

Research article

Open Access

Systematic identifiability testing for unambiguous mechanistic modeling – application to JAK-STAT, MAP kinase, and NF- κ B signaling pathway models

Tom Quaiser^{1,2} and Martin Mönnigmann*¹

Address: ¹Institute for Heat and Fuel Technology, Technical University Braunschweig, D-38106 Braunschweig, Germany and ²AVT – Process Systems Engineering, RWTH Aachen University, D-52064 Aachen, Germany

Email: Tom Quaiser - tom.quaiser@avt.rwth-aachen.de; Martin Mönnigmann* - m.moennigmann@tu-bs.de

* Corresponding author

Published: 9 May 2009

Received: 14 October 2008

BMC Systems Biology 2009, 3:50 doi:10.1186/1752-0509-3-50

Accepted: 9 May 2009

This article is available from: <http://www.biomedcentral.com/1752-0509/3/50>

© 2009 Quaiser and Mönnigmann; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: When creating mechanistic mathematical models for biological signaling processes it is tempting to include as many known biochemical interactions into one large model as possible. For the JAK-STAT, MAP kinase, and NF- κ B pathways a lot of biological insight is available, and as a consequence, large mathematical models have emerged. For large models the question arises whether unknown model parameters can uniquely be determined by parameter estimation from measured data. Systematic approaches to answering this question are indispensable since the uniqueness of model parameter values is essential for predictive mechanistic modeling.

Results: We propose an eigenvalue based method for efficiently testing identifiability of large ordinary differential models and compare this approach to three existing ones. The methods are benchmarked by applying them to models of the signaling pathways mentioned above. In all cases the eigenvalue method proposed here and the orthogonal method find the largest set of identifiable parameters, thus clearly outperforming the other approaches. The identifiability analysis shows that the pathway models are not identifiable, even under the strong assumption that all system state variables are measurable. We demonstrate how the results of the identifiability analysis can be used for model simplification.

Conclusion: While it has undoubtedly contributed to recent advances in systems biology, mechanistic modeling by itself does not guarantee unambiguous descriptions of biological processes. We show that some recent signal transduction pathway models have reached a level of detail that is not warranted. Rigorous identifiability tests reveal that even if highly idealized experiments could be carried out to measure all state variables of these signaling pathways, some unknown parameters could still not be estimated. The identifiability tests therefore show that the level of detail of the investigated models is too high *in principle*, not just because too little experimental information is available. We demonstrate how the proposed method can be combined with biological insight, however, to simplify these models.

Background

Several large and detailed mathematical models for signal transduction pathways exist in the literature. Lipniacki et al. [1] model the NF- κ B pathway using 15 state variables and 29 parameters. Yamada et al. [2] introduce a system of ordinary differential equations that describe the JAK-STAT pathway with 31 state variables and 52 parameters, and Schoeberl et al. [3] describe the EGF pathway with a model that comprises 103 variables and 98 parameters. Models of this kind can provide a concise and unambiguous representation even of very complex signaling pathways. However, since usually some of their parameters are not known, these models pose demanding parameter estimation problems. Two fundamental problems must be considered in this context: 1) the larger the number of unknown parameters in a model, the larger the amount of quantitative data necessary to determine meaningful values for these parameters. 2) Even if appropriate experimental data are available, model parameters may not be uniquely identifiable [4]. Ultimately, reliable predictive mathematical models can only be created after addressing these two problems.

Parameter identifiability has been paid little attention to in the recent systems biology literature. Exceptions exist, however. Swameye et al. [5] carefully check the identifiability of their JAK-STAT model. The authors propose a small model (four state variables and six parameters) and demonstrate that the model provides quantitatively reliable predictions despite its small size. In the field of HIV/Aids modeling differential algebraic techniques have been used to prove identifiability [6-8] for models of similar size. Other examples can be found in [9,10].

The problem of testing identifiability amounts to answering the following question: given a mathematical model of a system together with system input and output data, are the model parameters uniquely determined? The identifiability tests investigated here can be carried out before experimental data are available. To this end a model is used first to generate simulated data. Subsequently, it can be checked whether the model parameters are uniquely defined by the simulated data. Only if identifiability can be assured for the model and the simulated data it is reasonable to continue with lab experiments, identifiability tests with experimental data, and, eventually, parameter estimation. In the present paper we exclusively use simulated data.

Several notions of identifiability exist. We attempt a brief and informal summary of the essential ideas here. Stricter terminology is introduced in subsequent sections. Essentially, a model is called globally identifiable if a unique value can be found for each model parameter such that the model reproduces the measured or simulated output

data. If, in contrast, a finite number of points in the model parameter space can be found, for which the model reproduces the output data the model is called locally identifiable. Finally, if an infinite number of parameter values exists that reproduce the model input-output behavior, the model is considered to be unidentifiable. Independently of these three notions we need to distinguish between at-a-point identifiability on the one hand and structural identifiability on the other hand [11,12]. These two concepts distinguish two classes of methods for identifiability testing from one another. Methods for at-a-point identifiability testing can only be applied if candidate values for the model parameters are known a priori. This situation arises, for example, when parameter values for a model have already been published in the literature, as is the case for the signaling pathway models investigated here. In contrast, all mathematically and biologically possible parameter values must be considered as candidate values, if no information on the model parameter values is known a priori. In this case we speak of structural identifiability testing.

Several techniques exist for analyzing structural identifiability. These methods are based on power-series expansion [13], transfer function analysis [14], differential algebra [6,12,15], interval arithmetics [16], state isomorphisms [17-19] or semi-infinite programming [20,21]. These methods are, however, either restricted to linear models or to models with less than 10 states and parameters in the nonlinear case [22].

For large nonlinear models only methods for testing local at-a-point identifiability are feasible. A number of methods to test local at-a-point identifiability have been proposed in the literature. Some of these aim at determining the largest subset of identifiable parameters [23,24], other methods are tailored to finding the unidentifiable parameters [25,26], or to finding parameters that do not affect the input-output behavior of the model [27,28]. All these methods are based on the sensitivity matrix of the model responses (for details see section Methods for local at-a-point identifiability testing). In contrast to the approaches mentioned so far the method introduced in [29] does not depend on sensitivity information. This method repeatedly estimates parameters with randomly chosen start values and extracts dependencies between parameters with a statistical method. The data used for the parameter estimation steps is created by simulating the model at a nominal parameter point. The approach does not belong to the class of methods for at-a-point identifiability testing, since it is delocalized by using the above mentioned random multistart approach for parameter estimation. Consequently, the method is more rigorous than the discussed at-a-point identifiability tests, but not as rigorous as the structural methods.

The examples treated here turn out not to be locally at-a-point identifiable, therefore it is not reasonable to consider stricter concepts of identifiability. In the sequel the term identifiability will refer to local at-a-point identifiability if not noted otherwise.

We compare the methods for identifiability testing introduced by Yao et al. [24], Jacquez and Greif [25], and Degenring et al. [27], which for ease of reference we refer to as the orthogonal method, the correlation method, and the principle component analysis (PCA) based method, respectively. Specifically, we would like to establish which of these approaches is the method of choice for identifiability testing of large signal transduction pathway models. We omit the method published by Brun et al. [23] due to its combinatorial complexity. Similarly, the method proposed by Hengl et al. [29] is omitted, because it proved to be too computationally demanding for the examples treated here. The method was tested with default settings. When applied to the smallest of the models treated here, the wall clock computation time was about 15 hours in contrast to wall clock computation times of less than a minute for the methods compared here. When applied to the JAK-STAT model, the method proposed by Hengl et al. [29] did not finish within a week. It was therefore not included in the comparison. Finally, we note an interesting approach has only very recently been proposed by Chu and Hahn [30]. Since this approach solves a problem that is similar to, but after all different from, the identifiability tests addressed here, it is not included in the comparison. All results are compared to results found with our own method [31], which we refer to as the eigenvalue method for short. The eigenvalue method is an extension of an approach published by Vajda et al. [26]. We note the eigenvalue method has independently appeared in a recent paper by Schittkowski [32].

This paper has three contributions: 1) we reveal that three well established models of signal transduction are not identifiable, and demonstrate how results from identifiability studies can be used to simplify these models. 2) We suggest an efficient numerical method for identifiability testing of large nonlinear systems of ordinary differential equations in general, and 3) we compare our method to three previously published methods.

This paper is structured as follows. We start by reviewing the theory of identifiability and introduce the four methods for local at-a-point identifiability testing compared here. Subsequently, the three pathway models used in the case studies are introduced and each of the models is analyzed with each of the methods. Insight from this analysis is used to simplify the models.

Methods

In this section we will first introduce the mathematical system class and give a more concise definition of identifiability. We focus on local at-a-point identifiability and describe the four methods for testing local at-a-point identifiability compared here. Finally, we summarize the three signaling pathways and the corresponding models.

System Class

The concept of identifiability applies to a large system class. Here we treat models of the form

$$\begin{aligned} \dot{x}(t) &= f(x(t), p, u(t)), & x(0) &= x_0 \\ \gamma(t) &= h(x(t), p, u(t)), \end{aligned} \tag{1}$$

since many biological signaling pathways can be represented by systems of this form. In equation (1) $x \in \mathbb{R}^{n_x}$, $p \in \mathbb{R}^{n_p}$, $u \in \mathbb{R}^{n_u}$ and $\gamma \in \mathbb{R}^{n_\gamma}$ denote the state variables, the parameters, the inputs, and the outputs of the dynamical system, respectively. A biological parameter, for example a kinetic constant, corresponds to a component p_k of the parameter vector. The functions f and h map from an open subset $\mathcal{U}_x \times \mathcal{U}_p \times \mathcal{U}_u \subseteq \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \times \mathbb{R}^{n_u}$ onto \mathbb{R}^{n_x} and \mathbb{R}^{n_γ} , respectively, and are assumed to be smooth. Note both a lab experiment and a simulation are uniquely defined by the initial conditions $x(0) = x_0$ and the values of the inputs $u(t)$ from $t = 0$ to the final time $t = t_f$. Both an experiment and a simulation result in values of the outputs $\gamma(t)$ at successive points $t_1 \leq \dots \leq t_{n_t}$ in time

$$\gamma(t_i), \quad i = 1, \dots, n_t, \tag{2}$$

where n_t denotes the number of measurements or stored simulation results, respectively. In the present paper, output data of the form (2) is obtained from simulations with models that have been adjusted to experimental data by other authors before [1-3].

The solution of equation (1) that results for a particular choice of initial conditions x_0 , parameters p , and inputs $u(t)$ is denoted by

$$x(t) = \chi(t, p, x_0, u(t)). \tag{3}$$

The solution of equation (1) and its derivatives with respect to the parameters are computed with the integrator DDASPK [33]. The output behavior of the model is given by the response function

$$r(t, p, x_0) = h(\chi(t, p, x_0, u(t)), p, u(t)). \tag{4}$$

Concept of identifiability

Here we summarize the notions of identifiability as necessary for the paper (see [4] for a review). Assuming that the inputs $u(t)$, the initial conditions x_0 , and the measurement times t_i are given, the different notions of identifiability can be defined as follows.

Definition 1: The parameter p_k of the model (1) is called globally structurally identifiable if, for all admissible values $p^* \in \mathcal{U}_p$ and all $p' \in \mathcal{U}_p$,

$$\begin{aligned} r(0, p^*, x_0) &= r(0, p', x_0) \\ r(t_1, p^*, x_0) &= r(t_1, p', x_0) \\ &\vdots \\ r(t_{n_i}, p^*, x_0) &= r(t_{n_i}, p', x_0) \end{aligned} \tag{5}$$

implies

$$p_k^* = p'_k. \tag{6}$$

Definition 2: The parameter p_k of the model (1) is called locally structurally identifiable if, for all $p^* \in \mathcal{U}_p$, there exists a neighborhood $\mathcal{V}(p^*) \subseteq \mathcal{U}_p$ such that for all $p' \in \mathcal{V}(p^*)$, equations (5) imply equation (6).

