Distribution of a bifurcation parameter in a genetic network with uncertain parameters

Simon Borger, Wolfram Liebermeister, Edda Klipp
{borger|lieberme|klipp}@molgen.mpg.de
MPI for Molecular Genetics
Ihnestr. 73, D-14195 Berlin

Abstract

Mathematical models of dynamic biochemical processes are characterised by their network structure and their parameters. Since parameters are often not or not sufficiently well determined, we propose to use a Monte-Carlo sampling approach. Parameters are drawn from a distribution to study the resulting distribution of qualitative and quantitative properties of the system. A genetic circuit that may exhibit oscillations is used as example.

1 Introduction

Models with uncertain parameters

Mathematical modelling and dynamic simulation of metabolic networks, signal transduction cascades, and genetic networks is a central theme in systems biology and is attracting growing interest [1] [2]. Building mathematical models comprises mainly two aspects: (a) deciding on the model structure and (b) estimating the involved parameter values. The model structure reflects the investigated processes and the respective reaction network. Given a model structure, parameter estimation remains the limiting step in the modelling and simulation of biological systems. Parameter estimation for non-linear dynamic systems has been studied comprehensively [3][4], but in most cases, the amount of data available for a specific process and the respective model

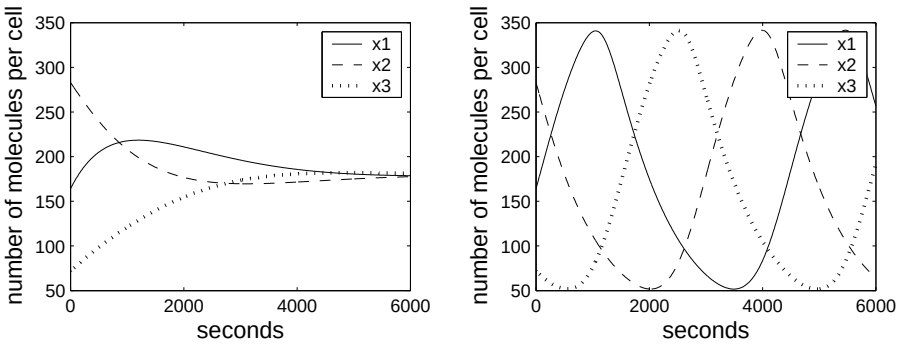

Figure 1: Dynamics of the repressilator below and above the Hopf bifurcation. The Hill coefficient h is a bifurcation parameter. For $h < h_{\mathrm{crit}}$, the system always tends to a steady state, while for $h > h_{\mathrm{crit}}$, oscillations can arise.

is by far too sparse to determine the parameters reliably. Moreover, a part of the system parameters may not be identifiable from the set of experiments, or parameter values remain unreliable due to measurement errors, dependence on experimental conditions, and individual variations in cell composition and state.

Therefore, we would like to change the perspective and investigate to which extent the network structure determines the quantitative and qualitative behaviour of the system - irrespective, or almost irrespective, of the parameter values. In this situation, an approach based on uncertain parameters can be helpful: one may assess the effects of measurement errors, find results that remain valid for a wide range of parameter values, and derive probabilities for different model outcomes. Besides this, probabilistic models may also be employed to describe the natural variability of parameters when studying the variability and robustness of biological systems.

Uncertainties in the qualitative and quantitative behaviour

The presented approach is to assume that the parameters are not fixed or estimated, but drawn from a distribution, which has to be specified [5] [6]. This parameter distribution, together with the fixed network structure, leads to distributions of the observables. Repeated simulations of the model, with parameters drawn from their distributions, will yield many realisations of the

dynamic profiles and observable quantities of the system. If the distribution of a quantity is sharp, we conclude that this quantity is strongly determined by the network structure, at least for the ensemble of parameters considered. This allows to study which kind of quantitative and qualitative behaviour can be expected from the model.

