy cows. *Math. Biosci.*, **174**, 77–90.

Yao,K. *et al.* (2003) Modeling ethylene/butene copolymerization with multi-site catalysts: parameter estimability and experimental design. *Polym. React. Eng.*, **11**, 563–588.