The parameter vector p is called globally structurally or locally structurally identifiable if all its components p_k are globally structurally or locally structurally identifiable, respectively. Global and local at-a-point identifiability are defined as follows.

Definition 3: Let $p^* \in \mathcal{U}_p$ be a point in the parameter space of the model (1). The parameter p_k of the model (1) is called globally at-a-point identifiable at the point p^* if, for all $p' \in \mathcal{U}_p$, equations (5) imply equation (6).

Definition 4: Let $p^* \in \mathcal{U}_p$ be a point in the parameter space of the model (1). The parameter p_k of the model (1) is called locally at-a-point identifiable at the point p^* if there exists a neighborhood $\mathcal{V}(p^*) \subseteq \mathcal{U}_p$ such that for all $p' \in \mathcal{V}(p^*)$, equations (5) imply equation (6).

The following subsection will introduce the four methods for at-a-point identifiability testing compared in this work.

Methods for local at-a-point identifiability testing

The compared numerical methods for local at-a-point identifiability testing are based on the sensitivity of the model outputs at discrete time points $t_k, k \in \{1, \dots, n_i\}$, with respect to the parameters. The sensitivity information is stored in the $n_y n_t \times n_p$ dimensional sensitivity matrix S . S is a block matrix that consists of time dependent blocks $s(t_i)$ of size $n_y \times n_p$ [34]:

$$S = \begin{pmatrix} s(t_1) \\ s(t_2) \\ \vdots \\ s(t_{n_i}) \end{pmatrix}, s(t_i) = \begin{pmatrix} s_{11}(t_i) & \dots & s_{1n_p}(t_i) \\ \dots & \ddots & \dots \\ s_{n_y,1}(t_i) & \dots & s_{n_y n_p}(t_i) \end{pmatrix}. \tag{7}$$

The entries of $s(t_i)$ are called sensitivity coefficients. For a nominal parameter vector p^* , given x_0 , and fixed time t_k with $k \in \{1, \dots, n_i\}$, they are defined as

$$s_{ij}(t_k) = \frac{\partial r_i(t_k, p^*, x_0)}{\partial p_j}. \tag{8}$$

Essentially, the sensitivity coefficients describe how sensitive the system output is to changes in a single parameter. If the model output is highly sensitive to a perturbation in one parameter we can consider this parameter to be important for the system behavior. In contrast, a parameter that has no influence on the outputs is a candidate for an unidentifiable parameter. Linearly dependent columns of the sensitivity matrix imply that a change in the system outputs due to a change in one parameter, say p_j , can be compensated by changing some or all of the dependent parameters p_k with $k \neq j$. If dependencies exist, parameter estimation will fail or result in non-unique parameter values [26]. The correlation and orthogonal method are based on this idea.

For ease of presentation, we first give a detailed introduction to the eigenvalue method proposed here. The remaining methods are summarized briefly only, and the reader is referred to the appendix for more details. Since all methods analyze at-a-point identifiability, nominal values for the parameters are required. For the test cases treated here parameter values are available from the literature [1-3] and denoted by p^{lit} .

Eigenvalue method

The eigenvalue method is similar to the method published by Vajda et al. [26]. The eigenvalue method proposed here differs from the approach proposed by Vajda and coworkers in that a manual inspection of eigenvalues and eigenvectors is not necessary here. As a result, an automatic analysis of large models becomes feasible. We first describe the eigenvalue method and subsequently

describe differences to the method published by Vajda et al. [26] in more detail.

We consider the least squares parameter estimation problem that amounts to minimizing cost functions of the form

$$\phi(p) = \frac{1}{2} \sum_{i=1}^{n_y} \sum_{j=1}^{n_t} (r_i(t_j, p, x_0) - \gamma_i(t_j))^2 \quad (9)$$

with respect to p , where the $\gamma_i(t_j)$ are the measured or simulated data introduced in equation (2). In Gaussian approximation, the Hessian matrix H of equation (9) has the entries

$$H_{kl} = \sum_{i,j} \frac{\partial r_i(t_j, p, x_0)}{\partial p_k} \frac{\partial r_i(t_j, p, x_0)}{\partial p_l} \quad (10)$$

$$= (S^T S)_{kl}$$

for $k, l \in \{1, \dots, n_p\}$. As indicated in equation (10), H can be expressed in terms of the sensitivity matrix S introduced in equation (7). H is a symmetric matrix and therefore its eigenvalues are real. Furthermore, H is positive semi-definite.

In order to study the identifiability of the system in equation (1) for given x_0 , $u(t)$ and for nominal parameters $p^* \in \mathbb{R}^{n_p}$ we first solve equation (1) by numerical integration and subsequently calculate H as given by equation (10). We set p^* to parameter values taken from the literature p^{lit} in all calculations. All identifiability tests carried out in the present paper therefore amount to asking whether the investigated signaling models are identifiable at the literature parameter values p^{lit} . As pointed out already in the Background section, identifiability tests of this type clarify whether a given model is reasonably detailed even before experiments are carried out. In this sense, the identifiability of a model at nominal or literature values for its parameters is considered to be a necessary condition for its practical identifiability with measured data.

Let λ_j and w denote the j th eigenvalue and the corresponding eigenvector of H , respectively. Assume the eigenvalues to be ordered such that $\lambda^1 \leq \lambda^2 \leq \dots \leq \lambda^{n_p}$, assume the eigenvectors w to be normalized such that $w^T w = 1$, and assume p^* to minimize equation (9). Consider the change of ϕ when moving from p^* in a direction αw for some real α . Gaussian approximation yields

$$\begin{aligned} \phi(p^* + \alpha w) &= \phi(p^*) + \frac{1}{2} \alpha^2 w^T H w \\ &= \phi(p^*) + \frac{1}{2} \alpha^2 \lambda^j, \end{aligned} \quad (11)$$

where the linear order is zero, since p^* is a minimum by assumption. Since $Hw = \lambda w$ and $w^T w = 1$, the last equality in equation (11) holds. Equation (11) implies that ϕ does not change when moving from p^* to $p^* + \Delta p$ if $\Delta p = \alpha w$ for any eigenvector w with $\lambda = 0$ and any real α . In words, the directions w corresponding to $\lambda = 0$ are those directions in the parameter space along which the least squares cost function is constant. We call these directions degenerate for short. In the particular case of

$$w^j = (0, \dots, 0, 1, 0, \dots, 0)^T, \quad (12)$$

where the entry 1 is in position k , the model is not identifiable with respect to the k th component p_k^* . The approach proposed here will therefore remove this parameter from consecutive calculations by fixing $p_k^* = p_k^{lit}$. In general however, w will not be of the special form (12), but w will be a vector with more than one non-zero entry, therefore the choice of k is not obvious. In this case, we select a k such that $|u_k^j| = \max(|u_1^j|, |u_2^j|, \dots, |u_{n_p}^j|)$, and remove the k th component of p^* from the parameter estimation problem by fixing it to its literature value. In the examples treated below this choice of k turns out to be appropriate in all cases. In general, this might not be the case, however. There might be entries of u_l^j , $l \neq k$, with an absolute value $|u_l^j|$ close or equal to the maximal value $|u_k^j|$. We call such entries u_l^j and the corresponding parameters co-dominant, since together with u_k^j they dominate the degenerate direction. A simple example for a system with co-dominant parameters is given in Appendix A. We will discuss the issue of co-dominant parameters for the examples treated here in the Results section.

In practical applications the smallest eigenvalue will typically not be zero but close to zero. In this case an eigenvalue cut-off value $\epsilon \approx 0$, $\epsilon > 0$ needs to be specified. An eigenvalue $\lambda^1 < \epsilon$ is considered to be small enough to be treated as zero. Note $\lambda^j \geq 0$ for all j since H is positive semi-definite.

The proposed algorithm splits the set of parameter indices $\{1, \dots, n_p\}$ of a given model into two disjoint subsets which we denote by I and U . Upon termination, the sets I and $U = \{1, \dots, n_p\} - I$ contain the indices to the identifiable and the unidentifiable parameters, respectively. We denote the current number of elements in I by n_I and the current elements in I by $I = \{i_1, \dots, i_{n_I}\}$. The algorithm proceeds as follows.

1. Choose $\Delta t, t_f, \epsilon$. Set $p^* = p^{lit}$. Set $I = \{1, \dots, n_p\}$, $n_I = n_p$, and $U = \emptyset$.
2. If I is empty, stop. The model is not identifiable with respect to any parameter in this case.
3. Fix the parameters $p_k, k \in U$ to their literature values and consider only the $p_k, k \in I$ to be variable. Formally, this corresponds to setting p to $(p_{i_1}, \dots, p_{i_{n_I}})^T$ and treating all remaining parameters $p_k, k \in U$ as fixed numbers.
4. Calculate $\phi(p^*)$ according to equation (9). Calculate the Hessian matrix of ϕ with respect to the parameters p_k with $k \in I$ and evaluate it at $p = p^*$. Calculate the eigenvalues λ^j and corresponding eigenvectors w^j of the Hessian matrix. Assume the eigenvalues to be ordered such that $\lambda^1 \leq \lambda^2 \leq \dots \leq \lambda^{n_I}$. Assume the eigenvectors to be normalized.
5. If $\lambda^1 \geq \epsilon$, stop. The model is identifiable with respect to the parameters p_k for all $k \in I$.
6. If $\lambda^1 < \epsilon$, select k such that $|u_k^1| = \max(|u_1^1|, |u_2^1|, \dots, |u_{n_I}^1|)$. Remove k from the set I , add k to the set U , set n_I to the current number of elements in I , and return to step 2.

The order in which parameters are removed from I determines the ranking of parameters from least identifiable to most identifiable.

The method is similar to the approach introduced by Vajda et al. [26]. These authors also analyze the eigenvectors that correspond to small eigenvalues of the Hessian matrix, but they focus on the dependencies between eigenvectors that arise due to special parameter combinations of the form p_1/p_2 or $p_1 \cdot p_2$ in the model. In a two step procedure Vajda et al. first decide which parameters to

lump together (for example, $p_{new} = p_1 \cdot p_2$) and subsequently recalculate the eigenvalues and eigenvectors for the system with the new lumped parameters. These two steps are repeated until the smallest eigenvalue is sufficiently large. The lumping step requires manual inspection of the eigenvalues and eigenvectors of the Hessian matrix. While the approach proposed by Vajda et al. worked very well for their example with 5 parameters, it becomes infeasible for models with considerably more parameters. The examples treated in the next section demonstrate that our approach, in contrast, can be applied to models with up to at least a hundred parameters. Moreover, our approach can be carried out automatically, while the approach suggested by Vajda et al. was never intended for this purpose.

Correlation method

The correlation method was first introduced by Jacquez and Greif [25] who compared identifiability results achieved with the correlation method to analytical results obtained with a transfer function method (see [14] for a review). Jacquez and Greif [25] consider only small linear compartmental models with up to 3 states and 5 parameters. In most of the examples the correlation and the transfer function method were in agreement.

In the systems biology literature the correlation method was applied by Zak et al. [9] to investigate identifiability of a large genetic regulatory network (nonlinear, 44 states and 97 parameters). More recently, Rodriguez-Fernandez et al. [10] embedded the correlation method into their framework for robust parameter estimation in order to exclude unidentifiable parameters from parameter estimation.

