PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Joining forces of Bayesian and frequentist methodology: a study for inference in the presence of non-identifiability

Andreas Raue, Clemens Kreutz, Fabian Joachim Theis and Jens Timmer

Phil. Trans. R. Soc. A 2013 **371**,
doi: 10.1098/rsta.2011.0544

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A

rsta.royalsocietypublishing.org

Research

Cite this article: Raue A, Kreutz C, Theis FJ, Timmer J. 2013 Joining forces of Bayesian and frequentist methodology: a study for inference in the presence of non-identifiability. Phil Trans R Soc A 371: 20110544.
http://dx.doi.org/10.1098/rsta.2011.0544

One contribution of 17 to a Discussion Meeting Issue 'Signal processing and inference for the physical sciences'.

Subject Areas:
parameter estimation, data analysis, identifiability, Markov chain Monte Carlo sampling, prediction uncertainty, systems biology

Keywords:
identifiability, profile likelihood, Bayesian Markov chain Monte Carlo sampling, posterior propriety, propagation of uncertainty, prediction uncertainty

Author for correspondence:
Jens Timmer
e-mail: jeti@fdm.uni-freiburg.de

Electronic supplementary material is available at http://dx.doi.org/10.1098/rsta.2011.0544 or via http://rsta.royalsocietypublishing.org.

Royal Society Publishing
Informing the science of the future

Joining forces of Bayesian and frequentist methodology: a study for inference in the presence of non-identifiability

Andreas Raue[1], Clemens Kreutz[1], Fabian Joachim Theis[2] and Jens Timmer[1,3,4]

[1]Institute for Physics, University of Freiburg, Freiburg, Germany
[2]Helmholtz Zentrum Munich, and Department of Mathematics, Technical University of Munich, Munich, Germany
[3]BIOSS Centre for Biological Signalling Studies, and Freiburg Institute for Advanced Studies (FRIAS), Freiburg, Germany
[4]Department of Clinical and Experimental Medicine, Linköping University, Sweden

Increasingly complex applications involve large datasets in combination with nonlinear and high-dimensional mathematical models. In this context, statistical inference is a challenging issue that calls for pragmatic approaches that take advantage of both Bayesian and frequentist methods. The elegance of Bayesian methodology is founded in the propagation of information content provided by experimental data and prior assumptions to the posterior probability distribution of model predictions. However, for complex applications, experimental data and prior assumptions potentially constrain the posterior probability distribution insufficiently. In these situations, Bayesian Markov chain Monte Carlo sampling can be infeasible. From a frequentist point of view, insufficient experimental data and prior assumptions can be interpreted as non-identifiability. The profile-likelihood approach offers to detect and to resolve non-identifiability by experimental design iteratively. Therefore, it allows one to better constrain the posterior probability distribution until Markov chain Monte Carlo sampling can be used securely. Using an application from cell biology, we compare both methods and show that a successive application of the two methods facilitates a realistic assessment of uncertainty in model predictions.

1. Introduction

In many scientific situations, mathematical models are used to predict the properties of a system under study using explicitly formulated assumptions and hypotheses. This is especially important for complex applications that do not allow for an intuitive understanding of the processes involved. Applications range from analyses in particle physics [1] or in climate research [2,3] to quantifying dynamical processes in cell biology [4]. Often, these mathematical models contain parameters that are unknown or known only with large uncertainty. For example, in the biochemical models that are used in cell biology, parameters such as reaction rate constants, amount of molecular compounds, detection sensitivities or measurement backgrounds are often unknown. Before a model can be used for prediction reliably, the unknown parameters have to be estimated by comparing model output to experimental data.

For a realistic assessment of the accuracy of model predictions, it is important that uncertainties in the experimental data and in prior assumptions are propagated correctly via the parameters to the desired predictions. Bayesian Markov chain Monte Carlo (MCMC) sampling facilitates this propagation of uncertainties by sampling from the posterior probability distribution [5]. However, if experimental data are limited and the mathematical models are nonlinear and contain many unknown parameters, the posterior probability distribution can be insufficiently constrained. For such an insufficiently constrained posterior, the probability mass can be distributed widely in a high-dimensional parameter space. Consequently, MCMC sampling can quickly become infeasible.

Alternatively, one can resort to frequentist methods in this situation. Here, insufficient experimental data and prior assumptions can be interpreted as non-identifiability of the model parameters [6]. We applied a generic approach that uses the profile likelihood to detect both structural and practical non-identifiability [7]. Furthermore, this approach allows one to design new experiments that resolve non-identifiability. Therefore, it is beneficial to further constrain the posterior probability distribution until MCMC sampling can be applied reliably and efficiently.

We compared the results of both MCMC sampling and profile-likelihood methods. In the absence of non-identifiability, the results of both methods are in good agreement. However, in the presence of non-identifiability their results can be substantially different. Our results imply that MCMC sampling in the presence of non-identifiability can be misleading. Therefore, we suggest a successive application of the two methods that ensures a realistic assessment of uncertainty in model predictions.