If the parameter variance is zero, then also the observables are, of course, exactly determined. For small parameter variability, the width of the variable distribution can be computed by expanding the variables around the mean parameter values. To first order, a log-normal parameter distribution leads to a corresponding log-normal distribution of the observables. The variance depends on the sensitivities, that is, the slopes of logarithmic observables seen as functions of the logarithmic parameters [6].

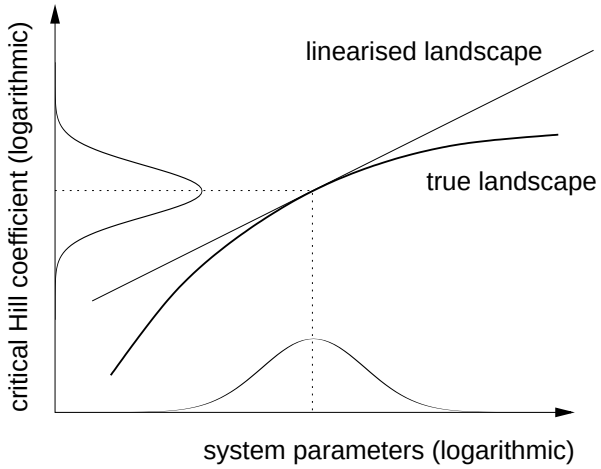

Figure 2: The critical Hill coefficient depends on the parameters α, β, and k. The diagram shows the function $h_{\mathrm{crit}}(\alpha, \beta, k)$ (schematic). The parameters α, β, and k (schematically represented by the abscissa) are drawn from a random distribution (a Gaussian distribution for logarithmic values). By computing the value of h_{crit} for each realisation, we can sample distribution of this parameter (plotted at the ordinate). A linear function $h_{\mathrm{crit}}(\alpha, \beta, k)$ would, again, yield a log-normal distribution.

Monte Carlo simulations with random parameters have been used to compute the distributions of metabolic concentrations, metabolic fluxes, control coefficients, and other variables [5] [7]. The same approach has been applied to gene regulatory circuits [8] and a MAP kinase cascade [9]. Here we focus on another feature of dynamic systems, namely a the location of a bifurcation

point. We study a simple genetic network as has been analysed by Elowitz and Leibler [10]. It shows a parameter-dependent transition from stable behaviour to oscillations, known as a Hopf bifurcation. The Hill coefficient in the kinetic equations is a critical parameter. By sampling all other parameters from predefined distributions, we compute the distribution of the critical value.

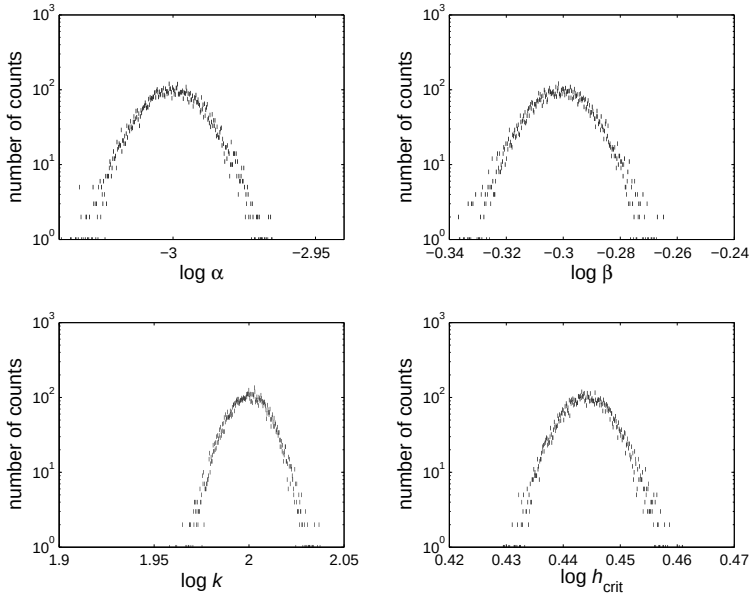

Figure 3: Three of the four plots show histograms of the distribution the parameters α, β and k were drawn from. The width was in each case $\sigma = 0.01$. The bottom right plot shows the histogram for the resulting distribution of the value h_{crit}.