The central idea of the correlation method is to find unidentifiable parameters by investigating the linear dependence of the columns of S (for details see Appendix B). The correlation method approximates linear dependence by calculating the sample correlation [35] of two columns $S_{\cdot i}, S_{\cdot j} \in \mathbb{R}^{n_y n_t}$ of S . The sample correlation is given by

$$\text{corr}(S_{\cdot i}, S_{\cdot j}) = \frac{\text{cov}(S_{\cdot i}, S_{\cdot j})}{\sigma(S_{\cdot i})\sigma(S_{\cdot j})}, \tag{13}$$

where

$$\text{cov}(S_{\cdot i}, S_{\cdot j}) = \frac{1}{n_y n_t - 1} \sum_{k=1}^{n_y n_t} (S_{ki} - \bar{S}_{\cdot i}) \cdot (S_{kj} - \bar{S}_{\cdot j}), \tag{14}$$

$$\sigma(S_{.i}) = \sqrt{\frac{1}{n_y n_t - 1} \sum_{k=1}^{n_y n_t} (S_{ki} - \bar{S}_{.i})^2}, \tag{15}$$

$$\text{and } \bar{S}_{.i} = \frac{1}{n_y n_t} \sum_{k=1}^{n_y n_t} S_{ki}. \tag{16}$$

In equations (13)–(16) $corr(S_{.i}, S_{.j})$, $cov(S_{.i}, S_{.j})$, $\sigma(S_{.i})$, and $\bar{S}_{.i}$ denote the sample correlation between $S_{.i}$ and $S_{.j}$, the sample covariance between $S_{.i}$ and $S_{.j}$, the sample standard deviation of $S_{.i}$, and the mean of the entries of $S_{.i}$, respectively.

Two linearly dependent columns $S_{.i}$ and $S_{.j}$ of the sensitivity matrix give rise to an absolute value of the correlation $corr(S_{.i}, S_{.j}) = 1$. In the examples treated here, none of the parameters exhibit a correlation of exactly ± 1 , however. In the numerical implementation of the method, two columns of S are considered correlated if the absolute value of their correlation is equal to or larger than $1 - \epsilon_c$, where $\epsilon_c \in [0, 1]$ is a parameter of the algorithm. We stress that while ϵ_c is a tuning parameter, the comparison of the identifiability methods is not affected by the choice of ϵ_c . As explained below, we systematically vary ϵ_c over its entire range $0 \leq \epsilon_c \leq 1$ in the comparison.

When applying the correlation method to the pathway model examples, two problems arise regularly that have not been discussed by Jacquez and Greif [25]. For one, the method detects pairs of correlated parameters, but there is no criterion which one of the parameters from a pair to consider unidentifiable. Moreover, if more than one pair of correlated parameters is detected, there is no criterion to choose among the pairs.

In order to mitigate these ambiguities we introduce the total correlation

$$c_i^{tot}(L) = \sum_{j \in L, j \neq i} corr^*(S_{.i}, S_{.j}), \tag{17}$$

where

$$corr^*(S_{.i}, S_{.j}) = \begin{cases} |corr(S_{.i}, S_{.j})|, & \text{if } |corr(S_{.i}, S_{.j})| \geq 1 - \epsilon_c \\ 0, & \text{else.} \end{cases} \tag{18}$$

and L denotes the set of indices of those parameters that have not been found to be unidentifiable in any previous iteration. We select the parameter with the highest total correlation for removal. Some cases remain, however, in

which two parameters have the same total correlation. In these cases we pick one parameter at random.

In order to be able to compare results to those of the other methods, we use the correlation method to rank model parameters from least identifiable to most identifiable. More precisely, we initially set $\epsilon_c = 0$ and $L = \{1 \dots n_p\}$ and calculate the total correlations (17) for all parameters in L . If parameters with nonzero total correlations exist we select the parameter with highest total correlation to be unidentifiable and remove its index from L . If no such parameter exists we increase ϵ_c until nonzero total correlations occur. The order in which parameters are removed from L creates a ranking of parameters from least to most identifiable. The algorithm is described in detail in Appendix B. Based on the ranking of parameters the method is compared to the other three approaches. The comparison is explained in detail in the subsection entitled Method comparison.

Principal component analysis (PCA) based method

Degenring et al. [27] introduce three criteria based on PCA that rank the influence of parameters on model output. Parameters that do not affect the model outputs are unidentifiable by definition and can be removed from the model equations. Degenring et al. successfully use their approach to simplify a complex metabolism model of Escherichia coli K12. A complete description how to get a full ranking using the three criteria can be found in [28]. A pseudo code description of the method can be found in Appendix C.

The PCA based method differs from the other methods described here in that it does not use the sensitivity matrix S as defined in equation (7) but a truncated matrix which we refer to by \tilde{S}^i . This matrix \tilde{S}^i is defined as

$$\tilde{S}^i = \begin{pmatrix} s_{i1}(t_1) & \dots & s_{in_p}(t_1) \\ \vdots & \ddots & \vdots \\ s_{i1}(t_{n_i}) & \dots & s_{in_p}(t_{n_i}) \end{pmatrix},$$

where $i \in \{1, \dots, n_y\}$. While the sensitivity matrix S as defined in equation (7) describes all responses in one matrix, \tilde{S}^i is created for each response r_i separately.

Three PCA based criteria are applied to each of the n_y matrices $\tilde{S}^{iT} \tilde{S}^i$. As a result $3n_y$ parameter rankings are created, which we represent by $3n_y$ lists

$$M_j^i = (m_{j,1}^i, m_{j,2}^i, \dots, m_{j,n_p}^i) \tag{19}$$

of integers $m_{j,k}^i \in \{1, \dots, n_p\}$, where $j \in \{1, 2, 3\}$ and $i \in \{1, \dots, n_y\}$. The first ranking position $m_{j,1}^i$ and the last ranking position m_{j,n_p}^i contain the index of the parameter with the lowest and the highest influence on the model output r_i , respectively. The lower index $j \in \{1, 2, 3\}$ denotes the criterion by which the ranking is calculated. We will explain one of the criteria in more detail in order to give the reader an idea of the method. The remaining two criteria are summarized in Appendix C. All $3n_y$ rankings are integrated into one ranking by applying a strategy described in [28], which we briefly summarize below.

Let, $\tilde{\lambda}^{i,j}$ and $\tilde{u}^{i,j}$ denote the j th eigenvalue and corresponding eigenvector of $\tilde{S}^{iT}\tilde{S}^i$, respectively, and assume the eigenvalues to be ordered such that $\tilde{\lambda}^{i,1} \leq \tilde{\lambda}^{i,2} \leq \dots \leq \tilde{\lambda}^{i,n_p}$. For each $i \in \{1, \dots, n_y\}$ evaluate the first criterion as follows.

1. Set $L_1 = \{1, \dots, n_p\}$.
2. For q from 1 to n_p do:
 - Find r_l such that $|u_{r_l}^{i,q}| = \max_{j \in L_q} (|\tilde{u}_j^{i,q}|)$.
 - Set $m_{1,q}^i = r_l$.
 - Set $L_{q+1} = L_q - \{r_l\}$.

The first criterion loops over the eigenvectors starting with the eigenvector that belongs to the smallest eigenvalue. It identifies the largest absolute entry of each eigenvector and selects the corresponding parameter for removal, if this parameter has not been selected yet. The index of the k th parameter selected for removal is ranked at position k , where $k \in \{1, \dots, n_p\}$.

The remaining two criteria of the PCA based method are similar but differ with respect to the process for choosing r_l . For brevity we omit details and refer the reader to Appendix C. A final ranking J that incorporates all $3n_y$ previously calculated rankings is obtained as follows. Set $N_q = \emptyset$, for all $q = 1, \dots, n_p$. Set J to an empty list.

For q from 1 to n_p do:

- For all $j \in \{1, 2, 3\}$ and all $i \in \{1, \dots, n_y\}$ set $W_j^i(q) = \{m_{j,k}^i \mid k \in \{1, \dots, q\}\}$.

- Set $\bullet N_q = \bigcap_{\substack{1 \leq j \leq 3 \\ 1 \leq i \leq n_y}} W_j^i(q) - N_{q-1} \dots - N_1$.

- Only if $N_q \neq \emptyset$ append the N_q to the list J .

First note that an iteration q might exist, in which no N_q is appended to J . Therefore the final number n_j of elements in J is not necessarily equal to but might be smaller than n_p . Further note that entries of J do not necessarily correspond to indices of single parameters but to sets of parameter indices. Therefore, a ranking position J_i might contain several parameter indices an internal ranking of which cannot be obtained. This ambiguity hampers not only the interpretation of the ranking result but also the comparison with the rankings produced by the other methods. We describe a strategy to deal with the latter problem in section entitled Method comparison.

Orthogonal method

The orthogonal method was developed by Yao et al. [24] to analyze parameter identifiability of an Ethylene/Butene copolymerization model with 50 parameters. In the field of systems biology the method has been applied in a framework for model identification [36] and in the identifiability analysis of a NF- κ B model [37].

We summarize the concept of the method and refer the reader to Appendix D for details. The method iterates over the columns $S_{\cdot k}$ of the sensitivity matrix S , with $k \in \{1, \dots, n_p\}$. Those columns of S that correspond to identifiable parameters are collected in a matrix X^q where q is the iteration counter. The algorithm starts by selecting the column of S with highest sum of squares in the first iteration. In iteration $q + 1$, q columns of S have been selected. In the order of their selection these columns form the matrix X^q . The next step essentially amounts to selecting the column $S_{\cdot k}$ that exhibits the highest independence to the vector space V spanned by the columns of X^q . More precisely, an orthogonal projection $S_{\cdot k}^{proj}$ of $S_{\cdot k}$ onto V is calculated and $S_{\cdot k}^\perp = S_{\cdot k} - S_{\cdot k}^{proj}$ is interpreted as the shortest connection from V to $S_{\cdot k}$. The squared length of $S_{\cdot k}^\perp$ is used as measure of independence. If the length of $S_{\cdot k}^\perp$ is near zero, $S_{\cdot k}$ is nearly linearly dependent to the columns of X^q . Conversely, a large value of $\|S_{\cdot k}^\perp\|_2^2$ indicates that parameter p_k is linearly independent to the columns of X^q . In the orthogonal method the parameter p_l where

$\|S_{\cdot j}^\perp\|_2^2 = \max_k (\|S_{\cdot k}^\perp\|_2^2)$ is selected as the next identifiable parameter.

Essentially, the orthogonal method finds those columns of S that are as independent as possible. The quantity $\|S_{\cdot k}^\perp\|_2^2$ can be interpreted as a measure of independence that has been freed from dependencies on previously selected parameters. Therefore not only the influence of a parameter on the outputs, but also the dependencies between the parameters are accounted for by using $\|S_{\cdot k}^\perp\|_2^2$ as a measure. A visualization of the orthogonal method can be found in Appendix D.

Yao et al. [24] stop the selection process once the length of the maximal $S_{\cdot k}^\perp$ drops below a cut-off value $\epsilon \approx 0$, the choice of which is rather arbitrary. In the present paper we do not apply this cut-off criterion but merely rank all parameters from most to least identifiable where the terms most identifiable and least identifiable are used in the same sense as in the correlation method. A comparison to the results of the other methods is therefore not affected by the choice of ϵ . Our approach to comparing the methods is explained in the section entitled Method comparison.

Method comparison

The central idea behind our comparison of the methods is the equivalence between the local identifiability of a model and the local positive definiteness of the Hessian matrix (10) at a minimum of the least squares parameter cost function (9) for this model. Due to this theoretical result, which has been established by Grewal and Glover [38], we can test local identifiability at a point in the model parameter space by testing the Hessian matrix (10) for positive definiteness at this point. Since the Hessian matrix is positive definite if and only if all its eigenvalues are strictly positive, we consider a model to be identifiable if the eigenvalues are bounded below by a small strictly positive number. Note the eigenvalue method proposed here is based on the same idea in that it selects those parameters to be unidentifiable that cause eigenvalues smaller than .