(a) Frequentist methods

Maximum-likelihood estimation of model parameters is a theoretically well-developed area [8]. Based on an assumption on the distribution of the measurement noise, the likelihood function $L(y|\theta)$ of the data y given the parameters θ describes the agreement of model output and experimental data. In the case of normally distributed measurement noise $\epsilon \sim \mathrm{N}(0, \sigma^2)$, the likelihood reads as

$$L(y|\theta) = \prod_{k=1}^{m} \prod_{l=1}^{d_k} \frac{1}{\sqrt{2\pi\sigma_{kl}^2}} \exp\left(-\frac{1}{2}\left(\frac{y_{kl} - y_k(t_l, \theta)}{\sigma_{kl}}\right)^2\right), \tag{1.1}$$

where m model outputs $y_k(t_l, \theta)$ and d_k data instances for each model output, such as time points t_l, can be considered. The maximum of $L(y|\theta)$, i.e. the best fit of the model to the data, provides a point estimate $\hat{\theta}$ of the parameters. This maximum-likelihood estimate (MLE) can be calculated for nonlinear models by numerical optimization methods; see e.g. the trust region algorithm in [9] and the general introductions in [10,11]. The uncertainty of the estimate $\hat{\theta}$ is buried in the shape of the likelihood function. Figure 1 shows an illustration of the likelihood for three typical cases.

If the amount of model quantities that can be accessed by experiments is limited, a subset of parameters can be structurally non-identifiable. This indicates that the parametrization of the model is such that two or more parameters can compensate their effects and yield exactly the

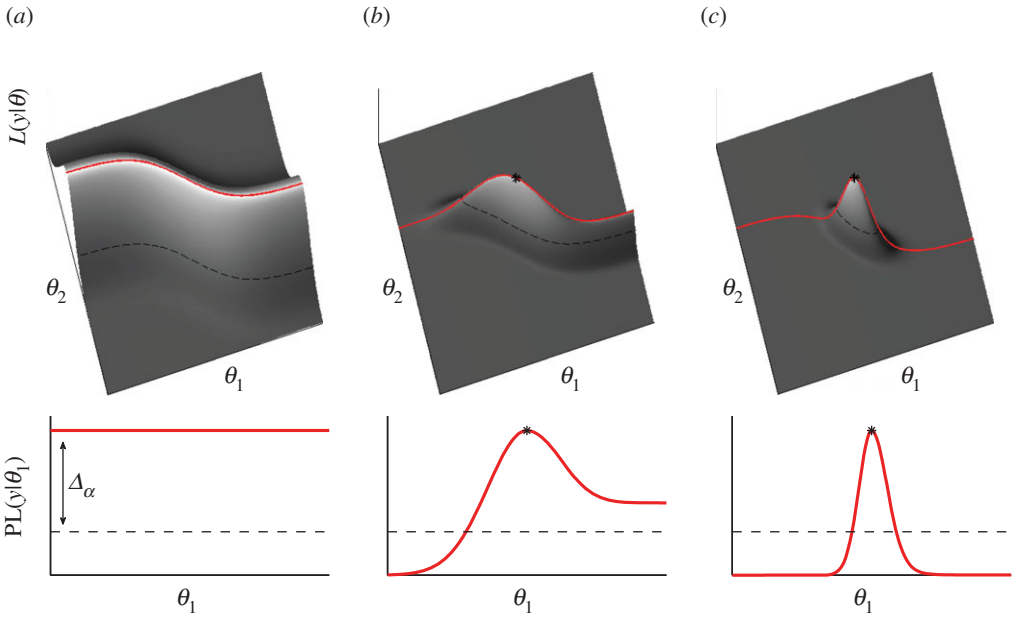

Figure 1. Identifiability analysis using likelihood profiles: the upper panels show as illustration the shape of the likelihood $L(y|\theta)$ in two dimensions for three typical cases. The traces of the profiles in parameter space are indicated as full red lines. The lower panels show the respective profile likelihood $\mathrm{PL}(y|\theta_1)$ for the dimension of parameter θ_1 as full red lines. In all panels, the asterisks denote the MLE in cases where a unique solution exists and the dashed lines denote the threshold Δ_α that yields a likelihood-based confidence region [10,12]. Three typical cases can arise: (*a*) A flat profile indicates structural non-identifiability. In this case, no unique solution for MLE exists. (*b*) A profile that decreases but tails out to a plateau to one or both sides indicates practical non-identifiability. (*c*) A profile that tails out to zero on both sides quickly enough, i.e. at least exponentially fast, indicates structural and practical identifiability. The confidence interval of θ_1 for (*a*) is infinite, for (*b*) has only a lower bound, and for (*c*) has a finite range. (Online version in colour.)

same model outputs $y_k(t_l, \theta)$. This in turn results in a constant likelihood value on a sub-manifold (figure 1*a*). Consequently, the MLE for the parameters cannot be determined uniquely. The parameter relations that cause the structural non-identifiability are akin to gauge invariances in physical theories. However, for complex models, it can be difficult to detect structural non-identifiability. For example, if models can only be evaluated by numerical simulation, such as in the case of detector models used in particle physics [13] or dynamical models that are used in cell biology [4], structural non-identifiability cannot be detected directly. In the latter case, methods for *a priori* structural non-identifiability analysis exist that analyse the structure of the ordinary differential equations (ODEs) without having an analytical solution. A comparison of these methods can be found in Chis *et al*. [14]. However, *a priori* methods are often limited to linear ODE systems or are impracticable for models containing many parameters [15]. Arguably one of the most practicable *a priori* methods that has also been proved to work for larger models applies a probabilistic algorithm [16,17].

