



The Microsoft Research - University of Trento
Centre for Computational
and Systems Biology

Technical Report CoSBI 20/2007

A new model for kinetic parameter estimation in biochemical reactions

Paola Lecca

*The Microsoft Research - University of Trento
Centre for Computational and Systems Biology*

lecca@cosbi.eu

Guido Sanguinetti

Dept. of Computer Science, University of Sheffield, UK

guido@dcs.shef.ac.uk

Alida Palmisano

*The Microsoft Research - University of Trento
Centre for Computational and Systems Biology
DISI - University of Trento*

palmisano@cosbi.eu

Corrado Priami

*The Microsoft Research - University of Trento
Centre for Computational and Systems Biology
DISI - University of Trento*

priami@cosbi.eu

A new model for kinetic parameter estimation in biochemical reactions.

Paola Lecca, Guido Sanguinetti, Alida Palmisano, and Corrado Priami

Abstract

We present a novel method for estimating rate coefficients from noisy observations of concentration levels at discrete time points. This is traditionally done by computing the least-squares estimator. However, estimation of the error function generally requires solving the reaction rate equations, which can become computationally unfeasible. Here we present an alternative approach based on a probabilistic, generative model of the variations in reactant concentration. Our method returns the rate coefficients, the level of noise and an error range on the estimates of rate constants. Its probabilistic formulation is key to a principled handling of the noise inherent in biological data, and it allows a number of further extensions.

1 Introduction

The ability to infer kinetic parameters of biochemical reactions is emerging as a crucial problem in system biology. Estimates of reaction rates are often crucial for modelling and simulating the biochemical dynamics, yet their direct measurement is an outstanding experimental problem. The current need for new techniques for integrative and predictive modelling [1, 7, 6] is based on the realisation that traditional reductionist approaches have to be complemented by the reconstruction of a model of how systems function *as a whole*. The importance and the great interest that the scientific community attributes to this problem are evidenced by the increasing number of recent works concerning this research topic.

In this paper we present a novel method for estimating rate coefficients from noisy observations of concentration levels at discrete time points. Given a number of N reactant species, we observe time series concentrations for each of the species, gathered in N state vectors $\vec{X}_1, \dots, \vec{X}_N$. The relation between the instantaneous rate of reaction and the concentrations of the reactants at any moment is given by the law of mass action: *i.e.* the rate at which a substance takes part in a reaction is proportional to its concentration raised to a power which represents the number of molecules taking part in the reaction. Such formulation is made for simultaneous as well as isolated reactions, and for heterogeneous as well as homogeneous systems. The goal is the estimation of the constant of proportionality. This is traditionally done by computing the least-squares estimators (LSE), resulting from the minimisation of the sum of squared differences between the observed data and the model. However, estimation of the error function generally requires solving the reaction rate equations, which can easily become computationally unfeasible. Here we present an alternative approach based on a probabilistic, generative model of the *variations* in reactant concentration. Our method discretizes the law of mass action and provides a tool to predict the values of the variables \vec{X}_i at time t , conditioned on their values at the previous time point. The variations of the species concentration at different time points are conditionally independent by the Markovian nature of the discrete model of the law of mass action. Assuming the observation noise to be Gaussian with variance σ^2 , the probability of observing a variation D_i for the concentration $[X]_i$ of species i between time t_{k-1} and t_k is a Gaussian with variance depending on σ and mean the expectation value of the law mass action function under the noise distribution. The discretization of the law of mass action provides a model for the variations of the species concentration, rather than a model for the time-trajectory of the species concentrations. This makes the evaluation of the expectation value of law mass action function simpler and analytically tractable. The rate coefficients and the level of noise σ^2 are then obtained by maximising the likelihood function defined by the observed variations. Our method returns

the rate coefficients, the level of noise and an error range on the estimates of rate constants. Its probabilistic formulation is key to a principled handling of the noise inherent in biological data, and it allows a number of further extensions, such as a fully Bayesian treatment of the parameter inference and automated model selection strategies based on the comparison between marginal likelihoods of different models. Finally, the implementation of this method may be used as an interface tool, connecting the outcomes of the wet-lab activity for the concentration measurements and the softwares for the simulation of chemical kinetics. We tested our algorithm on some real case-study, whose rate coefficients were experimentally determined and well documented in literature for simple chemical reactions and for more complex biochemical interactions.