2 Analysis of the symmetric repressilator

The model

We study a simplified version of the repressilator described in [10]. Three genes form a negative-feedback loop, each gene repressing the transcription of the gene it acts on. Their mRNA concentrations $x_i, i = 1, 2, 3$ we model

by the differential equations

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = -\alpha x_i + \frac{\beta}{1 + (\frac{x_{j(i)}}{k})^h} \tag{1}$$

where $j(1) = 3, j(2) = 1, j(3) = 2$ indicates the upstream gene, respectively. The parameters α and β describe the strength of degradation and full transcription, while k and h characterise the mutual repression of transcription. Here we assume that each of the genes is characterised by the same parameter values. Given values for α, β, and k, this system can show a Hopf bifurcation at a certain value $h = h_{\mathrm{crit}}(\alpha, \beta, k)$: for values $h < h_{\mathrm{crit}}$, the system has a stable steady state, while for $h > h_{\mathrm{crit}}$, the steady state becomes unstable and a stable limit cycle shows up (see Figure 1). For the values $\alpha = 0.001$, $\beta = 0.5$, $k = 100$, we find $h_{\mathrm{crit}} \approx 2.78$.

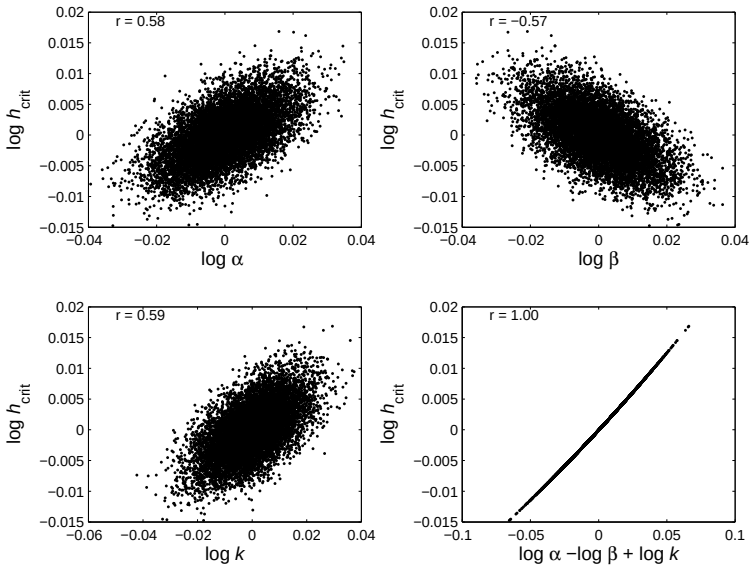

Figure 4: Correlation plots for the three parameters α, β and k versus h_{crit}. The bottom right plot shows the strong correlation between $\ln\alpha - \ln\beta + \ln k$ and h_{crit}.

Parameter sampling and bifurcation analysis

The parameters α, β, and k were drawn from log-normal distributions such that $\log_{10}\alpha$, $\log_{10}\beta$, and $\log_{10} k$ are independent and normally distributed

with a standard deviation of σ and mean values $\bar{\alpha} = 0.001$, $\bar{\beta} = 0.5$, and $\bar{k} = 100$, respectively. We performed 10000 simulations for distributions with widths $\sigma = 0.01$ and $\sigma = 0.2$. By drawing from these a set of parameters, each time we obtain a realisation of the dynamic system in which the Hill coefficient h is still undetermined. For each such realisation, we run a bifurcation analysis to determine the critical value $h_{\mathrm{crit}}(\alpha, \beta, k)$ of the Hill coefficient. By repeating the drawing from the parameter distributions, we can sample the distribution of critical Hill coefficients (see Figure 2). Dynamic simulations were performed with matlab. The critical parameter $h = h_{\mathrm{crit}}(\alpha, \beta, k)$ was determined using the package MATCONT [**?**] and by searching at least the interval from 1 to 50.