By virtue of the relation between local at-a-point identifiability and the eigenvalues of the Hessian matrix we can compare all four identifiability testing methods to one another with the same criterion and an unique parameter value. For a given model we first create a ranking of the parameters from least to most identifiable with each method as described in the previous four sections. For

each parameter ranking we then fix the parameter considered to be the least identifiable to its literature value, calculate the Hessian matrix (10) with respect to the remaining parameters, and determine the smallest eigenvalue. In all following steps we additionally fix the next least identifiable parameter and recalculate the Hessian and smallest eigenvalue. This process is carried out until the smallest eigenvalue exceeds . The parameters that need to be fixed in order for the smallest eigenvalue to exceed are considered to be the unidentifiable parameters. Clearly, the smaller the set of unidentifiable parameters, the larger the set of identifiable parameters or, equivalently, the larger the number of parameters for which the least squares parameter estimation problem can be solved. In this sense we consider the method to be the best one that results in the smallest number of unidentifiable parameters in our comparison.

Note in contrast to the other methods, the ranking created with the PCA based method does in general not contain one but several parameters. In such a case we fix all parameters ranked at the same position at once and reevaluate the smallest eigenvalue of the Hessian matrix. The last ranking position J_{n_j} created by the PCA based method contains the indices of all parameters that have not been ranked previously. Fixing of parameters p_k , with $k \in J_{n_j}$, would therefore not leave any parameter unfixed. For this reason we consider the PCA method to have failed, if fixing of all parameters with indices in $\bigcup_{1 \leq i \leq n_j - 1} J_i$ does not lead to a value of the smallest eigenvalue larger than or equal to .

Pathway model description

Having introduced the methods to be compared, we briefly summarize the three pathways and the corresponding models used for the identifiability studies.

JAK-STAT pathway

The Janus Kinase (JAK) – signal transducer and activator of transcription (STAT) pathway is triggered by cytokines and growth factors. The pathway has an impact on the expression of genes that regulate diverse cellular processes, such as cellular proliferation, differentiation, and apoptosis. A detailed model of the Interferon-induced JAK-STAT pathway is given by Yamada et al. [2]. This model describes the following signal transducing steps. Upon ligand binding the receptor dimerizes thus triggering the activation of receptor-associated JAK. Activated JAK immediately phosphorylates the receptor complex, enabling the binding and subsequent phosphorylation of cytoplasmic STAT. Phosphorylated STAT can dimerize,

and finally, dimerized STAT is imported into the nucleus. Here it activates target genes, one of which is the suppressor of cytokine signaling 1 (SOCS1), an important mediator of negative feedback.

EGF induced MAP kinase pathway

Mitogen-activated protein (MAP) kinases are involved in mitosis and the regulation of cellular proliferation. The Epidermal Growth Factor (EGF)-induced MAP kinase pathway is centered on a three kinase cascade that terminates with the dual phosphorylation of the third, so-called MAP kinase. EGF stimulation leads to dimerization of the EGF receptor. The dimerized EGF receptor triggers the phosphorylation of Raf, the first kinase of the cascade. Raf, in turn, phosphorylates the mitogen extracellular kinase, MEK, the second kinase in the cascade, which finally phosphorylates the MAP kinase (extracellular signal-related kinase, ERK). Phosphorylated ERK regulates several proteins and nuclear transcription factors and controls the expression of target genes. The model published by Schoeberl et al. [3] additionally includes the internalization of EGF receptor and two further modules: an SH2-domain containing protein (SHC) dependent module and an SHC independent one.

NF-κB pathway

Nuclear Factor-κB (NF-κB) is a transcription factor, which regulates genes involved in inflammation, immune response, cell proliferation, and apoptosis. In the unstimulated cell NF-κB is captured in the cytoplasm by the protein inhibitor of NF-κB (IκBα). Upon stimulation by pathway activating signals such as the Tumor Necrosis Factor (TNF), the IκB kinase (IKK) is activated. Activated IKK phosphorylates IκBα thus inducing its degradation. As a consequence NF-κB is released, translocates into the nucleus and regulates effector genes. NF-κB induces its own downregulation by inducing the production of two proteins: 1) IκBα that inhibits NF-κB by relocating it to the cytoplasm and 2) the zink-finger protein A20 that represses IKK and consequently indirectly inhibits NF-κB.

Results

In this section the results of the identifiability analysis are reported. All simulations are carried out with literature values p^{lit} for the model parameters, which are taken from the original publications of the models [1-3]. We assume that all state variables are available as outputs, i.e. $\gamma(t) = x(t)$ and $n_\gamma = n_x$. Clearly, it is far from realistic to assume all state variables could be measured in an actual experiment for any of the signaling pathways. Since even under this strong assumption the models turn out not to be identifiable, however, it is not reasonable to investigate the identifiability with fewer outputs. In fact our identifiability results show that the models are overparameterized even under the very strong assumption of all state variables

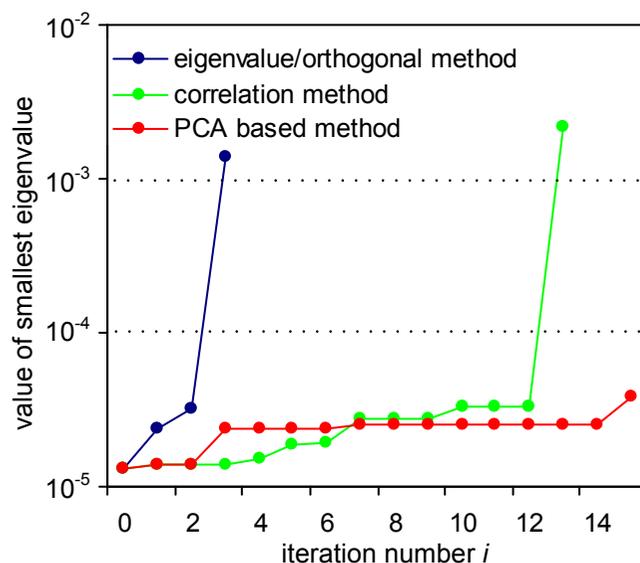


Figure 1
Results of the methods applied to the JAK-STAT model. The graph illustrates how the value of the smallest eigenvalue changes, when one parameter is fixed in each iteration. The y-axis shows the value of the smallest eigenvalue of the Hessian matrix *H*. Parameters are fixed until the value of the smallest eigenvalue is larger than $= 10^{-4}$.

being measurable. The algorithmic parameter is set to $= 10^{-4}$. While this value is motivated by our experience with numerical parameter estimation problems, we admit that it is somewhat arbitrary. Note, however, the comparison of the identifiability testing methods is consistent in that the same value $= 10^{-4}$ is used throughout. Output values $\gamma(t)$ are recorded from the simulations at equidistant points in time. The JAK-STAT model is simulated for 28800 seconds (8 hours) and the response function is recorded with a time step of $\Delta t = 360$ s. The corresponding values for the MAP kinase and NF-κB models are 3600 s (1 hour), $\Delta t = 120$ s, and 21600 s (6 hours), $\Delta t = 360$ s, respectively. These values are not critical. Similar results are obtained with other values of t_f and Δt (data not shown).

JAK-STAT pathway analysis

Figure 1 summarizes the results of the analysis of the JAK-STAT model. The figure shows the value of the smallest eigenvalue as a function of the number of fixed parameters. Before discussing the parameters that are considered to be unidentifiable by the various methods in detail, some results can already be inferred from Figure 1. In iteration $i = 0$, i.e. before any parameter is fixed in the model, the smallest eigenvalue λ^1 equals $1.3 \cdot 10^{-5}$. This implies the model is not identifiable even if all state variables $x(t)$ are assumed to be available. The smallest eigenvalue exceeds $= 10^{-4}$ after the first three parameters regarded as

unidentifiable by the eigenvalue and the orthogonal method are fixed (specifically, kb7, kb30, and kb5). In contrast, the first 13 parameters ranked by the correlation method need to be fixed to obtain a smallest eigenvalue larger than the threshold $= 10^{-4}$. Since the eigenvalue and orthogonal method find an identifiable model by fixing a much smaller subset of the parameters, we conclude these methods outperform the correlation method in this particular case. The PCA based method fails for this example.

Some of the results generated by the four methods are worth inspecting more closely. In the first and second iteration of the eigenvalue method there exists more than one dominant contribution to the degenerate direction, cf. Table 1. In the first iteration there are two co-dominant entries (last column in Table 1). The corresponding parameters are kb5 and kb30. In the second iteration there exists one co-dominant entry, which corresponds to the parameter kb5. In the last iteration no co-dominant entries exist and the parameter kb5 is solely responsible for the degenerate direction. It turns out that selecting the parameter corresponding to the maximum entry of $|u^1|$ as done here is sufficient to find a minimal set of identifiable parameters. Interestingly, we can see in this example that after three iterations our method finds the parameters corresponding to the co-dominant entry of all previous iterations.

The correlation method indicates that 13 parameters need to be fixed in order for the model to be locally identifiable (Figure 1). The sets of parameters removed by the eigenvalue and orthogonal method have only kb7 in common with the parameters selected by the correlation method. In the correlation method, the removal of parameter kb7 occurs in the last iteration and results in a clear increase of the smallest eigenvalue by a factor of about 100 (from $3.3 \cdot 10^{-5}$ to $2.2 \cdot 10^{-3}$, see Figure 1). Surprisingly, previous iterations of the correlation method hardly have an effect on the minimum eigenvalue.

Table 2 shows ranking J as produced by the PCA based method for the JAK-STAT model. Each ranking position J_i corresponds to a set of parameter indices, the size of which is shown in the table. The major problem of the

PCA based method becomes apparent here. Ranking J has 5 positions. Fixing the 15 parameters of the first four positions of J is not sufficient to get an identifiable model. Fixing the set of 36 parameters at the last ranking position would not leave any parameter unfixed. Therefore the method has failed in finding a set of identifiable parameters. The ambiguity of the method is apparent from positions four and five of the ranking. The sets J_4 and J_5 contain 8 and 36 parameters, respectively, which cannot be ranked internally by the PCA based method. It is interesting to note that the first ranking position of J is not filled until iteration 45 (data not shown), six iterations before the last iteration. No nonempty intersection of the first 44 positions of all rankings M_j^i with $j \in \{1, 2, 3\}$ and $i \in \{1, \dots, n_y\}$ exists. This indicates that large parts of the three different rankings $M_j^i, j = 1, 2, 3$ do not agree with respect to the order of parameters.

The parameters removed by the eigenvalue and orthogonal methods point the way to possible simplifications of the pathway models. The parameter kb7 describes the dissociation of phosphorylated STAT from the activated receptor complex. This reaction, however, is not the dissociation that immediately follows the phosphorylation of STAT by the activated receptor (Figure 2, kf6). Apart from the association of unphosphorylated STAT to the receptor (Figure 2, kf5), the phosphorylation of STAT by the activated receptor and the subsequent dissociation of phosphorylated STAT from the activated receptor (Figure 2, kf6), the model also permits the reassociation of already phosphorylated STAT from the cytoplasm to the activated receptor (Figure 2, kf7). The dissociation described by parameter kb7 is the reverse reaction to the reassociation described by kf7. We claim these steps are less important, since the key event is the phosphorylation of cytoplasmic STAT by the activated receptor. Removing kb7 alone creates a sink for phosphorylated STAT and activated receptor, since the dissociation of both species does not exist anymore. Therefore the model can only be simplified by removing both the reaction kb7 is involved in and the reaction kf7 is involved in. The main effects of these reac-

Table 1: Results of the eigenvalue method for the JAK-STAT model

iteration	removed parameter	max. entry of $ u^1 $	co-dominant entries of $ u^1 $
1	kb7	0.725	0.56, 0.4
2	kb30	0.728	0.68
3	kb5	0.997	-

Table 2: Ranking created by the PCA based method for the JAK-STAT model

ranking position i	size of J_i
1	1
2	2
3	4
4	8
5	36

tions are 1) to reduce the amount of free phosphorylated STAT and 2) to reduce the amount of free activated receptor. After removing both reactions, effect 1) can be imitated by appropriately decreasing those rate constants involved in reactions with free phosphorylated STAT. Effect 2) can be compensated by appropriately decreasing those rate constants that are involved in reactions with free activated receptor.