2 The model for inference

Consider N reactant species, S_1, S_2, \dots, S_N , with concentrations X_1, X_2, \dots, X_N , that evolve according to a system of rate equations

$$\frac{dX_i}{dt} = f_i(\mathbf{X}^{(i)}(t); \theta_i) \quad (1)$$

where θ_i , $i = 1, 2, \dots, N$, is the vector of the rate coefficients, which are present in the expression of the function f_i . We wish to estimate the set of parameters $\Theta = \cup \theta_i$ ($i = 1, 2, \dots, N$), whose element θ_i is the set of rate coefficients appearing in the rate equations of i -th species, therefore

$$\theta_1 = \{\theta_{11}, \theta_{12}, \dots, \theta_{1N_1}\}, \dots, \theta_N = \{\theta_{N1}, \theta_{N2}, \dots, \theta_{NN_N}\}$$

$\mathbf{X}^{(i)}$ is the vector of concentrations of chemicals that are present in the expression of the function f_i for the species i . According to the law of mass action, the functions f_i have the general form

$$\begin{aligned} f_i(\mathbf{X}^{(i)}(t); \theta_i) &= \\ &= \theta_{i1} \prod_{w \in S_1 \subseteq [1, N]} X_w^{\alpha_w} + \dots + \theta_{iN_i} \prod_{w \in S_{N_i} \subseteq [1, N]} X_w^{\alpha_w} = \sum_{h=1}^{N_i} \left(\theta_{ih} \prod_{w \in S_h} X_w^{\alpha_w} \right) \end{aligned} \quad (2)$$

where $\alpha_w \in \mathbf{R}$, and N_i is the number of parameter in the f_i rate equation. The rate equations in (2) form the so-called Generalized Mass Action law. We assume we have noisy observations $\hat{X}_i = X_i + \epsilon$ at times t_0, \dots, t_M , where $\epsilon \sim \mathcal{N}(0, \sigma^2)$ is a Gaussian noise term with mean zero and variance σ . With this choice we are assuming that the concentration measurements are not significantly affected by systematic errors, but by uncontrolled random

errors and that an error is equally likely to occur in either positive or negative direction with respect to the symmetry axis of the distribution.

We also assume a number M of concentration measurements for each considered species. Approximating the rate equation (1) as a finite difference equation between the observation times, gives

$$X_i(t_k) = X_i(t_{k-1}) + (t_k - t_{k-1})f_i(\mathbf{X}^{(i)}(t_{k-1}); \theta_i) \quad (3)$$

where $k = 1, \dots, M$. In Eq. (3) the rate equation is viewed as a model of increments/decrements of reactant concentrations; i.e., given a value of the variables at time t_{k-1} , the model can be used to predict the value at the next time point t_k . Increments/decrements between different time points are conditionally independent by the Markov nature of the model (3). Therefore, given the Gaussian model for the noise, it is possible to estimate the probability to observe the value $\hat{X}_i(t_k)$ given the model at time t_{k-1} , $X_i(t_{k-1})$, and the set of parameters θ_i , as

$$p(\hat{X}_i(t_k)|X_i(t_{k-1})) = \mathcal{N}(X_i(t_{k-1}) + (t_k - t_{k-1})f_i(X_i(t_{k-1}), \theta_i), \sigma^2) \quad (4)$$

We then also have that the true value of $X_i(t_k)$ is normally distributed around the observed value $\hat{X}_i(t_k)$, so that

$$\begin{aligned} p(X_i(t_{k-1})|\hat{X}_i(t_{k-1})) &= \mathcal{N}(\hat{X}_i(t_{k-1}), \sigma^2) = \\ &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(X_i(t_{k-1}) - \hat{X}_i(t_{k-1}))^2}{2\sigma^2}\right] \end{aligned} \quad (5)$$

Therefore, the probability to observe a variation $D_i(t_k) = X_i(t_k) - X_i(t_{k-1})$ for the concentration of the i -th species between the time t_{k-1} and t_k , given the parameter vector θ_i is

$$p(D_i(t_k)|\theta_i, \sigma) = \mathcal{N}(E[f_i(\mathbf{X}^{(i)}(t_{k-1}), \theta_i)], 2\sigma^2) \quad (6)$$

and

$$E[f_i(\mathbf{X}^{(i)}(t_{k-1}), \theta_i)] = \int_{\Omega_{\mathbf{X}^{(i)}}} f_i(\mathbf{X}^{(i)}(t_{k-1}), \theta_i) \prod_{i=1}^{K_i} \left[p_i(X_i(t_{k-1})|\hat{X}_i(t_{k-1})) \right] d\mathbf{X}^{(i)} \quad (7)$$