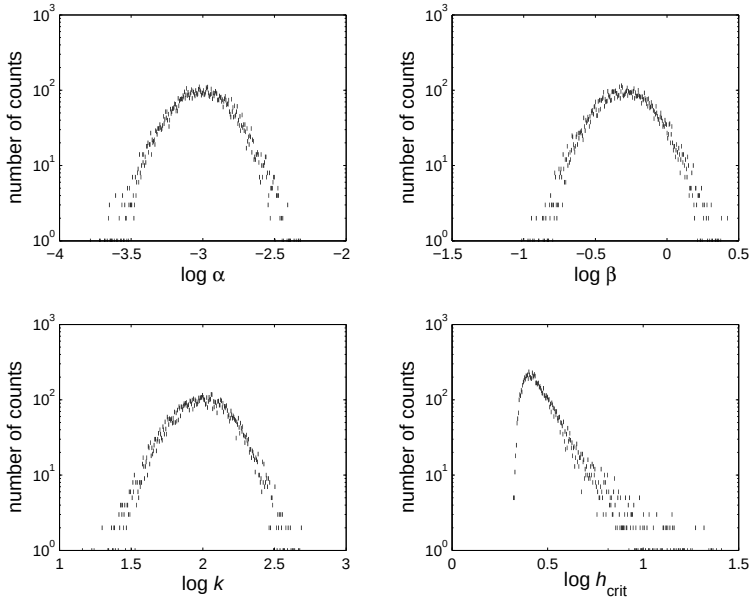

Figure 5: Three of the four plots show histograms of the distribution the parameters α, β and k were drawn from. The width was in each case $\sigma = 0.01$. The bottom right plot show the histogram for the resulting distribution of the value h_{crit}.

Results

For a parameter width $\sigma = 0.01$ we were always able to determine the bifurcation points. Figure 3 shows histograms of the random parameters and

the resulting values of h_{crit}. In 10^4 simulations, no critical value lower than ~ 2.68 was found. The correlation between the individual parameters and h_{crit} is shown in the (logarithmic) scatter plots in figure 4.

We find positive correlation values for α and k, respectively, while β is negatively correlated with h_{crit}. This shows that a lower damping (small α) and a stronger coupling (high β or low k) between genes makes the system more prone to oscillations.

For a parameter width $\sigma = 0.2$ we found a bifurcation point in 9779 out of 10000 simulations. Figures 5 and 6 show the the parameter histograms, the resulting histogram for h_{crit} and the correlations plots. The lowest value h_{crit} we found was ~ 2.09.

The qualitative behaviour of the cycle does not depend on the absolute scaling of time and concentration. This implies that h_{crit} can only depend the linear combination $\ln\alpha - \ln\beta + \ln k$ which is confirmed by our simulation results (see figures 4 and 6).