In addition to structural non-identifiability, model parameters can be practically non-identifiable [7]. This type of non-identifiability arises if the amount and quality of experimental data are limited. It cannot be detected by *a priori* methods. However, practical non-identifiability is of equal importance. A generic approach that allows one to detect both structural and practical non-identifiability at the same time uses the concept of the profile likelihood [7,18]. The profile likelihood (PL) can be calculated for each parameter θ_i individually by

$$\mathrm{PL}(y|\theta_i) = \max_{\theta_{j\neq i}}[L(y|\theta)]. \tag{1.2}$$

The equation indicates that for each value of θ_i all of the remaining parameters θ_j are re-optimized; see figure 1 for illustration. The profiles $\mathrm{PL}(y|\theta_i)$ break down the uncertainty contained in the high-dimensional likelihood $L(y|\theta)$ to a footprint in one dimension. It allows for reliable conclusions about whether a parameter can be inferred from the experimental data. Three typical cases arise and can be detected from the profiles. A flat profile with a constant value indicates structural non-identifiability (cf. figure 1*a*). A profile that decreases but tails out to a plateau to one or both sides indicates practical non-identifiability (cf. figure 1*b*). A profile that tails out to zero on both sides quickly enough, i.e. at least exponentially fast, indicates structural and practical identifiability (cf. figure 1*c*). Experimental design and model reduction strategies based on the profile likelihood allow one to resolve parameter non-identifiabilities iteratively; for an application see [19].

Furthermore, the profile likelihood allows one to assess likelihood-based confidence intervals [20–22]. A confidence interval $[\sigma_i^-, \sigma_i^+]$ to a confidence level $\alpha = 0.95$ signifies that the true value of the parameter θ_i^* is expected to be inside the interval with 95% probability. Using a threshold $\Delta_\alpha = \mathrm{Q}(\chi^2_{\mathrm{d.f.}}, \alpha)$, which is the α quantile of the χ^2-distribution with d.f. degrees of freedom [10], confidence intervals can be determined from the profiles by $\{\theta_i \mid -2\log(\mathrm{PL}(y|\theta_i)/L(y|\hat{\theta})) < \Delta_\alpha\}$.

(b) Bayesian methods

By applying Bayes' theorem, the likelihood function (1.1) is extended by the prior probability density function (PDF) of the parameters $P(\theta)$ and normalized by a factor c, yielding the posterior PDF of the parameters,

$$P(\theta|y) = c\,L(y|\theta)\,P(\theta). \tag{1.3}$$

In analogy to the MLE, the maximum *a posteriori* (MAP) estimate is defined by maximizing $P(\theta|y)$. The only decisive difference to frequentist methodology is the choice of the prior $P(\theta)$. The direct computation of the normalization factor c is not feasible for a high-dimensional parameter space. However, extensive MCMC sampling offers a way to evaluate $P(\theta|y)$ despite the unknown c. This intriguing feature opens the prospect of considering the full high-dimensional posterior PDF for statistical inference. Most prominently, the Metropolis–Hastings algorithm [23,24] defines a Markov process where transitions $\theta \to \theta'$ are generated using a proposal function $q(\theta|\theta')$ that eventually produces a series of samples of the posterior PDF $P(\theta|y)$. The transitions are accepted with probability

$$\alpha(\theta,\theta') = \min\left[1, \left(\frac{L(y|\theta')}{L(y|\theta)}\right)\left(\frac{q(\theta|\theta')}{q(\theta'|\theta)}\right)\right]. \tag{1.4}$$

For efficiency of the sampling, the choice of the proposal function $q(\theta|\theta')$ is important. Often, proposals drawn from a multivariate normal distribution are convenient. One of the most simple implementations uses $q(\theta|\theta') \sim \mathrm{N}(0, s \cdot \mathbb{I})$ where $s \cdot \mathbb{I}$ is a scaled identity matrix (SIM). Too small a choice of s in relation to the actual shape of the posterior PDF will cause the process to converge slowly and to produce correlated samples. Too large a choice of s will lead to rejection of too many proposals θ' yielding a slow sampling. Assuming that the posterior PDF is a multivariate normal distribution, the optimal acceptance rate of proposals is ≈ 0.23 [25]. Nevertheless, in the light of complex and nonlinear models with possibly limited amount and accuracy of experimental data, this assumption is problematic because the shape of the posterior PDF can be far from the PDF of a normal distribution. For a high-dimensional parameter space, computational efficiency also becomes an important issue. In these cases, the Markov chain may have to move along complex structures [12, fig. 8.2.2]. To increase efficiency, more sophisticated methods take into account the natural geometry of the posterior PDF, e.g. the manifold Metropolis adjusted Langevin algorithm (MMALA) takes into account the local gradient and curvature information [26].

