Available online at www.sciencedirect.com
ScienceDirect
Procedia Computer Science 1 (2012) 1645–1653
Procedia Computer Science
www.elsevier.com/locate/procedia

International Conference on Computational Science, ICCS 2010

Towards experimental design using a Bayesian framework for parameter identification in dynamic intracellular network models

Andrei Kramer and Nicole Radde

Institute for Systems Theory and Automatic Control, University of Stuttgart, Pfaffenwaldring 9, 70550 Stuttgart, Germany

Abstract

Biological measurements of intracellular regulation processes are typically noisy, and time resolution is low. In practice often only steady state measurements of perturbation experiments are available. Since data acquisition is expensive, a framework for experimental design that allows the inclusion of prior knowledge and takes uncertainty into account is highly desirable. We introduce a framework for the experimental design problem to infer parameters from steady state observations of intracellular networks. Our network model consists of (nonlinear) ordinary differential equations based on chemical reaction kinetics. We consider sets of structural perturbation experiments, that is, steady state measurements of the system subject to gene knockout or mutations. The model is stochastically embedded by introducing Gaussian measurement errors. This allows the application of a statistical Bayesian framework and usage of information-theoretic measures for experimental design. We propose to choose the optimal experiments with respect to identifiability of model parameters by maximizing the information content of the expected outcome, measured as the entropy of the posterior distributions. In this setting the posterior has no closed form and an analysis requires efficient sampling methods. We introduce a simulation-based experimental design framework for the identification of network parameters with an efficient entropy estimation approach. First results are shown on a network model for secretory pathway control. Secretion of proteins from cells involves the budding of vesicles at the Golgi. For this process PKD activity is central.

Keywords: experimental design, Bayesian learning, parameter estimation, secretory pathway

1. Introduction

A main goal in systems biology is to reveal regulatory mechanisms underlying complex behavior of intracellular networks at the molecular level. For this purpose, differential equation models based on chemical reaction kinetics have become a standard model class to describe the dynamic of signaling networks, metabolic pathways or regulatory

Email address: {kramer,radde}@ist.uni-stuttgart.de ()

1877-0509
doi:10.1016/j.procs.2010.04.184

mechanisms. In order to develop reliable predictions with these models, parameters have to be estimated from experimental data. This inverse problem is usually formulated as an optimization problem. Since the data available for this purpose is usually scarce and noisy, and time resolution is low [1], the optimization problem is ill-posed, and the performance of standard methods such as least-squares or maximum likelihood estimation is poor [2, 1]. Data acquisition for the identification of intracellular processes is often expensive and time-consuming. Therefore, a model-based optimal design of experiments with respect to identifiability of parameters or, more generally, model selection, can highly facilitate the development of reliable models by a pre-selection of the most informative experiments.

Our focus is on Bayesian frameworks for experiment planning. Such a statistical framework has several advantages in our setting. First, it uses a stochastic modeling framework, which generally allows to integrate intrinsic stochastic effects and measurement errors. Second, Bayesian approaches can include prior knowledge and beliefs about the outcome in terms of an a-priori distribution over model parameters. Since this prior distribution does not depend on the data, it can stabilize the solution with respect to data noise and avoids overfitting [2]. As such additional knowledge is often available for real biological systems, the Bayesian framework is generally appropriate here. Third, the result is not only a point estimate, but a posterior distribution over model parameters, which can be analyzed in different respects. It contains information about the reliability of the estimated parameter values in terms of its variance and about the degree of identifiability. Furthermore, stochastic models can generally cope with unobserved states, missing values and hidden variables, which is typically encountered in biological data sets.

In design problems a cost functional is optimized with respect to a set of potential experiments (for an introduction into design problems and a discussion of different cost functions we refer to [3]). A statistical framework allows the application of information-theoretic measures for this cost function such as the entropy of the posterior distribution or Kullback-Leibler divergence (KLD).