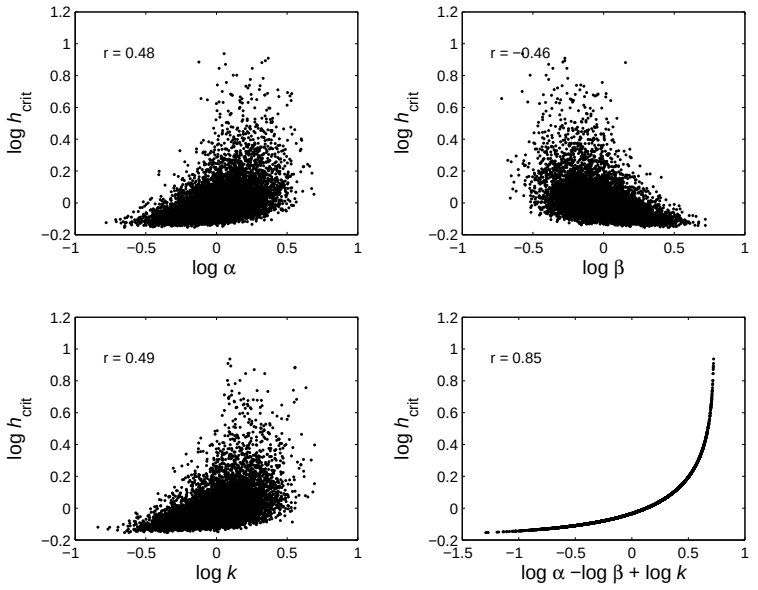

Figure 6: Correlation plots for the three parameters α, β and k versus h_{crit}. The bottom right plot still shows that h_{crit} is a function of $\ln\alpha - \ln\beta + \ln k$.

3 Discussion

Parameter estimation for complex dynamic models is a challenge in current systems biology. To study the potential dynamic behaviour of a given model with uncertain parameters, we use a Monte-Carlo sampling approach. We draw the parameters from a distribution and observe the distribution of the variables in the simulated system.

In the example model describing a small genetic network, the incidence of a Hopf bifurcation (a qualitative trait) and the distribution of the critical values of the Hill coefficient, at which the Hopf bifurcation occurs (a quantitative measure), have been determined. We may also ask a slightly different question: if all parameters (including the Hill coefficients) are drawn from distributions, what is the probability for the system to oscillate? Given our distribution of h_{crit}, this can be easily answered by sampling h and h_{crit} independently from their distributions and counting how often $h >$ crit .

The presented analysis can be considered as first step towards a thoroughly parameterized model. It gives hints, which types of qualitative behaviour can be expected at all and at which parameter combination. It enrolls which parameter values have a strong influence on the dynamics, which points points to parts of the model where exact measurements are necessary or where fluctuations are important or nonrelevant.

References

[1] H. Kitano. Systems biology: a brief overview. *Science*, 295(5560):1662–4, 2002. 1095-9203 Journal Article Review Review, Tutorial.

[2] E. Klipp, R. Herwig, A. Kowald, C. Wierling, and H. Lehrach. *Systems Biology in Practice. Concepts, Implementation and Application*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2005.

[3] P. Mendes and D. Kell. Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation. *Bioinformatics*, 14(10):869–83, 1998. 1367-4803 Journal Article.

[4] C. G. Moles, P. Mendes, and J. R. Banga. Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome Res*, 13(11):2467–74, 2003. 1088-9051 Journal Article.

[5] E. Klipp, W. Liebermeister, and C. Wierling. Inferring dynamic properties of biochemical reaction networks from structural knowledge. *Genome Inform Ser Workshop Genome Inform*, 15(1):125–37, 2004. 0919-9454 Journal Article.

[6] W. Liebermeister and E. Klipp. Biological networks with uncertain parameters. *submitted*.

[7] E. K. Ainscow and M. D. Brand. Errors associated with metabolic control analysis. application of monte-carlo simulation of experimental data. *J Theor Biol*, 194(2):223–33, 1998. 0022-5193 Journal Article.

[8] P. M. Kim and B. Tidor. Limitations of quantitative gene regulation models: a case study. *Genome Res*, 13(11):2391–5, 2003. 1088-9051 Journal Article Validation Studies.

[9] N. Blüthgen and H. Herzel. How robust are switches in intracellular signaling cascades? *J Theor Biol*, 225(3):293–300, 2003. 0022-5193 Journal Article.

[10] M.B. Elowitz and S. Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403:335–338, 2000.

[11] YA Kuznetsov A Dhooge, W Govaerts. MATCONT: A MATLAB package for numerical bifurcation analysis of ODEs.