Before applying an MCMC sampling method, the prior PDF of the parameters needs to be specified. Given that empirical evidence exists about the distribution of a parameter, such as a previous measurement or estimation, $P(\theta)$ should incorporate this prior knowledge accordingly [27]. If no empirical evidence about the parameter value is available, the prior should be chosen

as uninformative [12]. This requires a flat metric in parameter space, i.e. one that does not artificially favour certain parameter values. Depending on the parametrization of the model, it can in practice be difficult to obtain a flat metric and hence to specify an uninformative prior [27].

The most crucial problem that MCMC sampling faces is to ensure that the samples obtained realistically represent the actual posterior PDF. One instance where the convergence of the Markov chain fails is if the posterior PDF is not proper [28]. Posterior PDFs are called *proper* if they are integrable [12]. This means that it has to tail out to zero sufficiently fast over the appreciable range of the parameters such that its integral can be normalized to one. Non-identifiable parameters cause the posterior PDF to be non-proper [28]. This indicates that neither the prior assumptions nor the likelihood that represents the experimental data constrain the posterior PDF sufficiently. A practical consequence is that the Markov chain cannot converge and hence gives inaccurate results [29]. For convergence, it is required that the Markov chain is positive recurrent, which is not given in the case of non-identifiability [30]. The user must ensure parameter identifiability before an MCMC technique can be used securely. If models contain many unknown parameters and the posterior PDF is not sufficiently constrained, the convergence of the Markov chain can be impracticably slow even for a proper posterior PDF.

2. Results

For reliable inference in the presence of non-identifiability, we propose a joint approach that take advantage of both profiling and MCMC sampling methods. The profile likelihood is suitable to detect parameter non-identifiability [7]. Furthermore, it allows for experimental design that helps to resolve parameter non-identifiability [19]. This ensures that the posterior PDF is proper and well constrained. Subsequently, efficient MCMC sampling [26] can be used reliably to generate samples of the posterior PDF. Finally, the uncertainty in model predictions can be assessed realistically. The workflow of this joint approach is displayed in figure 2*a*.

Both profiling and MCMC sampling methods separately are used for inference in various fields, e.g. in particle physics [1,31] or in climate research [2,3]. Here, we present results of using the joint approach that take advantage of both methods for an application from cell biology [32]. An ODE model was used to describe the concentration dynamics of six molecular compounds involved in the interplay of the hormone erythropoietin (Epo) and its corresponding membrane receptor (EpoR) (figure 2*b*). Epo is an important factor in the differentiation of blood cells. The dynamics of the molecular compounds are formulated as

$$\frac{\mathrm{d}\boldsymbol{x}(t,\theta)}{\mathrm{d}t} = \boldsymbol{f}(\boldsymbol{x}(t,\theta),\theta), \tag{2.1}$$

and a mapping of the dynamics to experimentally accessible quantities by the model output function

$$\boldsymbol{y}(t,\theta) = \boldsymbol{g}(\boldsymbol{x}(t,\theta),\theta) \tag{2.2}$$

that enters the likelihood function (1.1) was used. The model comprises six ODEs and 10 unknown parameters, including one nuisance parameter; for details about the model equations see the electronic supplementary material. All parameters are positive by definition, hence the logarithmic space yields a flat metric [12]. Since no empirical evidence about the values of the nine parameters of interest was available, the prior was chosen uninformative.

Radioactively labelled Epo facilitated the measurement of Epo concentration in the extracellular medium and of Epo bound to the cell membrane; see electronic supplementary material, table S1. The MAP estimates of the model parameters were obtained by numerical optimization; see electronic supplementary material, table S2. The agreement of model outputs and experimental data for this initial experimental set-up is shown in figure 2*c*. In accordance with the experimental data, the model describes the binding of Epo to the Epo receptor on the membrane, internalization and recycling of the Epo–EpoR complex. As first step of the proposed joint approach, the posterior PDF was screened for non-identifiability using the profiling method.