Using a Bayesian framework in order to estimate parameters of nonlinear differential equations is particularly challenging, because the likelihood function does not have a conjugate prior distribution and thus the posterior cannot be computed in closed form but has to be approximated by Markov Chain Monte Carlo (MCMC) sampling strategies [1]. When evaluating the cost function for a given experimental setting, this requires estimating the posterior densities from these samples, which can for example be done by using kernel density estimators. The relation between the parameters to be estimated and the observations y can be quite complicated. For example, if steady states are measured under different perturbations, these are implicitly described by the solution set $\bar{x}(\theta)$ of the implicit equation $f(\bar{x}, \theta) = 0$. This usually requires simulations with an appropriate numerical integration method in each optimization step, which can be very expensive computationally.

Bayesian experimental design has recently been applied to biological networks in different settings. For example, Steinke et al. [1] use linear models with independently and identically distributed Gaussian noise and sparsity enforcing Laplace priors for the identification of gene regulatory networks from steady state measurements. This approach is a compromise between sparsity and computational tractability. On the one hand, the log-posterior density is concave in this setting and thus has a single maximum which allows using a normal approximation, for which many quantities such as the relative entropies can easily be calculated. On the other hand, compared to Gaussian prior distributions, it favors sparse solutions.

An analysis of the posterior distribution is much harder with nonlinear differential equation models. The posterior might be multimodal and standard approximating inference methods become algorithmically unstable or are limiting from the computational side, i.e. require an excessive amount of time. One of the first approaches towards nonlinear models is described in Calderhead & Girolami [4]. They investigate the effect of the number and choice of observed variables on the information content of the posterior distribution for the well-known repressilator model described in [5]. The posterior is investigated by MCMC sampling with time-varying transition kernels, and the KLD of prior and posterior distribution is approximated by using a kernel density estimator for the posterior and numerical Monte Carlo integration for calculating the expectation $E_P(\log Q)$ of the posterior logarithm $\log Q$ with respect to the prior density P.

A Bayesian experimental design method for the identification of network structures was recently introduced by Busetto & Buhmann in combination with a novel clustering technique [6]. The likelihood function was build from nonlinear stochastic differential equations, and the KLD of prior and posterior was investigated via MCMC with importance sampling. The method was evaluated on a Goodwin model. A similar framework was developed by Busetto et al. [7] for model selection of nonlinear dynamical systems, which requires averaging over the set of parameters that represents a single model. The utility function to be optimized in that work was the KLD as well.

An optimal design method was applied to guide live-cell fluorescence microscopy experiments in Bandara et al. [8], which shows that model-based experiment planning can highly improve identifiability of model parameters in a real setting and thus be a valuable tool for building accurate quantitative models with as few experiments as possible. Finally, a more general review about several aspects in experimental design and model discrimination is given by Kreutz & Timmer [9].

Those articles described indicate that a Bayesian framework is generally advantageous for the identification of biological networks, and it has also proven useful in first applications. However, while linear Gaussian models and static data are well-investigated, Bayesian design remains a fairly unexplored field in the context of nonlinear dynamic systems [7].

In this paper we introduce a Bayesian approach for calculating the expected information content of a set of experiments with respect to identifiability of model parameters. This is the basis for experimental design, which seeks to optimize this information content by a pre-selection of appropriate experiments. Information content is optimized by minimizing the entropy of the posterior distribution. The entropy is estimated by first sampling from the posterior via a Metropolis algorithm. This sample is then used to normalize the posterior by approximating the density using a multivariate Gaussian kernel density estimator with data-driven bandwidth. Within this setting we suggest a simulation-based approach to evaluate the likelihood function, which is particularly suited to deal with hidden variables and marginal likelihoods. We will show first results on the identification of two and three parameters of a regulatory network for secretory pathway control at the trans-Golgi network in mammalian cells. In contrast to previous works, we use realistic, easily obtainable data sets (relative equilibrium concentrations) and aim to limit computational costs by limiting the number of evaluations of the kernel density estimator, which in our setup is used for normalization purposes only.