where $\Omega_{\mathbf{X}^{(i)}}$ is the sample space of $\mathbf{X}^{(i)}$, and K_i is the number of chemical species in the expression for f_i . While the increments/decrements are conditionally independent given the starting point $X_i(t_k)$, the random variables $D_i(t_k)$ are not independent of each other. Intuitively, if $X_i(t_k)$ happens to be below its expected value because of random fluctuations, then the following increment $D_i(t_{k+1})$ can be expected to be bigger as a result, while the

previous one $D_i(t_k)$ will be smaller. A simple calculation allows us to obtain the covariance matrix of the vector of increments for the i -th species. This is a banded matrix $\mathbf{C}_i \equiv \mathbf{C} = \text{Cov}(\mathbf{D}_i)$ with diagonal elements given by

$$E\left[D_i^2(t_k) - E[D_i^2(t_k)]\right] = 2\sigma^2$$

and a non-zero band above and below the diagonal given by

$$E\left[(D_i(t_k) - E[D_i(t_k)])(D_i(t_{k-1}) - E[D_i(t_{k-1})])\right] = -\sigma^2$$

with all other entries zero. The likelihood for the observed increments/decrements therefore will be

$$p(\mathbf{D}|\Theta) = \prod_{i=1}^N \mathcal{N}(\mathbf{D}_i|\mathbf{m}_i(\Theta), \mathbf{C}) = \left(\frac{1}{\sqrt{2\pi \det(\mathbf{C})}}\right)^N e^{-\frac{1}{2}(\mathbf{D}_i - \mathbf{m}_i)^T \mathbf{C}^{-1}(\mathbf{D}_i - \mathbf{m}_i)} \quad (8)$$

where $\mathbf{D} = \{\mathbf{D}_1, \dots, \mathbf{D}_N\}$, $\mathbf{D}_i = D_i(t_1), D_i(t_2), \dots, D_i(t_M)$ ($i = 1, 2, \dots, N$), and $\mathbf{m}_i(t_{k-1}) \equiv E\left[f_i(\mathbf{X}(t_{k-1}), \theta_i)\right]$.

The Eq. (8) can be optimized w. r. t. the parameters $\Theta = (\theta_1, \theta_2, \dots, \theta_N)$ of the model to yield estimates of the parameters themselves and of the noise level. The chief numerical problem of this approach is the computation of the expectations of the rate functions given by equation (7). Non-integer values of the coefficients α can make estimating the integral analytically difficult. We propose an approximate method in which the Gaussian noise is replaced by an approximate uniform (white) noise, with the amplitude of the uniform noise being obtained as a sample from the Gaussian cumulative distribution function. At the first order, for small σ , we can approximate the Gaussian with zero mean and variance σ with an uniform distribution defined on the interval $[-\frac{\sqrt{2\pi}\sigma}{4}, \frac{\sqrt{2\pi}\sigma}{4}]$, so that

$$\prod_{i=1}^{K_i} p_i = \prod_{i=1}^{K_i} \chi_i \quad (9)$$

where

$$\chi_i(X_i) = \begin{cases} \frac{2}{\sqrt{2\pi}\sigma} & \text{if } -\frac{\sqrt{2\pi}\sigma}{4} \leq X_i \leq \frac{\sqrt{2\pi}\sigma}{4} \\ 0 & \text{otherwise.} \end{cases}$$

This approximation makes the calculation of the expectation value of the rate equation (Eq. (7)) simpler and reduces the computational time of the procedure. Moreover, experiments not illustrated in this paper demonstrate that it does not influence the accuracy of the parameter estimates until σ is less than 30% of the concentration measurement.

Substituting Eq. (9) in Eq. (7) gives

$$E[f_i(\mathbf{X}^{(i)}(t_{k-1}), \theta)] = \left(\frac{2}{\sqrt{2\pi\sigma}} \right)^{K_i} \int_{\hat{X} - \frac{\sqrt{2\pi\sigma}}{4}}^{\hat{X} + \frac{\sqrt{2\pi\sigma}}{4}} f_i(\mathbf{X}^{(i)}(t_{k-1}), \theta_i) d\mathbf{X}^{(i)} \quad (10)$$