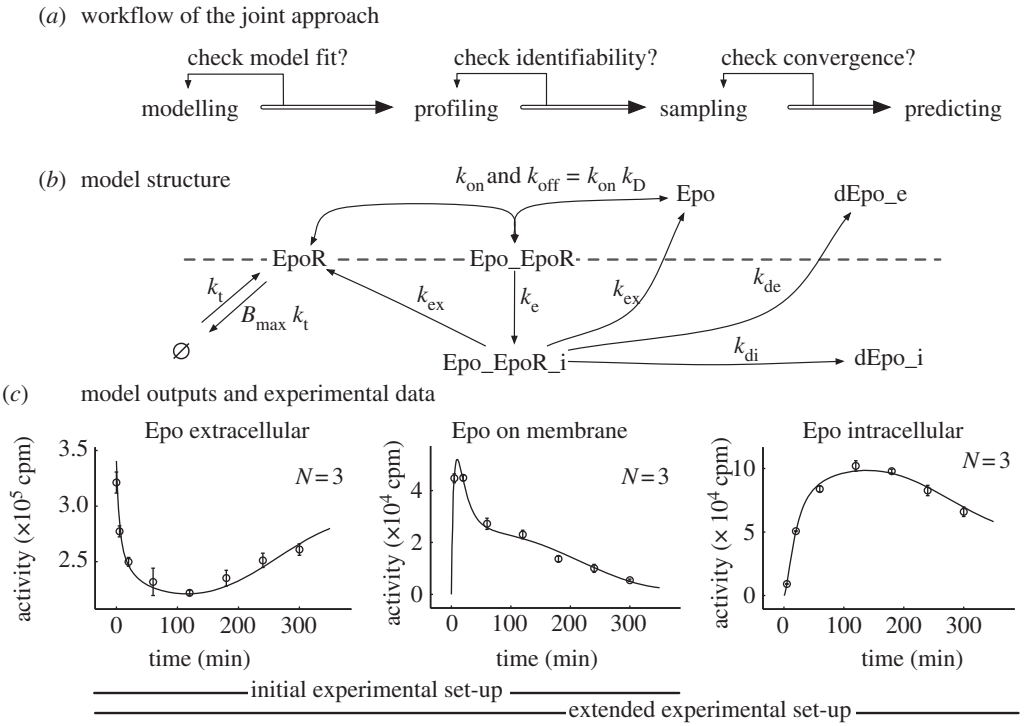

Figure 2. (*a*) Workflow of the joint approach that combines profile likelihood and MCMC sampling methods: the first step is setting up an appropriate mathematical model (*b*) that contains the available prior knowledge and explains the experimental data sufficiently well (*c*). The profile-likelihood approach allows one to detect and to resolve non-identifiability by experimental design. Therefore, it helps to ensure that the posterior PDF is proper and sufficiently well constrained. Subsequently, MCMC sampling can be applied securely. Finally, the uncertainty in model predictions can be assessed realistically using the obtained samples. (*b*) The structure of the ODE model describing Epo and EpoR interactions. The dashed line indicates the cell membrane; the part above is the outside, the part below is the inside of the cell. (*c*) The agreement of experimental data and model output for MAP parameter values for both experimental set-ups is displayed. The extended set-up was derived by experimental design considerations, see [19].

(a) Identifiability analysis using posterior profiles

The profile-likelihood approach for identifiability analysis [7] can be translated into a profile posterior approach. In analogy to equation (1.2), profiling can be applied to the unnormalized posterior PDF by defining the profile posterior

$$\mathrm{PP}(\theta_i|y) = \max_{\theta_{j\neq i}}[P(\theta|y)] \tag{2.3}$$

(cf. figure 1). Technical details on the implementation of the methods are given in the electronic supplementary material notes. For the initial experimental set-up, the profiles of the posterior PDF reveal four structurally non-identifiable, two practically non-identifiable and three identifiable parameters. One of each case is displayed as full red lines in figure 3*a*; the remaining profiles are shown in the electronic supplementary material, figure S1.

(b) Markov chain Monte Carlo sampling

The results of the posterior profiling approach indicate that the posterior PDF for the initial experimental set-up is not proper, i.e. the posterior PDF cannot be normalized to one and is therefore not a valid PDF. In this situation, an MCMC sampling cannot converge and hence