2. Bayesian experimental design for parameter identification in kinetic network models

2.1. The likelihood function - Stochastic model

We consider ordinary differential equations of the form

$$\dot{x} = f(x, \theta), \quad x \in \mathbb{R}^n, \theta \in \mathbb{R}^m, f : \mathbb{R}^n_+ \times \mathbb{R}^m_+ \to \mathbb{R}^n \in C^1(x, \theta) \tag{1}$$

with a vector field f that is continuously differentiable, such that a unique and smooth solution $\phi(t, x_0, \theta)$ exists for each initial state $x_0 \in \mathbb{R}^n_+$ and parameter vectors $\theta \in \mathbb{R}^m_+$ [10]. This model describes the dynamics of a regulatory network and is based on chemical reaction kinetics, which is not further specified here. We only assume that the system has a trapping region in the positive orthant, i.e. a region that is eventually reached from all initial conditions, and that it has at least one stable steady state, which is not a strong restriction, since it is fulfilled for most chemical reaction networks. The vector θ contains the parameters to be estimated from data, such as for example rate constants or Michaelis-Menten constants.

The data set we will use for learning consists of steady state measurements under different structural perturbations $p = 1, \ldots, P$ such as for example knockout of components or manipulation of regulation strengths via mutations that affect the regulatory activity of components. In our modeling framework, these perturbations translate into changes in the values of the parameter vector θ, which we denote by θ^p. The perturbed steady states are measured in units of the steady state of the unperturbed system, $z^p(\theta^p) = \frac{\bar{x}^p(\theta^p)}{\bar{x}(\theta)}$. Note that this identification problem is generally very difficult regardless of how the differential equations are stochastically embedded or how the prior distributions are chosen in a Bayesian setting. First of all, the parameters and the observations are related via an implicit function, $f(\bar{x}^p, \theta^p) = 0$. For the inverse problem, we can think of the dependence of this output on the parameters in terms of bifurcation diagrams, which inherits several challenges for the optimization problem [2]. For example, at bifurcation points the observations can change nonsmoothly with parameters [11], which translates into a nonsmooth likelihood function in a stochastic framework. Furthermore, there might be several bifurcation branches for certain parameter values, which introduces a dependence of the long-term behavior, here the steady states, on the initial state x_0.

Measurement error is described by an additive and normally distributed noise term with mean zero and variance σ^2,

$$y^p_i = z^p_i(\theta^p) + \epsilon_i, \quad \epsilon \sim \mathcal{N}(0, \sigma^2_i). \tag{2}$$

We measure a subset I of all network components. This stochastic framework relies on the assumption that different independent error sources contribute additively to the overall noise term, and clearly any other distribution would be possible here as well. In contrast to stochastic differential equation models or Markov processes, the noise term does not affect the underlying deterministic system. Thus our modeling framework only includes noise due to measurement errors, but not intrinsic stochasticity (for relations between these model classes we refer to [12]). Given data $\mathcal{Y} = \{y_i^p : i \in I, p = 1, \ldots, P\}$ the probability of seeing the data $\mathcal{Y}$ given parameters θ reads

$$p(\mathcal{Y}|\theta) = \prod_{p=1}^{P} \prod_{i \in I} p(y_i^p | z_i^p(\theta^p)) = \prod_{p=1}^{P} \prod_{i \in I} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{1}{2} \left(\frac{y_i^p - z_i^p(\theta^p)}{\sigma_i} \right)^2 \right]. \tag{3}$$

In general this does not permit a conjugate modeling framework, and the posterior distribution has to be investigated via sampling or other approximating methods [13]. In practice the setup is often even worse, since usually not all variables can be observed. A stochastic framework provides a solution to that via marginalizing over the hidden states, that is, calculating the expectation value of the observed states with respect to the distribution over the hidden ones [14, 15, 16, 17]. This partial observability often comes along with an additional problem, the structural and practical identifiability of model parameters [18]. In our setting, structural non-identifiability comes from the fact that steady state measurements do not always contain the information required to identify all parameters. A simple example is the steady state of a reversible reaction (A⇌B) with rate constants k_1 and k_{-1} described by mass action kinetics. According to the law of mass action, the ratio of concentrations on the right and the left hand sides of the reactions equal the ratio of the rate constants. Thus steady state measurements of this system only allow to identify the ratio of the two rate constants rather than their individual values. Practical identifiability refers to the fact that sparse data does not always contain the information for a unique identification of the parameters.