The second fixed parameter, kb30, corresponds to the rate of dissociation of STAT from the complex formed by SOCS1 and the activated receptor complex. In order to get a simplified model that is identifiable, we may remove this step. This is justified, because the key function of

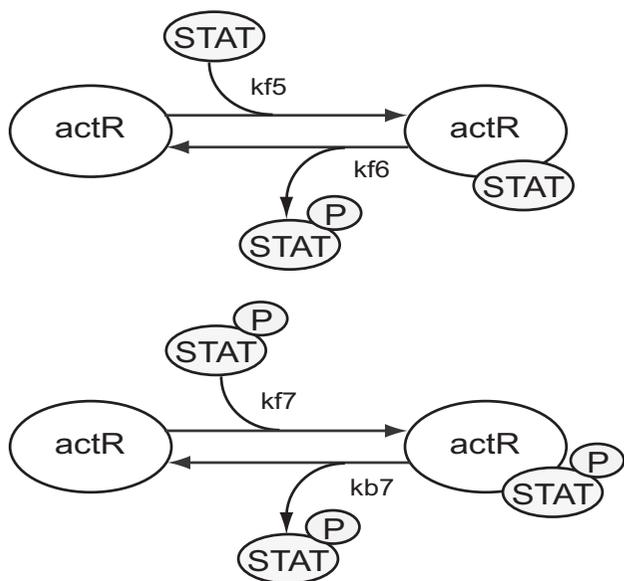


Figure 2
Part of the JAK-STAT model to be simplified. The activated receptor is denoted by actR. The encircled P indicates that STAT is phosphorylated.

SOCS1 is not its binding to STAT, but its inhibition of the activated receptor complex [39]. The latter is still ensured in the simplified model.

The third fixed parameter, kb5, describes the dissociation of receptor-bound unphosphorylated STAT. Assuming that STAT binds irreversibly to the activated receptor, the model can be simplified. In this scenario, STAT dissociates from the receptor only after STAT phosphorylation has occurred. Due to the presence of PPX, which can dephosphorylate cytoplasmic STAT, this change to the model need not result in an increase of the concentration of phosphorylated STAT.

We stress that the fact that kb7, kb30, and kb5 render the model unidentifiable does not imply the corresponding reactions do not exist. Our results imply, in contrast, that the model is too complex in the sense that these parameters cannot be determined by model based parameter estimation. Note that this result holds true even if we assume data for all the state variables to be available.

Map Kinase pathway analysis

Before any parameter of the MAP kinase model is fixed, the smallest eigenvalue of the Hessian matrix (10) equals $3.3 \cdot 10^{-13}$. This implies the model is not identifiable. The eigenvalue and the orthogonal method again outperform the correlation method in that they find a significantly smaller set of parameters to be responsible for the unidentifiability of the model. The PCA based method does not lead to an identifiable method.

In the eigenvalue method the maximal components of the respective eigenvectors are near 1 (data not shown) in both iterations. As a result, two parameters, k15 and k1, can unambiguously be fixed in the first and second iteration, respectively.

The correlation method needs to fix 36 parameters to arrive at an identifiable model. Fixing the first 35 parameters does not significantly increase the smallest eigenvalue (Figure 3). Only when the last parameter is fixed a dramatic increase of the smallest eigenvalue from about $4.0 \cdot 10^{-13}$ to $4.3 \cdot 10^2$ results. The parameter k15 is also removed by the eigenvalue and the orthogonal method, which indicates the importance of k15 for identifiability. The other parameter removed by the eigenvalue method, k1, is selected in the 2nd iteration of the correlation method. Therefore, the set of 36 parameters found by the correlation method is a superset of the 2 parameters selected by the eigenvalue and the orthogonal method. Closer inspection of the first iteration in the correlation method reveals a high maximal total correlation, c_i^{tot} , of

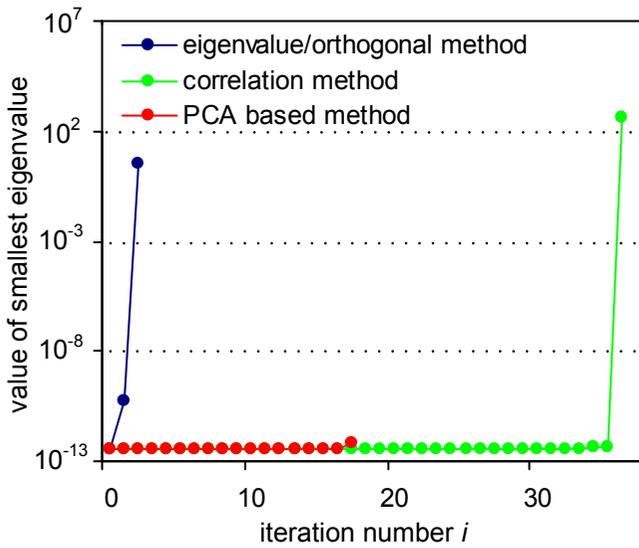


Figure 3
Results of the methods applied to the MAP kinase model. Until iteration number 17 the green curve lies below the red curve.

7.98. This high value is caused by couplings of the selected parameter with eight other parameters. Despite this high total correlation value of 7.98, the removal of the corresponding parameter does not cause a significant increase of the smallest eigenvalue. Just as in the previous example, removing parameters with high correlations does not necessarily improve identifiability.

The ranking found by the PCA based method has five positions (Table 3). When the 17 parameters that correspond to $J_1, J_2, J_3,$ and J_4 are fixed, the model remains unidentifiable. Fixing the parameters that correspond to J_5 is not reasonable, since doing so would not leave any parameters unfixed. Due to this result, we consider the PCA based method to have failed.

Table 3: Ranking created by the PCA based method for the MAP kinase model

ranking position i	size of J_i
1	1
2	1
3	3
4	12
5	78

In summary, the eigenvalue method and the orthogonal method identify the smallest set of parameters that needs to be fixed in order to create an identifiable model. The first parameter, k_{15} , is involved in internalization processes represented by more than twenty reactions [3]. Comprehensive changes to this part of the model that are necessary to ensure identifiability cannot be addressed here. The second parameter, k_1 , describes the binding of EGF to its receptor. This essential step cannot be simplified. However, experimental data on this reaction exists in the literature [40,41]. Our analysis suggests using this kind of independent information to fix the value of k_1 before attempting to determine the remaining parameters by parameter estimation.

NF- κ B pathway analysis

For the NF- κ B pathway model, the smallest eigenvalue $\lambda^1 = 1.9 \cdot 10^{-5}$ results if no parameters are fixed (Figure 4). This indicates that the model is not identifiable for the published parameter values [1]. The eigenvalue and the orthogonal method again find the same 3 unidentifiable parameters, $c_{4a}, c_4,$ and t_1 . The correlation method selects a considerably larger set of 21 parameters. The PCA based method fails.

In the first three iterations of the eigenvalue method, the largest components of u^1 have absolute values close to one (data not shown). The method therefore selects these parameters unambiguously.

The correlation method needs to fix 21 parameters until the smallest eigenvalue is greater than . The last parame-

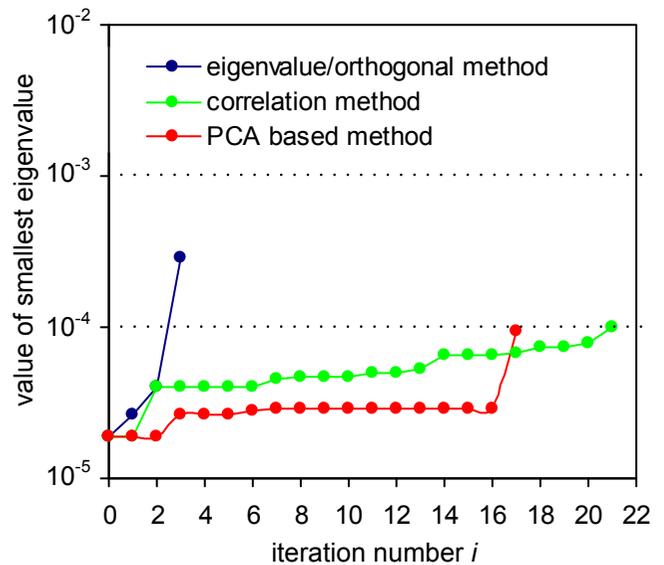


Figure 4
Results of the methods applied to the NF- κ B model.

ter, the fixing of which leads to an identifiable model (Table 4), has been selected with a value of $c_c = 0.55$. This is not reasonable anymore, since two columns of S with a correlation of $1 - c_c = 0.45$ can hardly be considered correlated. Therefore we conclude that the correlation method fails for this example.

The PCA based method creates a ranking J with six positions (Table 5). Fixing the parameters that correspond to J_1 to J_5 does not result in an identifiable model. Fixing the parameters corresponding to J_6 will not leave any parameters unfixed. Therefore the method has failed in producing an identifiable model. After fixing all parameters from J_1 to J_5 the resulting smallest eigenvalue $\lambda^1 = 9.5 \cdot 10^{-5}$ is close to 0. Having a look at the parameters in the ranking J we can see that J_1 contains parameter $t1$, and J_5 contains parameter $c4$. Both parameters are also found by the eigenvalue and the orthogonal method.

The three parameters selected for removal by the eigenvalue and the orthogonal method indicate how to simplify the model. The first two parameters, $c4a$ and $c4$, describe the only two translation steps in the model: the translation of $I\kappa B\alpha$ mRNA and A20 mRNA, respectively. By lumping together transcription and translation into one protein synthesis step the two unidentifiable parameters $c4a$ and $c4$ can be removed from the model. The third parameter selected by the eigenvalue and orthogonal method, $t1$, describes the dissociation of activated IKK

Table 4: Results obtained with the correlation method for the NF- κ B model

rank	removed parameter*	$ u_k^1 \cdot 10^{-05}$	c_i^{tot}
1	c1(0.99)	1.90	1.000
2	c1a(0.99)	3.98	1.000
3	c3(0.99)	3.99	0.998
4	a3(0.97)	4.06	1.958
5	c3a(0.96)	4.07	0.967
6	c3c(0.93)	4.07	1.874
7	c4a(0.93)	4.61	0.938
8	c4(0.92)	4.62	0.930
9	c1c(0.92)	4.62	0.927
10	kdeg(0.84)	4.65	0.849
11	c2a(0.81)	4.96	0.819
12	c2(0.80)	4.96	0.805
13	c5a(0.74)	5.34	0.743
14	a2(0.72)	6.54	0.727
15	a1(0.64)	6.64	0.643
16	i1(0.56)	6.64	0.565
17	c6a(0.50)	6.82	0.506
18	k2(0.50)	7.31	0.503
19	e2a(0.46)	7.42	0.470
20	e1a(0.45)	7.79	0.459
21	kv(0.45)	10.08	0.455

* The value of $1 - c_c$ is given in brackets. See the text for a discussion.

Table 5: Ranking created by the PCA based method for the NF- κ B model

ranking position i	size of J_i
1	1
2	2
3	3
4	1
5	10
6	12

from $I\kappa B\alpha$. Since this reaction is a key step of the model the only option for simplifying this part of the model is to remove IKK entirely. As a consequence A20 becomes obsolete, since the only function of A20 is the regulation of IKK. If we also remove A20, the resulting simplified model closely resembles the minimal model proposed by Krishna et al. [42].

Discussion

The eigenvalue, orthogonal, and correlation methods find the desired subset of identifiable parameters and the desired subset of unidentifiable parameters in all cases. The eigenvalue and orthogonal methods give the same results for each of the models. The subsets of unidentifiable parameters that result from the correlation method are larger than those that result from the eigenvalue or orthogonal method in all cases. The PCA based method fails for all three examples.