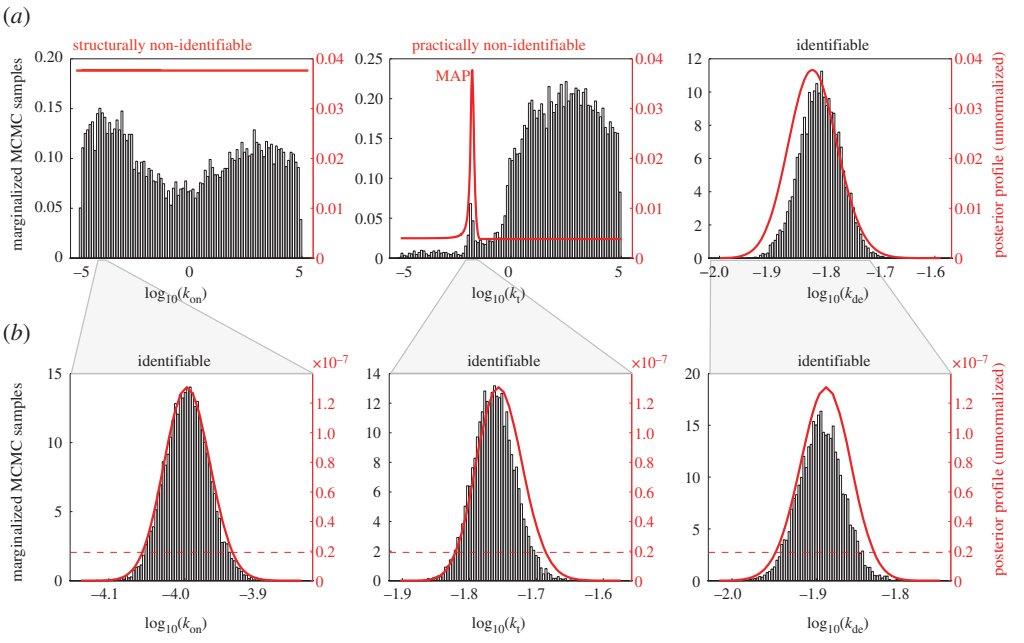

Figure 3. Comparison of profiling and sampling results. (*a*) For the initial experimental set-up, the posterior profiles indicated by full red lines reveal that parameter k_{on} is structurally non-identifiable, parameter k_t is practically non-identifiable and parameter k_{de} is identifiable. The histograms display the marginalized MCMC samples obtained by the MMALA algorithm. For the identifiable parameter k_{de}, results of both profiling and sampling agree quite well. Also for the structurally non-identifiable parameter k_{on} the agreement is acceptable. For the practically non-identifiable parameter k_t, the results are substantially different. The profile shows that the MAP point is located at $\log_{10}(k_t) \approx -1.8$. However, the lion's share of the marginalized MCMC samples propose $\log_{10}(k_t)$ to be > 0. (*b*) Taking into account more experimental data, the posterior profiles for the extended experimental set-up indicate that all parameters are now identifiable. The results of MCMC sampling and profiling are in good agreement. Interestingly, the MCMC samples for parameter k_t for the extended set-up are localized close to the MAP point of the initial set-up; note the different scales on the *x*-axis for (*a*) and (*b*). The dashed red lines indicates the threshold Δ_α that can be used to assess likelihood-based confidence intervals. (Online version in colour.)

gives inaccurate results [29]. In order to allow for a comparison between profiling and MCMC sampling, the prior PDF was restricted artificially to a uniform distribution with support from 10^{-5} to 10^{+5}. The resulting posterior PDF is now proper and the Markov chain can in principle converge. It is important to note that the results of MCMC sampling can potentially be biased due to the artificial specification of this prior PDF. The new posterior PDF, though proper, still exhibits a complicated structure. The probability mass is distributed widely in the 10-dimensional parameter space; see the posterior profiles in figure 3*a*. In this situation, the SIM sampling algorithm, which uses a scale identity matrix for generating proposals, was too inefficient and did not yield reasonable results within an acceptable time; see the electronic supplementary material, figures S1–S3.

To improve efficiency, the MMALA algorithm, which takes into account the local geometry of $P(\theta|y)$, was used [26]. Figure 3*a* shows the histograms of marginalized MCMC samples. The remaining results are displayed in the electronic supplementary material, figures S4–S6. For the identifiable parameter k_{de} and for the structural non-identifiable parameter k_{on}, the results of the MCMC sampling are in good agreement with the results of the profiling. For the practically non-identifiable parameter k_t, the results are substantially different. The profile shows that the MAP point is located at $\log_{10}(k_t) \approx -1.8$. However, the lion's share of the marginalized MCMC samples propose $\log_{10}(k_t)$ to be >0. In this region, the posterior profile reveals a plateau.

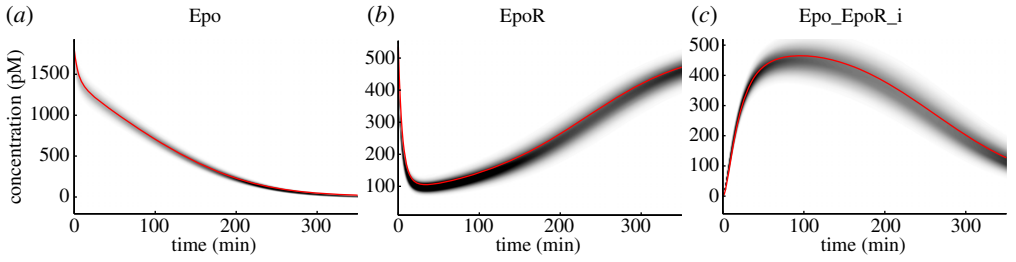

Figure 4. Propagation of parameter uncertainties to model predictions: the posterior samples that were acquired by MCMC sampling for the extended set-up (figure 3*b*) can now be used to assess prediction uncertainties accurately. The figure shows the predicted dynamics of three molecular components that could not be measured directly. The full red line shows the prediction that corresponds to the MAP estimate of the parameters. The greyscale displays the posterior PDF of the prediction. (Online version in colour.)