Our framework allows a simple and efficient evaluation of the likelihood via simulation of the trajectory until convergence to a steady state is reached. Since the steady states of each variable are uniquely determined for any θ, and measurement noise does not feed back into the system, integration over hidden variables is not required although we cannot observe all variables of the system. In other words, the steady states $\bar{x}_i(\theta, x_0)$ do only depend on the parameters and the initial conditions and not on the steady state coordinates $\bar{x}_j$ of other network variables, and the same holds for the observations y. This comes from the fact that we assume a pure measurement noise model and simplifies analysis considerably. In order to calculate the likelihood, we determine the steady states $\bar{x}^p(\theta^p)$ via simulations, which directly allows to calculate the likelihood for any y^p via equation (3). This approach is similar to the likelihood-free methods introduced first in Marjoram et al. [19]. The only drawback is that the steady states determined in this way might depend on the initial conditions x_0 if the system has several distinct steady states. Since x_0 is generally unknown, simulations have to be conducted for several x_0 in the trapping region of the system.

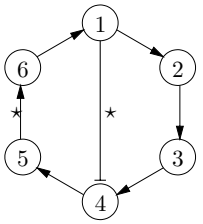
2.2. Bayesian framework - experimental design

For the prior densities over network parameters we use independent gamma distributions for each parameter,

$$f(x|b_1, b_2) = \begin{cases} \frac{b_1^{b_2}}{\Gamma(b_2)} x^{b_2 - 1} e^{-b_1 x} & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{4}$$

and fixed hyperparameters $b_{1/2}$. With this relatively simple prior we do not address the problem of learning the network topology, but assume it to be known, such that all models are of the same complexity, i.e. have the same number of parameters.

We propose to choose the most informative experiments over a finite set of perturbation experiments by minimizing the expected entropy of the posterior distribution Q. This requires estimating the posterior entropy. To this end, we generate 60.000 configurations from the posterior via Metropolis sampling: Given a current parameter vector θ^p and the data set $\mathcal{Y}$, suggest an update $\tilde{\theta}^p$ from a symmetric transition kernel $p(\tilde{\theta}^p|\theta^p)$, then use the stochastic model to simulate data $(z_i^p(\theta^p))$, and accept $\tilde{\theta}^p$ with probability $p = \min(1, \frac{p(\tilde{\theta}^p)p(\mathcal{Y}|\tilde{\theta}^p)}{p(\theta^p)p(\mathcal{Y}|\theta^p)})$. To estimate the expected entropy $E_Q[\log Q]$, use this parameter sample to approximate the posterior density with a Gaussian kernel estimator and use the approximate distribution to find the correct normalizing factor for the posterior sample using only a small subset of the sample. Evaluate the normalized sample to determine the entropy, the details are given in [20]. Compared



$$
\begin{aligned}
\dot{x}_1 &= \theta_1 x_6 - x_1 && \text{PKD} \\
\dot{x}_2 &= \theta_2 x_1 - x_2 && \text{PI(4)KIII}\beta \\
\dot{x}_3 &= \theta_3 x_2 - x_3 && \text{PI(4)P} \\
\dot{x}_4 &= \theta_4 x_3 - x_4 - \theta_8 \frac{x_1 x_4}{1 + x_4} && \text{CERT} \\
\dot{x}_5 &= \theta_5 x_4 - \theta_6 x_5 - \theta_9 \frac{x_5}{1 + x_5} && \text{ceramide} \\
\dot{x}_6 &= -\theta_7 x_6 + \theta_9 \frac{x_5}{1 + x_5} && \text{DAG.}
\end{aligned} \tag{5}
$$

Figure 1: Interactions of proteins and lipids for the regulation of secretory transport at the trans-Golgi network. *Left:* Regulatory network. Arrows indicate activation, blunt ends inhibitions, and enzyme-driven regulations are marked with a star (⋆). *Right:* Differential equation model.

to previous work [4, 6, 7], instead of calculating the entropy with the estimated posterior density, we only use this approximation to normalize the posterior, and the entropies are calculated with the normalized posterior values at the sample points. Moreover, we improved the kernel density estimation by using multivariate Gaussian kernels whose covariance is a priori estimated from a subsample of the posterior.