The four identifiability testing methods use different approaches to finding unidentifiable parameters. In order to carry out a meaningful comparison of the methods, we need a single identifiability criterion that can be applied to all methods. The positive definiteness of the Hessian matrix (10) of the least squares parameter estimation problem (9) is an appropriate and concise criterion for this purpose. Ultimately, the use of this criterion is justified by the equivalence between local at-a-point identifiability and the positive definiteness of the Hessian matrix (10). This equivalence was established by Grewal and Glower [38].

Since the Hessian matrix is positive definite if and only if all its eigenvalues are strictly positive, we consider a model to be identifiable if the eigenvalues of the Hessian matrix (10) are bounded below by a small strictly positive number ϵ . Based on experience with parameter estimation problems we set $\epsilon = 10^{-4}$. While this choice is arguably arbitrary,

trary, the comparisons of the methods are meaningful since the same value of ϵ is applied in all comparisons.

The numbers of unidentifiable parameters are summarized in Figure 5 for the four methods and the three models. More precisely, the Figure shows how many parameters are selected by each method to be fixed in order to obtain a Hessian matrix with all eigenvalues larger than $\epsilon = 10^{-4}$. From Figure 5 it is apparent that fewer parameters need to be fixed than those selected by the correlation method to obtain a positive definite Hessian matrix or, equivalently, an identifiable model in all examples. In this sense we conclude that the eigenvalue and orthogonal method outperform the correlation method. Since the PCA based method fails in all three examples, we consider it to be inferior to the other methods.

In the remainder of the section we discuss the differences between the methods.

Eigenvalue method vs. orthogonal method

The eigenvalue method can be motivated in two different ways. For one, it can be interpreted as a convexity analysis of the minimization of the least squares cost function ϕ (9) at a nominal point p^* , in parameter space. If ϕ is too flat at p^* , the method identifies the parameters that cause this flatness. Secondly, the method can be looked upon as an approach to identifying linearly dependent columns of S . If $S^T S$ has zero eigenvalues it is not invertible, which implies that S has linearly dependent columns. By fixing those parameters that cause $S^T S$ to be not invertible, the linear independence of columns of S can be ensured. The

orthogonal method selects those columns of S that have the largest possible orthogonal distance to previously selected columns. This way linearly independent columns are found. Essentially, the eigenvalue and orthogonal method use the same criterion, but the eigenvalue method discards those parameters that render a model unidentifiable while the orthogonal method selects the identifiable parameters. Both methods give the same results for the examples treated here.

A formal analysis of the two methods reveals the computational complexity to depend on the magnitudes of $n_y n_t$ and n_p (see Appendix E). If $n_y n_t \approx n_p^2$, as is the case in the JAK-STAT and the NF- κ B example, the eigenvalue method is an order of magnitude in the number of parameters faster than the orthogonal method. If we assume $n_p \leq n_y n_t \ll n_p^2$, as is the case for the MAP kinase example, the eigenvalue method is slightly faster than the orthogonal method. If $n_y n_t \approx n_p$, both methods have the same computational complexity. We stress the complexity analysis is carried out under assumptions that are unfavorable for the eigenvalue method in that the eigenvalue method is assumed to require a full ranking to separate the identifiable from the unidentifiable parameters. In the three examples treated here this is clearly not the case, however. In fact the eigenvalue method stops after three, two, and three iterations in the JAK-STAT, the MAP kinase, and the NF- κ B case, respectively, and therefore turns out to be far less computationally expensive than under the worst case assumption used in Appendix E.

Technical drawbacks of the correlation method

From a technical point of view, the correlation method suffers from two drawbacks. These drawbacks cause the differences between the number of unidentifiable parameters found by the correlation method as compared to the orthogonal and eigenvalue method. First, the correlation $corr(S_{\cdot i}, S_{\cdot j})$ is only an approximation of the linear dependency of columns of S . While two columns that are linearly dependent result in a correlation of ± 1 , columns that have a correlation of ± 1 need not necessarily be linearly dependent (see Appendix F for an example). Therefore, the correlation method may find false positive unidentifiable parameters. Secondly, the correlation $corr(S_{\cdot i}, S_{\cdot j})$ is only a pair wise measure for any two columns $S_{\cdot i}$ and $S_{\cdot j}$ of S . Linear dependence of a set of more than two columns, which does not contain a linearly dependent pair, is not detected.

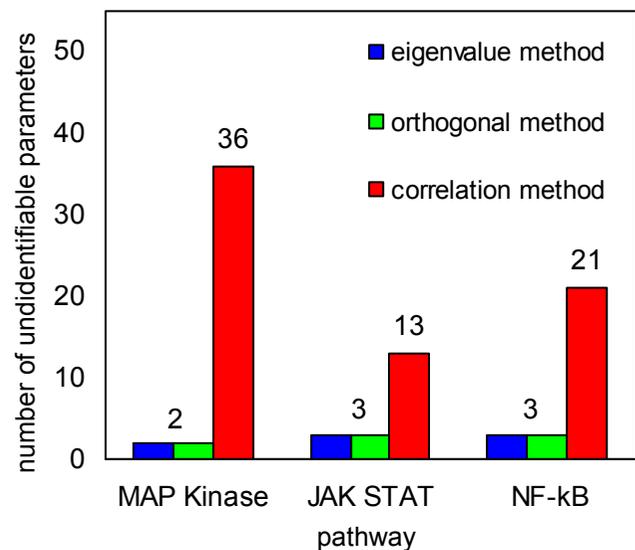


Figure 5
Summary of the method comparison. Since the PCA based method fails in all three cases, it is not included here.

Technical drawbacks of the PCA based method

The PCA based method differs from the three other methods in several aspects. The PCA based method does not use the sensitivity matrix S , but a sensitivity matrix is calculated for each component r_i of the response function separately. In contrast to the other methods, couplings between the components r_i are therefore not considered. A second important difference is that the PCA based method does not use a single criterion, but a combination of criteria which all must hold for a parameter to be considered unidentifiable. Due to this combination of criteria, the method does not select a single parameter in each iteration but a set of parameters. Finally, some parameters that cause S not to have full rank are not found by the PCA based method in all three examples. For this reason, we consider the PCA based method to be inappropriate for the identifiability testing of the three treated pathways.

Conclusion

We proposed a new approach to finding identifiable and unidentifiable parameters of large mechanistic models. The method, which we refer to as the eigenvalue method, was applied to existing models of the JAK-STAT, the MAP kinase, and the NF- κ B pathways. All three pathway models turned out not to be identifiable. In fact our results show that the models would not be identifiable even if all state variables could be measured precisely at a high frequency. The models must therefore be considered to be severely overparameterized. Note state identifiability investigated here is a necessary condition for output identifiability. Since the models turn out not even to be state identifiable, we refrained from further investigations of output identifiability.

The identifiability analysis reveals how to simplify the models in order to arrive at identifiable systems. For the JAK-STAT pathway an identifiable model can be obtained by simplifying only three reactions. For the NF- κ B pathway model the results lead to a simplified model that closely resembles the minimal model published by Krishna et al. [42].

We presented a detailed comparison of the proposed eigenvalue approach to three established methods for identifiability testing, namely, the correlation method [25], the PCA based method [27], and the orthogonal method [24]. These three methods and the new approach suggested are of general interest, because they can be applied to a broad class of nonlinear systems of ordinary differential and algebraic equations. Essentially, the methods were compared with respect to their ability to find an identifiable subset of model parameters for each of the three pathway models. Despite algorithmically different, the eigenvalue and the orthogonal method result in the same subsets of identifiable parameters for each of the

examples. Moreover, these two methods outperform the other methods in that they find the largest subsets of identifiable parameters.

Authors' contributions

TQ and MM developed the eigenvalue method. The discussed method were implemented and compared by TQ. TQ and MM wrote the article.

Appendix

A: Example with co-dominant parameters

We present an illustrative example to motivate the notion of co-dominant parameters. Consider the trivial system $\dot{x}_1 = (p_1 + p_2) \cdot x_1$ with $y_1 = x_1$. Using the notation introduced below equation (10), the eigenvalue λ^1 equals 0 (independent of the choice of p_1, p_2 , sampling times, and initial condition) with $|u^1| = (0.707, 0.707)^T$. From the simple model equation it is apparent that p_1 and p_2 are not identifiable separately, but only their sum $p_1 + p_2$ can be estimated from data. As a result, p_1 and p_2 are co-dominant as indicated by the result $u_1^1 = u_2^1$. Fixing either one of the parameters renders the other one identifiable.

B: Correlation method

We first summarize theoretical aspects of the correlation method and subsequently give a detailed description of the algorithm.

Theory

Let $\Delta p = p - p^*$ describe the deviation of a parameter vector p from the true parameter vector p^* . For sufficiently small $\|\Delta p\|_2$ the response function (4) can be approximated by its linearization at p^* :

$$\begin{aligned} r_i(t_j, p, x_0) &\approx r_i(t_j, p^*, x_0) + \sum_{k=1}^{n_p} \frac{\partial r_i(t_j, p^*, x_0)}{\partial p_k} \Delta p_k \\ &= r_i(t_j, p^*, x_0) + \sum_{k=1}^{n_p} s_{ik}(t_j) \Delta p_k \end{aligned}$$

for all $i \in \{1, \dots, n_y\}$ and $j \in \{1, \dots, n_t\}$. The sum of squared errors of the linearized response functions and the measured outputs $y_i(t_j)$ reads

$$\begin{aligned} R(\Delta p) &= \sum_{i=1}^{n_y} \sum_{j=1}^{n_t} \left(r_i(t_j, p^*, x_0) + \sum_{k=1}^{n_p} s_{ik}(t_j) \Delta p_k - y_i(t_j) \right)^2 \\ &= \sum_{i=1}^{n_y} \sum_{j=1}^{n_t} \sum_{k=1}^{n_p} (s_{ik}(t_j) \Delta p_k)^2. \end{aligned} \tag{20}$$

The last equality holds, because the $y_i(t_j)$ are assumed to be noise free measurements and p^* are the true parameter values by assumption. Therefore $r_i(t_j, p^*, x_0) - y_i(t_j) = 0$ for all $i \in \{1, \dots, n_y\}$ and $j \in \{1, \dots, n_t\}$.

Let $\Delta\hat{p} = (\Delta\hat{p}_1, \Delta\hat{p}_2, \dots, \Delta\hat{p}_{n_p})^T$ denote the least squares estimates of Δp , i.e. the particular values of Δp that minimize equation (20). The necessary condition for optimality $\nabla R(\Delta p)|_{\Delta p} = 0$ is equivalent to the so called normal equation $S^T S \Delta p = 0$. Solving the normal equation for Δp therefore results in the least squares estimate $\Delta\hat{p}$.

If $S^T S$ has full rank, the normal equation only admits the solution $\Delta p = 0$. This implies $\Delta p = 0$, or equivalently $\hat{p} = p^*$, is the unique minimizer of the least squares function $R(\Delta p)$. Consequently, the model parameters p are locally identifiable by definition. Note only local identifiability can be inferred, since an approximation by linearization was introduced. If, on the other hand, $S^T S$ does not have full rank, there exists a non-trivial solution $\Delta\hat{p} \neq 0$ of the normal equation. As a consequence, a parameter vector $\hat{p} \neq p^*$ exists that results in the same model response as p^* . Therefore the model parameters p are not identifiable by definition.

The matrix $S^T S$ has full rank if and only if the columns of S are linearly independent [43]. The central idea of the correlation method is to detect linear dependence of two columns of S by calculating the correlation of these columns. Parameters corresponding to columns of S with a correlation of ± 1 are considered unidentifiable.