(c) Taking into account additional experimental data

The results of both posterior profiling and MCMC sampling indicate substantial uncertainty in the parameter estimates for the initial experimental set-up. Based on the results of the profiling approach, additional experiments were suggested; see [19] for details. The target of the experimental design was to resolve non-identifiabilities. The additional experimental data were included in the estimation procedure, yielding an extended experimental set-up (figure 2*c*). Figure 3*b* shows the recomputed posterior profiles. They indicate, by tailing out to zero, that the identifiability problems were resolved for all parameters. The remaining profiles are shown in the electronic supplementary material, figure S7.

As a consequence, the posterior PDF is now proper without artificial assumptions on the prior PDF. In this situation, the SIM sampling algorithm was still too inefficient and did not yield reasonable results within an acceptable time; see the electronic supplementary material, figures S7–S9. To improve efficiency, the MMALA algorithm, which takes into account the local geometry of $P(\theta|y)$, was used again [26]. Figure 3*b* shows the histograms of marginalized MCMC samples. The results for the remaining parameters are shown in the electronic supplementary material, figures S10–S12. For all parameters, the results of the MCMC sampling and posterior profiling are now in good agreement. The MCMC samples for parameter k_t that showed substantial difference between profiles and MCMC samples for the initial set-up are now in good agreement as well. Interestingly, the MCMC samples for parameter k_t for the extended set-up are localized close to the MAP point of the initial set-up. This suggests that the large probability mass of MCMC samples for values of $\log_{10}(k_t) > 0$ obtained for the initial set-up was misleading.

MCMC sampling results are now reliable and allow one to consider the full high-dimensional posterior PDF for inference; see e.g. the nonlinear correlation between parameter k_{di} and k_{de} shown in the electronic supplementary material, figure S13. Using the generated MCMC samples, the parameter uncertainties contained in the posterior PDF can be propagated accurately to the prediction of the dynamics of molecular compounds that are not accessible by experiments directly (figure 4).

3. Summary

We introduced a joint approach that combines frequentist profiling and efficient Bayesian MCMC sampling methods. The proposed approach uses the analysis of posterior profiles that helps to determine the identifiability of the model parameters. Non-identifiability results in a posterior distribution that is not proper, i.e. that cannot be normalized to one. However, a proper posterior distribution is required for convergence of the MCMC sampling. The results of the posterior profiling can be used for experimental design that helps resolve non-identifiabilities

iteratively. Consequently, this ensures that the posterior distribution is proper. Having ensured identifiability of the parameters, the MCMC sampling results are reliable and can be used to propagate the uncertainty of parameter estimates to model predictions. Using an application from cell biology, we showed that this approach enables one to obtain accurate results.

We compared the results of posterior profiling and marginalizing over the MCMC samples for two stages of experimental set-up: an initial set-up that contains non-identifiabilities and an extended set-up where all parameters are identifiable. A uniform prior distribution ensured that the posterior is proper despite non-identifiability for the initial set-up. For the extended set-up, the results of posterior profiling and marginalizing over the MCMC samples are in good agreement. For the initial set-up, substantial differences between profiles and MCMC samples are observed. Interestingly, the profiles for the initial set-up reflect the posterior distribution for the extended set-up much better than the MCMC samples of the initial set-up. This indicates that MCMC sampling results in the presence of non-identifiability can be inaccurate despite a proper posterior distribution.

This work was supported by the German Federal Ministry of Education and Research (Virtual Liver (grant no. 0315766), LungSys (grant no. 0315415E) and FRISYS (grant no. 0313921)), the European Union (CancerSys (grant no. EU-FP7 HEALTH-F4-2008-223188)), the Initiative and Networking Fund of the Helmholtz Association within the Helmholtz Alliance on Systems Biology (CoReNe HMGU), and the Excellence Initiative of the German Federal and State Governments (EXC 294).