3. Application to a model for secretory pathway control at the trans-Golgi network in mammalian cells

3.1. Network model

Many proteins which are produced in a cell are secreted into the extracellular environment to become active for example as hormones, neurotransmitters, antibodies or digestive enzymes. These proteins play a crucial role in transmitting signals across cells and tissues, and secretory transport is highly regulated to ensure a proper activity. We focus on regulation of the activity of protein kinase D (PKD), which has been identified as a crucial regulator of secretory transport at the trans-Golgi network [21]. Specifically, PKD is critically involved in the fission of transport carriers, and was recently shown to play a dual role in CERT-mediated transport from the endoplasmic reticulum (ER) to the trans-Golgi network (TGN) [21, 22]. Here we use a differential equation model for this regulation via PKD, which is based on a description of the processes in Fugmann et al. [21]. According to [21], PKD is recruited to the TGN by diacylglycerol (DAG), which is generated via different pathways. PKD localized at the TGN activates Phosphatidylinositol 4-kinase IIIβ (PI4(K)IIIβ), which in turn increases Phosphatidylinositol(PI)4-phosphat (PI(4)P) levels at the TGN. PI(4)P recruits the transport protein CERT, which has a START domain specific for ceramide, to the Golgi complex. CERT contributes to PKD activation by providing ceramide as a precursor for further DAG production. Thus the system has a positive feedback loop. Additionally, CERT has been shown to be a PKD substrate. Active PKD decreases the affinity of CERT towards PI(4)P by direct phosphorylation, which counteracts the indirect positive regulation via PI(4)KIIIβ. The detailed mechanisms of this system are still unknown [23, 24].

From a mathematical point of view, this system has an interesting network structure, since it consists of two interrelated feedback loops. The interaction graph is shown in Figure 1. We describe this system by the set of equations on the *Right* of Figure 1. We used first order mass action kinetics for most reactions. Phosphorylation of CERT by PKD and conversion of ceramide to sphingomyelin and DAG, which is driven by the enzyme sphingomyelin-synthase, are described by Michaelis-Menten terms. The positive orthant is invariant for the flow of this system, and the vector field is continuously differentiable and has a unique fixed point in this region.

3.2. Results

Our approach was tested with simulated data of the secretory pathway model (5), for which we used one randomly selected parameter vector

$$\theta^* = (0.95, 1.42, 1.30, 1.48, 1.46, 1.20, 0.87, 1.42, 0.80)\,. \tag{6}$$

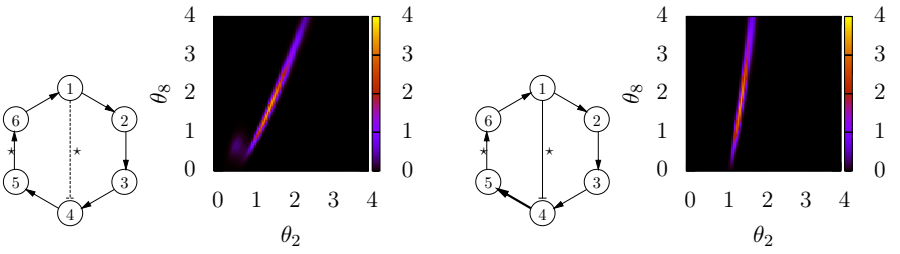


Figure 2: Comparison of two experiments; *Left:* The inhibition of CERT is deactivated; adjacent: the resulting posterior is shown. *Right:* the transport of Ceramide is enhanced by 1.5; The posterior densities clearly indicate a linear relation between both parameters. Including both perturbations into one experiment set confines the posterior to a single point (where the lines intersect). This point coincides with the true values $\theta^*_{2/8}$. The prior distribution (not shown), is a gamma distribution with mean 2 and unit variance.

parameters	θ_2	θ_8	θ_9
posterior maximum	1.28	1.61	1.43
posterior mean	1.6(2)	2.3(4)	1.4(2)
prior maximum	1.5	1.5	1.5
prior mean	2	2	2
true values θ^*	1.42	1.42	0.80

Table 1: The estimates of the unknown parameter subset. Only one perturbation was considered: $\theta_8 = 0$. The numbers in parentheses represent the standard deviation, i.e.: 1.23(1) = 1.23 ± 0.01