Algorithm for the correlation method

Let C denote the matrix that contains those absolute correlation values between columns of S that exceed the threshold $1 - c$. C is given by:

$$C = \begin{pmatrix} \text{corr}^*(S_{\cdot 1}, S_{\cdot 1}) & \dots & \text{corr}^*(S_{\cdot 1}, S_{\cdot n_p}) \\ \vdots & \ddots & \vdots \\ \text{corr}^*(S_{\cdot n_p}, S_{\cdot 1}) & \dots & \text{corr}^*(S_{\cdot n_p}, S_{\cdot n_p}) \end{pmatrix}$$

where $S_{\cdot j}$ denotes the j th column of S and $\text{corr}^*(S_{\cdot i}, S_{\cdot j})$ is defined in equation (18). The algorithm proceeds in the following way:

1. Choose Δt and an initial guess p^* . Set $c = 0$, $I = \{1, \dots, n_t\}$, and initialize U to an empty list. Calculate the sensitivity matrix S .

2. Calculate the matrix C .

3. For all $i \in I$ calculate $c_i^{\text{tot}}(I)$ as defined in equation (17).

4. If I is empty, stop. The list U contains the parameter indices in the order of the least to most identifiable parameter.

5. Find r_q such that $|c_{r_q}^{\text{tot}}| = \max_{i \in I}(|c_i^{\text{tot}}|)$.

6. If $c_{r_q}^{\text{tot}} = 0$, increase c by 0.01 and return to step 2.

Otherwise, remove r_q from I , append r_q to the list U , and return to step 3.

The list U contains the desired result as explained in step 4. Note c is incremented in steps of $\Delta_c = 0.01$. The choice of this increment is not critical as long as it is small enough. If Δ_c is chosen smaller than necessary the same ranking of parameters will be obtained but the algorithm may require more computation time.

C: PCA based method

We explain the second and third PCA based criterion and present the algorithm for the PCA based method. The first criterion is explained in the section entitled Methods.

Description of the second and third criterion of the PCA based method

Let $\tilde{\lambda}^{i,j}$ and $\tilde{u}^{i,j}$ denote the j th eigenvalue and corresponding eigenvector of $\tilde{S}^{iT} \tilde{S}^i$, respectively, and assume the eigenvalues to be ordered such that $\tilde{\lambda}^{i,1} \leq \tilde{\lambda}^{i,2} \leq \dots \leq \tilde{\lambda}^{i,n_p}$. For each $i \in \{1, \dots, n_p\}$ evaluate the second and third criterion as follows.

- Second criterion:

Set $L_1 = \{1, \dots, n_p\}$.

For q from 1 to n_p do:

- Find r_l such that $|u_{r_l}^{i,q}| = \max_{j \in L_q} \left(\sum_{k=1}^{n_p} (\tilde{u}_j^{i,k})^2 \right)$.

- Set $m_{2,q}^i = r_l$.

- Set $L_{q+1} = L_q - \{r_l\}$.

The second criterion is based on the sum of squared eigenvector entries

$$\sum_{k=1}^{n_p} (\tilde{u}_j^{i,k})^2. \tag{21}$$

Note in this sum, i and j are fixed, therefore the sum is not carried out over components of a particular eigenvector, but the l th entries of all eigenvectors are summed. The parameter with the largest value of (21) is selected.

• Third criterion:

Set $L_1 = \{1, \dots, n_y\}$.

For q from 1 to n_p do:

- Find r_l such that $|u_{r_l}^{i,q}| = \max_{j \in L_q} (|\tilde{u}_j^{i, n_p + 1 - q}|)$.

- Set $m_{3, n_p + 1 - q}^i = r_l$.

- Set $L_{q+1} = L_q - \{r_l\}$.

The third criterion loops over the eigenvectors but, in contrast to the first and second criterion, starts with the eigenvector that belongs to the largest eigenvalue. As in the first criterion, parameters are selected based on the absolute value of the eigenvector component. The index of the parameter selected first will be placed last in ranking M_3^i , the index of the parameter selected second will be placed at position $n_p - 1$, and so on.

Algorithm for the PCA based method

1. Choose Δt and an initial guess p^* . For all $i \in \{1, \dots, n_y\}$ let M_1^i , M_2^i , and M_3^i be empty lists. For all $q \in \{1, \dots, n_p\}$ set $N_q = \emptyset$. Set J to an empty list.

2. For all $i \in \{1, \dots, n_y\}$ calculate $\tilde{S}^{iT} \tilde{S}^i$, its eigenvectors $\tilde{u}^{i,j}$, and the corresponding eigenvalues $\tilde{\lambda}^{i,j}$, $j \in \{1, \dots, n_p\}$.

3. For $i \in \{1, \dots, n_y\}$ calculate M_1^i , M_2^i , and M_3^i by applying ranking criterion one to three as described in the Method section. Let $m_{j,k}^i$ denote the k th entry of M_j^i .

4. For q from 1 to n_p do:

• For $j \in \{1, 2, 3\}$ and $i \in \{1, \dots, n_y\}$ set $W_j^i(q) = \{m_{j,k}^i \mid k \in \{1, \dots, q\}\}$.

$$N_q = \bigcap_{\substack{1 \leq j \leq 3 \\ 1 \leq i \leq n_y}} W_j^i(q) - N_{q-1} \dots - N_1$$

• Only if $N_q \neq \emptyset$ append N_q to the list J .

The final ranking that combines the PCA results of all $\tilde{S}^{iT} \tilde{S}^i$ and all three criteria is given by J , where the first position in J corresponds to the least and the last position in J corresponds to the most identifiable parameter.

D: Orthogonal method

In this part we present a geometric interpretation the orthogonal method, followed by a description of algorithmic details.

Visualization of the concept of the orthogonal method

The orthogonal method is visualized in Figure 6 for $q = 3$. Two columns of S have already been selected and represent the first and the second column of X^2 denoted by X_1^2 and X_2^2 , respectively. These two columns span the vector space V_2 . Without restriction this vector space is assumed to be the xy plane and k is the next parameter to be chosen.

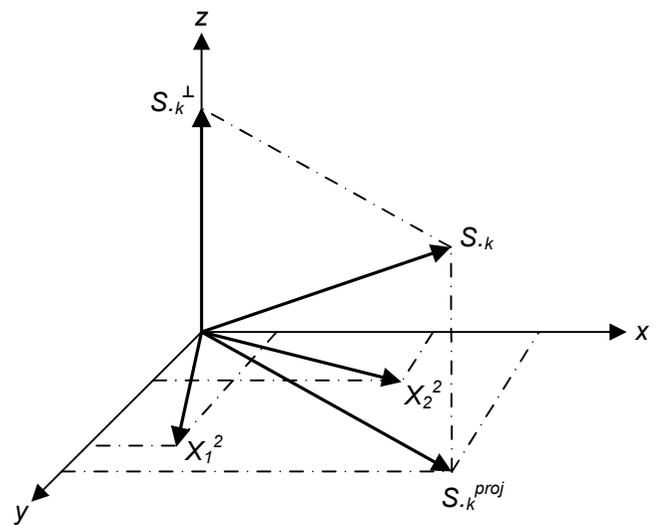


Figure 6
Visualization of the concepts of the orthogonal method.

Algorithm for the orthogonal method

Algorithmic details for the orthogonal method are given below.

1. Choose Δt and an initial guess p^* . Initialize $W^{ortho} \in \mathbb{R}^{n_p}$ to a list of zeros.
2. Calculate the sensitivity matrix S .
3. Find r_1 such that $\|S_{\cdot r_1}\|_2^2 = \max_{k \in \{1, \dots, n_p\}} (\|S_{\cdot k}\|_2^2)$, set $W_1^{ortho} = r_1, X^1 = S_{\cdot r_1}$, and $L_1 = \{1, \dots, n_p\} - \{r_1\}$.

4. For q from 1 to n_p do:

- Calculate P_q as given by equation (22).
- Find r_q such that $\|S_{\cdot r_q}^\perp\|_2^2 = \max_{k \in L_q} (\|S_{\cdot k}^\perp\|_2^2)$, with $S_{\cdot k}^\perp$ corresponding to the k th column of P_q .
- Set $S_{\cdot k}^\perp \geq o$:

$$\text{set } W_q^{ortho} = r_q,$$

$$\text{set } X^q = (X^{q-1}, S_{\cdot r_q}) \in \mathbb{R}^{n_y n_t} \times \mathbb{R}^q,$$

$$\text{and set } L_{q+1} = L_q - \{r_q\}.$$

The ranking is given by W^{ortho} . The shown implementation ignores the stopping criterion applied in the original literature, to ensure a complete ranking.

E: Computational complexity of the orthogonal and the eigenvalue method

The numerical operations of the two methods that dominate the number of necessary numerical operations are 1) matrix inversions and eigenvalue/eigenvector calculations, which both require $O(N^3)$ operations for $N \times N$ matrices, and 2) matrix multiplications, which require $N \cdot M \cdot L$ operations for the multiplication of an $N \times M$ by an $M \times L$ matrix.

In what follows, we assume the eigenvalue method is used to rank all model parameters. This overestimates the computational cost of the eigenvalue method, since the method stops once an identifiable set of parameters has been found.

Computational complexity of the orthogonal method

In compact notation the orthogonal projection in iteration i , where $i \in \{1, \dots, n_p\}$ is arbitrary but fixed, can be written as:

$$\begin{aligned} P_i &= (I - X^i (X^{iT} X^i)^{-1} X^{iT}) S \\ &= S - Y X^{iT} S \\ &= S - Y S' \end{aligned} \tag{22}$$

The major runtime costs result from the computation 1) of Y and 2) of $Y S'$ in equation (22). The number of operations necessary to calculate Y is:

$$\begin{aligned} time_Y &= O \left(\sum_{i=1}^{n_p-1} (n_y n_t i^2 + i^3 + i n_y n_t) \right) \\ &= O(n_y n_t n_p^3 + n_p^4 + n_p^2 n_y n_t) \end{aligned} \tag{23}$$

$$= O(n_y n_t n_p^3 + n_p^4), \tag{24}$$

where all iterations $i = 1, \dots, n_p$ have been accounted for. The first and second term in the sum of equation (23) account for the multiplication of X^i by $(X^{iT} X^i)^{-1}$ ($n_y n_t \times i$ matrix times $i \times i$ matrix) and the inversion of the matrix $X^{iT} X^i$ ($i \times i$ matrix), respectively. The third term accounts for the cost of multiplying X^{iT} by X^i ($i \times n_y n_t$ matrix times $n_y n_t \times i$). This term reduces from $i^2 n_y n_t$ to $i n_y n_t$, the cost of calculating the last row of $X^{iT} X^i$, since $X^{i-1T} X^{i-1}$ is available from the previous iteration:

$$X^{iT} X^i = \left(\begin{array}{c|c} X^{i-1T} X^{i-1} & X_1^{iT} X^i \\ \hline & \vdots \\ X_i^{iT} X_1^i & \dots & X_i^{iT} X^i \end{array} \right).$$

Here X_i^i denotes the i th column of X^i . The last row is equal to the last column, since $X^{iT} X^i$ is symmetric. The cost for i vector multiplications $X_i^{iT} X_k^i (X_k^i \in \mathbb{R}^{n_y n_t})$ with $1 \leq k \leq i$ amounts to $i n_y n_t$, which gives rise to the last term in the sum in equation (23).

The number of operations necessary to calculate $Y S'$ is:

$$\begin{aligned} time_{YS'} &= O \left(\sum_{i=1}^{n_p-1} (n_y n_t i n_p + n_y n_t n_p) \right) \\ &= O(n_y n_t n_p^3 + n_y n_t n_p^2) \\ &= O(n_y n_t n_p^3) \end{aligned} \tag{25}$$

The first term in the sum accounts for the multiplication of Y by S^i ($n_y n_t \times i$ matrix times $i \times n_p$ matrix). The second term accounts for the calculation of S^i from the product $X^{iT} S$ ($i \times n_y n_t$ matrix times $n_y n_t \times n_p$ matrix). The latter reduces to the cost $n_y n_t n_p$ for calculating the last row of S^i , since $X^{i-1T} S$ is available from the previous iteration:

$$S^i = X^{iT} S = \left(\frac{X^{i-1T} S}{X_i^{iT} S_{\cdot 1} \quad \dots \quad X_i^{iT} S_{\cdot n_p}} \right)$$

where X_i^i denotes the last column of X^i and $S_{\cdot i}$ denotes the i th column of S .