References

1. Feroz F, Cranmer K, Hobson M, Ruiz de Austri R, Trotta R. 2011 Challenges of profile likelihood evaluation in multi-dimensional SUSY scans. *J. High Energy Phys*. **2011**, 42. (doi:10.1007/JHEP06(2011)042)
2. Smith RL, Tebaldi C, Nychka D, Mearns LO. 2009 Bayesian modeling of uncertainty in ensembles of climate models. *J. Am. Stat. Assoc.* **104**, 97–116. (doi:10.1198/jasa.2009.0007)
3. Moncho R, Caselles V, Chust G. 2012 Alternative model for precipitation probability distribution: application to Spain. *Clim. Res.* **51**, 23–33. (doi:10.3354/cr01055)
4. Bachmann J, Raue A, Schilling M, Timmer J, Klingmüller U. 2011 Division of labor by dual feedback regulators controls JAK2/STAT5 signaling over broad ligand range. *Mol. Syst. Biol.* **7**, 516. (doi:10.1038/msb.2011.50)
5. Robert CP, Casella G. 2004 *Monte Carlo statistical methods*. Berlin, Germany: Springer.
6. Walter E, Pronzato L. 1997 *Identification of parametric models from experimental data*. Berlin, Germany: Springer.
7. Raue A, Kreutz C, Maiwald T, Bachmann J, Schilling M, Klingmüller U, Timmer J. 2009 Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics* **25**, 1923–1929. (doi:10.1093/bioinformatics/btp358)
8. Fisher RA. 1922 On the mathematical foundations of theoretical statistics. *Phil. Trans. R. Soc. Lond. A* **222**, 309–368. (doi:10.1098/rsta.1922.0009)
9. Coleman TF, Li Y. 1996 An interior, trust region approach for nonlinear minimization subject to bounds. *SIAM J. Optim.* **6**, 418–445. (doi:10.1137/0806023)
10. Press WH, Teukolsky SLA, Flannery BNP, Vetterling WMT. 1990 *Numerical recipes: FORTRAN*. Cambridge, UK: Cambridge University Press.
11. Seber GAF, Wild CJ. 2003 *Nonlinear regression*. New York, NY: Wiley.
12. Box GEP, Tiao GC. 1973 *Bayesian inference in statistical analysis*. New York, NY: Wiley.
13. Allison J *et al.* 1992 The detector simulation program for the OPAL experiment at LEP. *Nucl. Instrum. Methods Phys. Res. A* **317**, 47–74. (doi:10.1016/0168-9002(92)90593-S)
14. Chis O-T, Banga JR, Balsa-Canto E. 2011 Structural identifiability of systems biology models: a critical comparison of methods. *PLoS ONE* **6**, e27755. (doi:10.1371/journal.pone.0027755)
15. Raue A, Kreutz C, Maiwald T, Klingmüller U, Timmer J. 2011 Addressing parameter identifiability by model-based experimentation. *IET Syst. Biol.* **5**, 120–130. (doi:10.1049/iet-syb.2010.0061)
16. Sedoglavic A. 2002 A probabilistic algorithm to test local algebraic observability in polynomial time. *J. Symbol. Comput.* **33**, 735–755. (doi:10.1006/jsco.2002.0532)

17. Karlsson J, Anguelova M, Jirstrand M. 2012 An efficient method for structural identifiability analysis of large dynamic systems. In *Proc. 16th IFAC Symp. on System Identification (SYSID2012), Brussels, Belgium, 11–13 July*.
18. Murphy SA, van der Vaart AW. 2000 On profile likelihood. *J. Am. Stat. Assoc.* **95**, 449–485. (doi:10.1080/01621459.2000.10474219)
19. Raue A, Becker V, Klingmüller U, Timmer J. 2010 Identifiability and observability analysis for experimental design in non-linear dynamical models. *Chaos* **20**, 045105. (doi:10.1063/1.3528102)
20. Meeker WQ, Escobar LA. 1995 Teaching about approximate confidence regions based on maximum likelihood estimation. *Am. Stat.* **49**, 48–53. (http://www.jstor.org/stable/2684811)
21. Venzon DJ, Moolgavkar SH. 1988 A method for computing profile-likelihood-based confidence intervals. *Appl. Stat.* **37**, 87–94. (doi:10.2307/2347496)
22. Neale MC, Miller MB. 1997 The use of likelihood-based confidence intervals in genetic models. *Behav. Genet.* **27**, 113–120. (doi:10.1023/A:1025681223921)
23. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. 1953 Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**, 1087. (doi:10.1063/1.1699114)
24. Hastings WK. 1970 Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**, 97. (doi:10.1093/biomet/57.1.97)
25. Roberts GO, Gelman A, Gilks WR. 1997 Weak convergence and optimal scaling of random walk Metropolis algorithms. *Ann. Appl. Probab.* **7**, 110–120. (doi:10.1214/aoap/1034625254)
26. Girolami M, Calderhead B. 2011 Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *J. R. Stat. Soc. B (Stat. Methodol.)* **73**, 123–214. (doi:10.1111/j.1467-9868.2010.00765.x)
27. Efron B. 2005 Bayesians, frequentists, and scientists. *J. Am. Stat. Assoc.* **100**, 1–5. (doi:10.1198/016214505000000033)
28. Bayarri MJ, Berger JO. 2004 The interplay of Bayesian and frequentist analysis. *Stat. Sci.* **19**, 58–80. (doi:10.1214/088342304000000116)
29. Gelfand AE, Sahu SK. 1999 Identifiability, improper priors, and Gibbs sampling for generalized linear models. *J. Am. Stat. Assoc.* **94**, 247–253. (doi:10.1080/01621459.1999.10473840)
30. Hobert JP, Casella G. 1996 The effect of improper priors on Gibbs sampling in hierarchical linear mixed models. *J. Am. Stat. Assoc.* **91**, 1461–1473. (doi:10.1080/01621459.1996.10476714)
31. Aprile E *et al.* 2011 Dark matter results from 100 live days of XENON100 data. *Phys. Rev. Lett.* **107**, 131302. (doi:10.1103/PhysRevLett.107.131302)
32. Becker V, Schilling M, Bachmann J, Baumann U, Raue A, Maiwald T, Timmer J, Klingmueller U. 2010 Covering a broad dynamic range: information processing at the erythropoietin receptor. *Science* **328**, 1404–1408. (doi:10.1126/science.1184913)