Among these, the two, respectively three parameters θ_2, θ_8 and $\theta_2, \theta_8, \theta_9$ are assumed to be unknown. PKD is the only measured output (with a signal to noise ratio of 10). We determined the expected information content for two data sets, with $P = 1$:

1. no inhibition of CERT: $\theta^1_8 = \theta^{\text{pert}}_8 = 0$ (Figure 2 *left*)
2. enhanced Ceramide transport: $\theta^1_5 = \theta^{\text{pert}}_5 = 1.5\,\theta_5$ (Figure 2 *right*)

The hyperparameters for the gamma-prior were set such that mean and variance resulted to 2 and 1 respectively. This choice reflects a relatively vague a-priori knowledge about the true parameter values. Figure 2 shows a comparison of two experiments. The entropies of these two densities are very similar: −0.31(5) and −0.37(6) respectively. The combined posterior (P=2; not shown) has an entropy of −4.36(2), this experiment contains full information on the true parameter values within a small error margin. Since there is a large difference between prior and posterior in both cases, the expected data is generally very informative with respect to these two parameter values, especially for parameter θ_2. The second experiment, Figure 2 (right), appears to be slightly more informative and should be prefered to the first if only one experiment is to be performed. The combined dataset however is vastly more informative, thus performing both expermients is very rewarding. The figures also show that both parameters are positively correlated. This result is obvious in the context of our model, where the parameter θ_2 describes the rate of PKD-mediated phosphorylation of PI(4)KIIIβ, and θ_8 is the strength of CERT-phosphorylation via PKD. Increasing θ_2 increases the concentration of active PI(4)KIIIβ and thus strengthens at the same time the overall effect of the positive feedback loop in the network. As phosphorylation of CERT decreases its activity, an increase in θ_8 has the opposite effect. Hence the net effect on PKD activity can partly be compensated by increasing both parameters simultaneously, which is reflected in the posterior distributions.

Figure 3 shows a comparison of prior and posterior densities for three parameter values. Here we used parallel coordinates to plot the 30% of the respective samples with the highest values. Each label on the abscissa has its own ordinate where the value of the label-symbol is noted. The points on all ordinates are connected to produce a θ

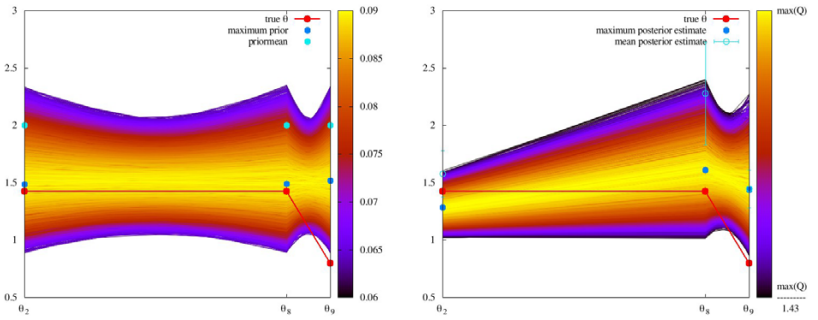


Figure 3: Prior - Posterior; On the *right*: A sample from the posterior distribution over the parameters θ_2, θ_8 and θ_9 plotted in parallel coordinates, each polyline represents a θ configuration with the color indicating the associated posterior, the y axis represents the values of the labeled entities on the x axis. Shown are the 30% samples with highest posterior values. It shows a positive correlation between θ_2 and θ_8 and a less pronounced negative correlation between θ_8 and θ_9. The expected information content of the perturbation experiment shown in Figure 2 *(left)* is highest for θ_2 and lowest for θ_9, as can be seen at the maxima and the variances of the posteriors. On the *left:* the prior distribution plotted in the same manner. The error bars attached to the posterior-mean represent the standard deviation of the whole distribution, not the standard deviation of the mean estimator. The values of the colormap are unnormalized.

configuration polyline. These plots also contain information about the correlation of parameters. An hourglass shape between two ordinates indicates negative correlation. positive correlations result in nonintersecting line sets.