The subtraction of S from $Y S^i$ ($n_y n_t \times n_p$ matrix minus $n_y n_t \times n_p$ matrix) in equation (22) requires $O(n_y n_t n_p)$ operations and therefore can be neglected.

In summary the dominant parts in equation (24) and (25) result in a computational complexity of order $O(n_y n_t n_p^3 + n_p^4)$.

Computational complexity of the eigenvalue method

For ease of presentation the description of the algorithm in the Methods section slightly differs from the actual implementation. In the implemented version the Hessian matrix $H = S^T S$ has to be calculated only once at the beginning of the algorithm, which requires $O(n_y n_t n_p^2)$ operations (multiplication of a $n_p \times n_y n_t$ matrix with a $n_y n_t \times n_p$ matrix). In each iteration i , the current Hessian H^i can be obtained by removing the rows and columns from H that correspond to fixed parameters. In each iteration the eigenvalues of the current Hessian matrix have to be calculated. The number of operations necessary for the eigenvalue method is as follows.

$$time = O \left(n_y n_t n_p^2 + \sum_{i=1}^{n_p-1} (n_p - i)^3 \right) \quad (26)$$

$$= O(n_y n_t n_p^2 + n_p^4) \quad (27)$$

The first term accounts for the initial calculation of the Hessian matrix. The second term accounts for the eigenvalue calculation.

Comparing computational complexity of the eigenvalue and the orthogonal method

The computational complexity of the eigenvalue method is of order $O(n_y n_t n_p^2 + n_p^4)$ (see equation (27)), whereas the computational complexity of the orthogonal method is of order $O(n_y n_t n_p^3 + n_p^4)$ (see equation (24)). For a comparison we have to distinguish between the following three cases. In the first case we assume $n_y n_t \leq n_p$. In this case the complexity of both methods is dominated by the term n_p^4 and both methods are equally efficient. In the examples treated here, this case does not occur. In the second case we assume $n_p \leq n_y n_t < n_p^2$. Here the computational complexity of the eigenvalue and the orthogonal method reduce to $O(n_p^4)$ and $O(n_y n_t n_p^3)$, respectively. This case applies to the MAP-kinase example and the eigenvalue method is faster than the orthogonal method. The difference is the more pronounced the larger $n_y n_t - n_p$. In the last case we assume that $n_y n_t \approx n_p^2$. This is true in the JAK-STAT and the NF- κ B example. Here the complexity of the eigenvalue method and the orthogonal method reduce to $O(n_p^4)$ and $O(n_p^5)$, respectively. In this case the eigenvalue method is an order of magnitude in the number of parameters faster than the orthogonal method.

F: Correlation is not equal to linear dependence

In order to show that a correlation of ± 1 between two vectors does not necessarily imply their linear dependence, we present a simple counter-example. Consider the vectors $x = (1, -1)^T$ and $y = (2, 0)^T$. According to equation (13) the correlation $corr(x, y) = 1$ results. Clearly, however, x and y are linearly independent. A set of parameters that are found to be unidentifiable with the correlation method may therefore contain false positives.

Acknowledgements

We gratefully acknowledge funding by Deutsche Forschungsgemeinschaft (DFG) under grant MO 1086/5-1. We thank Prof. F. Schaper for providing his insight on biological aspects and C. Michalik for fruitful discussions. A previous version of this paper was presented at the 2nd Foundations of Systems Biology in Engineering conference, Stuttgart, Germany, Sept. 2007.

References

1. Lipniacki T, Paszek P, Brasier ARAR, Luxon B, Kimmel M: **Mathematical model of NF-kappaB regulatory module.** *J Theor Biol* 2004, **228(2)**:195-215.
2. Yamada S, Shiono S, Joo A, Yoshimura A: **Control mechanism of JAK/STAT signal transduction pathway.** *FEBS Lett* 2003, **534(1-3)**:190-196.
3. Schoeberl B, Eichler-Jonsson C, Gilles ED, Mueller G: **Computational modeling of the dynamics of the MAP kinase cascade**

- activated by surface and internalized EGF receptors. *Nat Biotechnol* 2002, **20(4)**:370-375.
4. Walter E, Pronzato L: **Identification of parametric models from experimental data.** Springer Berlin; 1997.
 5. Swameye I, Muller TG, Timmer J, Sandra O, Klingmuller U: **Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by databased modeling.** *Proc Natl Acad Sci USA* 2003, **100(3)**:1028-1033.
 6. Xia X, Moog CH: **Identifiability of nonlinear systems with application to HIV/AIDS models.** *IEEE T Automat Contr* 2003, **48(2)**:330-336.
 7. Jeffrey A, Xia X: **Identifiability of HIV/AIDS models.** *Deterministic and Stochastic Models of AIDS Epidemics and HIV Infections with Intervention* 2005:255-286.
 8. Wu H, Zhu H, Miao H, Perelson A: **Parameter Identifiability and Estimation of HIV/AIDS Dynamic Models.** *Bulletin of Mathematical Biology* 2008, **70(3)**:785-799.
 9. Zak DE, Gonye GE, Schwaber JS, Doyle FJ: **Importance of input perturbations and stochastic gene expression in the reverse engineering of genetic regulatory networks: insights from an identifiability analysis of an in silico network.** *Genome Res* 2003, **13(11)**:2396-2405.
 10. Rodriguez-Fernandez M, Mendes P, Banga JR: **A hybrid approach for efficient and robust parameter estimation in biochemical pathways.** *Biosystems* 2006, **83(2-3)**:248-265.
 11. Jacquez JA, Perry T: **Parameter estimation: local identifiability of parameters.** *American Journal of Physiology-Endocrinology And Metabolism* 1990, **258(4)**:727-736.
 12. Ljung L, Glad T: **On global identifiability for arbitrary model parametrizations.** *Automatica* 1994, **30(2)**:265-276.
 13. Pohjanpalo H: **System identifiability based on power-series expansion of solution.** *Math Biosci* 1978, **41(1-2)**:21-33.
 14. Godfrey K, DiStefano J: **Identifiability of Parametric Models.** Pergamon, Oxford; 1987.
 15. Audoly S, Bellu G, D'Angiò L, Saccomani MP, Cobelli C: **Global identifiability of nonlinear models of biological systems.** *IEEE Trans Biomed Eng* 2001, **48**:55-65.
 16. Walter E, Braems I, Jaulin L, Kieffer M: **Guaranteed Numerical Computation as an Alternative to Computer Algebra for Testing Models for Identifiability.** *Lecture Notes in Computer Science* 2004:124-131.
 17. Vajda S, Rabitz H: **State isomorphism approach to global identifiability of nonlinear-systems.** *IEEE T Automat Contr* 1989, **34(2)**:220-223.
 18. Chappell M, Godfrey K: **Structural identifiability of the parameters of a nonlinear batch reactor model.** *Math Biosci* 1992, **108(2)**:241-51.
 19. Peeters RLM, Hanzon B: **Identifiability of homogeneous systems using the state isomorphism approach.** *Automatica* 2005, **41(3)**:513-529.
 20. Asprey SP, Macchietto S: **Statistical tools for optimal dynamic model building.** *Comput Chem Eng* 2000, **24(2-7)**:1261-1267.
 21. Asprey SP, Macchietto S: **Designing robust optimal dynamic experiments.** *J Process Control* 2002, **12(4)**:545-556.
 22. Asprey SP, Macchietto S: **Dynamic model development: Methods, Theory and Applications.** Elsevier, Amsterdam, The Netherlands; 2003.
 23. Brun R, Reichert P, Kunsch HR: **Practical identifiability analysis of large environmental simulation models.** *Water Resour Res* 2001, **37(4)**:1015-1030.
 24. Yao KZ, Shaw BM, Kou B, McAuley KB, Bacon DW: **Modeling ethylene/butene copolymerization with multi-site catalysts: Parameter estimability and experimental design.** *Polym React Eng* 2003, **11(3)**:563-588.
 25. Jacquez JA, Greif P: **Numerical parameter identifiability and estimability: Integrating identifiability, estimability, and optimal sampling design.** *Math Biosci* 1985, **77(1-2)**:201-227.
 26. Vajda S, Rabitz H, Walter E, Lecourtier Y: **Qualitative and quantitative identifiability analysis of nonlinear chemical kinetic models.** *Chem Eng Commun* 1989, **83**:191-219.
 27. Degenring D, Froemel C, Dikta G, Takors R: **Sensitivity analysis for the reduction of complex metabolism models.** *J Process Control* 2004, **14(7)**:729-745.
 28. Froemel C: **Parameterreduktion in Stoffwechselmodellen mit Methoden der Statistik.** In *Master's thesis Fachhochschule Aachen*, Abteilung Juelich, Fachbereich Physikalische Technik, Studienrichtung Technomathematik; 2003.
 29. Hengl S, Kreutz C, Timmer J, Maiwald T: **Data-based identifiability analysis of non-linear dynamical models.** *Bioinformatics* 2007, **23(19)**:2612-2618.
 30. Chu Y, Hahn J: **Parameter Set Selection via Clustering of Parameters into Pairwise Indistinguishable Groups of Parameters.** *Industrial & Engineering Chemistry Research* 2009 in press.
 31. Quaiser T, Marquardt W, Moennigmann M: **Local Identifiability Analysis of Large Signaling Pathway Models.** In *2nd Conference Foundation of Systems Biology in Engineering, Proceedings Plenary and Contributed Papers*, ISBN 978-3-8167-7436-5 Edited by: Allgoewer F, Reuss M. Stuttgart: Fraunhofer IRB Verlag; 2007:465-470.
 32. Schittkowski K: **Experimental design tools for ordinary and algebraic differential equations.** *Ind Eng Chem Res* 2007, **46(26)**:9137-9147.
 33. Li S, Petzold L: **Design of new Daspk for Sensitivity Analysis.** *Technischer Report* 1999.
 34. Beck J, Arnold K: *Parameter estimation in engineering and science* Wiley New York; 1977.
 35. Zill D, Cullen M: *Advanced Engineering Mathematics* Jones & Bartlett Publishers; 2006.
 36. Gadkar KG, Gunawan R, Doyle FJ: **Iterative approach to model identification of biological networks.** *BMC Bioinformatics* 2005, **6**:155.
 37. Yue H, Brown M, Knowles J, Wang H, Broomhead DS, Kell DB: **Insights into the behaviour of systems biology models from dynamic sensitivity and identifiability analysis: a case study of an NF-kappaB signalling pathway.** *Mol BioSyst* 2006, **2(12)**:640-649.
 38. Grewal MS, Glover K: **Identifiability of linear and nonlinear dynamical-systems.** *IEEE T Automat Contr* 1976, **21(6)**:833-836.
 39. Heinrich PC, Behrmann I, Haan S, Hermanns HM, Mueller-Newen G, Schaper F: **Principles of interleukin IL-6-type cytokine signalling and its regulation.** *Biochem J* 2003, **374(Pt 1)**:1-20.
 40. Berkers JA, en Henegouwen PM van Bergen, Boonstra J: **Three classes of epidermal growth factor receptors on HeLa cells.** *J Biol Chem* 1991, **266(2)**:922-927.
 41. Chung JC, Sciaky N, Gross DJ: **Heterogeneity of epidermal growth factor binding kinetics on individual cells.** *Biophys J* 1997, **73(2)**:1089-1102.
 42. Krishna S, Jensen MH, Sneppen K: **Minimal model of spiky oscillations in NF-kappaB signaling.** *Proc Natl Acad Sci USA* 2006, **103(29)**:10840-10845.
 43. Strang G: *Introduction to Linear Algebra* Wellesley Cambridge Pr; 2003.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