A comparison of the maximum posterior estimates with the true values (blue dots) indicates that the expected information content of the data is highest for the parameter θ_2. The parameter θ_9 describes the maximal DAG production rate from conversion of ceramide, which is triggered by a sphingomyelin synthase. While the posterior gives an almost perfect estimate for θ_2, i.e. the maximum is close to the true value and the variance in θ_2 direction is low, both the distance between estimated maxima and true values and the posterior variance are higher for θ_8 and θ_9. Moreover, while the positive correlation between θ_2 and θ_8 is still clearly visible here, there is a tendency for a negative correlation between θ_8 and θ_9. In the context of the secretory pathway model, the latter correlation is more subtle and not directly obvious from the model any more. First, both parameters are part of the negative feedback loop. Thus a stronger inhibition of CERT by increasing the maximal phosphorylation rate θ_8 has a negative effect on x_1 at first glance. In contrast, increasing the conversion rate of ceramide to DAG is expected to have a positive effect on x_1. However, this simple picture does not consider that a negative effect on PKD activity eventually turns into an unpredictable effect on CERT that depends on both the indirect positive and direct negative regulation, and it does also not take into account that the enzyme-driven reactions might be at its saturation, such that if for example ceramide raises, this might not at all have an effect on PKD.

4. Discussion and conclusions

In this paper we introduced a statistical Bayesian method for experimental design with respect to identifiability of model parameters. We consider differential equation models that describe the dynamic of regulatory networks, and the data used for learning correspond to steady state measurements of network variables under different perturbations. The described procedure provides an optimization criterion for experiment evaluation: the expected information content of the posterior distribution, measured as the inverse of the respective entropy. This value is estimated using MCMC sampling in combination with multivariate Gaussian kernel density estimators to approximate the posterior density. In this setting, the likelihoods and the expectation value of the entropy are both approximated via Monte Carlo simulations. We show first results for a model of secretory pathway control at the trans-Golgi network via the protein kinase D and demonstrate that an analysis of the posterior density gives, besides point estimates, also information about correlations across variables and identifiability.

Our method is generally able to deal with nonlinear systems and the problems typically encountered in biological models such as partial observability, hidden states and sparse and noisy data sets. Moreover, such a statistical learning

approach provides not only point estimates, but also allows for an estimation of uncertainties. Furthermore, it can in principle be applied to higher-dimensional problems, which is one of the major goals in the future. We expect that this step will currently be limited by excessively long computation times, since the sample size needed for an adequate representation of the posterior increases exponentially with the dimension, which requires for example further optimization with respect to the sampling strategies. Further issues will probably be the potentially complex relation between parameters and steady state measurements, which might translate into multiple modes, discontinuities and ruggedness of the posterior function. Instead of prior densities with fixed hyperparameters a fully Bayesian approach could be applied as well. These problems are currently investigated in our group. We are furthermore on the way to extent the secretory pathway model by collecting experimental data available in the literature, which we are going to use for parameter identification. We plan to support model building and parameter identification by applying the introduced design methods in cooperation with our experimental partners, the Institute for Immunology and Cell Biology at the University of Stuttgart.

Acknowledgements

We thank Dr. Angelika Hausser and Dr. Monilola Olayioye (Institute of Cell Biology and Immunology, University of Stuttgart) for helpful discussions about the model of secretory pathway control and the German Research Foundation (DFG) for financial support of the project within the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart.

References

[1] F. Steinke, M. Seeger, K. Tsuda, Experimental design for efficient identification of gene regulatory networks using sparse bayesian models, BMC Systems Biology 1 (51).
[2] H. Engl, C. Flamm, P. Kügler, J. Lu, S. Müller, P. Schuster, Inverse problems in systems biology, Inverse Problems 25 (2009) 123014(51pp).
[3] K. Chaloner, I. Verdinelli, Bayesian experimental design: A review, Statist. Sci. 10 (3) (1995) 273–304.
[4] B. Calderhead, M. Girolami, Sloppy parameters in oscillatory systems with unobserved species, in: Proc. of the 5th International Workshop on Computational Systems Biology (WCSB08), Leipzig, 2008, pp. 21–24.
[5] M. Elowitz, S. Leibler, A synthetic oscillatory network of transcriptional regulators, Nature 403 (2000) 335–338.
[6] A. Busetto, J. Buhmann, Structure identification by optimized interventions, in: Journal of Machine Learning Research Proceedings of the 12th Int. Conf. on Artificial Intelligence and Statistics (AISTATS) 2009, Vol. 5, Clearwater Beach, Florida, USA, 2009, pp. 49–56.
[7] A. Busetto, C. Ong, J. Buhmann, Optimized expected information gain for nonlinear dynamical systems, in: Proc. of the 26th International Conference on Machine Learning, Vol. 382, Montreal, Canada, 2009, pp. 97–104.
[8] S. Bandara, J. Schlöder, R. Eils, H. Bock, T. Meyer, Optimal experimental design for parameter estimation of a cell signaling model, PLoS Comp. Biol. 5 (11) (2009) e1000558.
[9] C. Kreutz, J. Timmer, Systems biology: experimental design, FEBS Journal 276 (2009) 923–942.
[10] J. Guckenheimer, P. Holmes, Nonlinear oscillations, dynamical systems, and bifurcations of vector fields, no. 42 in Applied Mathematical Sciences, Springer series, 1990.
[11] N. Radde, The impact of time-delays on the robustness of biological oscillators and the effect of bifurcations on the inverse problem, Eurasip Journal on Bioinformatics and Systems Biology 2009. doi:10.1155/2009/327503.
[12] N. Radde, N. Bar, A. Tresch, A comparison of likelihoods for dynamic stochastic models of biological networks, in: T. Manninen, C. Wiuf, H. Lähdesmäki, M. Grzegorczyk, J. Rahnenführer, M. Ahdesmäki, M.-L. Linne, O. Yli-Harja (Eds.), WCSB Conference proceedings, Aarhus, Denmark, 2009, pp. 131–134.
[13] M. Beal, Z. Ghahramani, A variational Bayesian EM algorithm for incomplete data: With application to scoring graphical model structures, Bayesian Statistics 7 (2003) 453–463.
[14] M. Beal, F. Falciani, Z. Ghahramani, C. Rangel, D. Wild, A Bayesian approach to reconstructing genetic regulatory networks with hidden factors, Bioinformatics 21 (3) (2005) 349–356.
[15] A. Kramer, N. Radde, A statistical framework for noise separation in dynamic models of intracellular networks, proc. of International Conference on Computing Anticipatory Systems (Casys09), August 3-8th, 2009, Liège, Belgium, in print.
[16] M. Quach, N. Brunel, F. d'Alché Buc, Estimating parameters and hidden variables in non-linear state-space models based on odes for biological networks inference, Bioinformatics 23 (23) (2007) 3209–3216.
[17] C. Sabatti, G. James, Bayesian sparse hidden components analysis for transcription regulation networks, Bioinformatics 22 (6) (2006) 739–746.
[18] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, J. Timmer, Structural and practical identifiability anaylsis of partially observed dynamical models by exploiting the profile likelihood, Bioinformatics 25 (15) (2009) 1923–1929.
[19] P. Marjoram, J. Molitor, S. Tavaré, Markov chain Monte Carlo without likelihoods, Proc. Natl. Acad. Sci. 100 (26) (2003) 15324–15328.
[20] CCA, Computation of the posterior entropy in a Bayesian framework for parameter estimation in biological networks, submitted.
[21] T. Fugmann, A. Hausser, P. Schöffler, S. Schmid, K. Pfizenmaier, M. Olayioye, Regulation of secretory transport by protein kinase d-mediated phosphorylation of the ceramide transfer protein, J. Cell Biol. 178 (1) (2007) 15–22.

[22] A. Hausser, S. Märtens, G. Link, K. Pfizenmaier, Protein kinase D regulates vesicular transport by phosphorylation and activation of phosphatidylinositol-4 kinase iii β at the Golgi, Nat. Cell Biol. 7 (9) (2005) 880–886.
[23] K. Hanada, K. Kumagai, N. Tomishige, T. Yamaji, CERT-mediated trafficking of ceramide, Biochimica et Biophysica Acta 1791 (2009) 684–691.
[24] R. Perry, N. Ridgway, Molecular mechanisms and regulation of ceramide transport, Biochimica et Biophysica Acta 1734 (2005) 220–234.