Now, substituting Eq. (2) in Eq. (10) leads to

$$\begin{aligned} E[f_i(\mathbf{X}^{(i)}(t_{k-1}), \theta_i)] &= \\ &= \left(\frac{2}{\sqrt{2\pi\sigma}} \right)^{K_i} \left\{ \sum_{h=1}^{N_i} \theta_{ih} \left[\left(\frac{\sqrt{2\pi\sigma}}{2} \right)^{\#(S-S_h)} \times \right. \right. \\ &\times \left. \left. \prod_{w \in S_h} \frac{1}{\alpha_w + 1} \left(\left(\hat{X}_w + \frac{\sqrt{2\pi\sigma}}{4} \right)^{\alpha_w + 1} - \left(\hat{X}_w - \frac{\sqrt{2\pi\sigma}}{4} \right)^{\alpha_w + 1} \right) \right] \right\} \quad (11) \end{aligned}$$

where S is the set containing the indexes referring to all the K_i species appearing in f_i , and $\alpha_w \neq -1$. In case some orders are equal to -1 Eq. (11) takes the following form

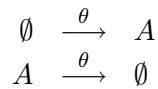
$$\begin{aligned} E[f_i(\mathbf{X}^{(i)}(t_{k-1}), \theta_i)] &= \left(\frac{2}{\sqrt{2\pi\sigma}} \right)^{K_i} \sum_{h=1}^{N_i} \theta_{ih} \left\{ \left(\frac{\sqrt{2\pi\sigma}}{2} \right)^{\#(S-S_h)} \times \right. \\ &\times \left[\prod_{w \in S'_h} \frac{1}{\alpha_w + 1} \left(\left(\hat{X}_w + \frac{\sqrt{2\pi\sigma}}{4} \right)^{\alpha_w + 1} - \left(\hat{X}_w - \frac{\sqrt{2\pi\sigma}}{4} \right)^{\alpha_w + 1} \right) \right] \times \\ &\times \left[\prod_{w \in S''_h} \ln \frac{\hat{X}_w + \frac{\sqrt{2\pi\sigma}}{4}}{\hat{X}_w - \frac{\sqrt{2\pi\sigma}}{4}} \right] \left. \right\} \quad (12) \end{aligned}$$

where S'_h is the set of indexes $\{h'_1, h'_2, \dots, h'_s\}$ such that $\alpha_{h'} \neq -1 \forall h' \in S'_h$, and S''_h is the set of indexes $\{h''_1, h''_2, \dots, h''_s\}$ such that $\alpha_{h''} = -1 \forall h'' \in S''_h$.

If in the Eq. (8), \mathbf{m}_i is substituted with the expression (11) or (12), Eq. (8) becomes more tractable and can be optimized w. r. t. the parameters $\Theta = (\theta_1, \theta_2, \dots, \theta_N)$ and σ . The values of the model's parameters for which $p(\mathbf{D}|\Theta)$ has a maximum are the most likely values giving the observed kinetics.

3 The estimation errors

For a zero-th order reaction

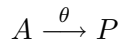


$$\Delta\theta = \left\{ \left(\frac{\partial\theta}{\partial A_0} \epsilon_{A_0} \right)^2 + \left(\frac{\partial\theta}{\partial A_f} \epsilon_{A_f} \right)^2 \right\}^{\frac{1}{2}} \quad (13)$$

$$\frac{\partial k}{\partial A_f} = \frac{(A_f - A_0)}{t_f - t_0} \quad (14)$$

$$\frac{\partial k}{\partial A_0} = \frac{(A_0 - A_f)}{t_f - t_0} \quad (15)$$

For a first order reaction



if A_0 (A_f) is the initial (final) concentration of the molecules of the species A , the error that the algorithm makes in estimating θ is given by the following

$$\Delta\theta = \left\{ \left(\frac{\partial\theta}{\partial A_0} \epsilon_{A_0} \right)^2 + \left(\frac{\partial\theta}{\partial A_f} \epsilon_{A_f} \right)^2 \right\}^{\frac{1}{2}} \quad (16)$$

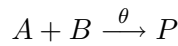
Since

$$\theta = \frac{1}{t_f - t_0} (\ln A_0 - \ln A_f) \quad (17)$$

Eq. 16 becomes

$$\Delta\theta = \left\{ \left(\frac{\partial\theta}{\partial A_0} \epsilon_{A_0} \right)^2 + \left(\frac{\partial\theta}{\partial A_f} \epsilon_{A_f} \right)^2 \right\}^{\frac{1}{2}} \quad (18)$$

For a second order reaction



if A_0 (A_f) and B_0 (B_f) are the initial (final) concentrations of the molecules of species A and B , respectively, the error that the algorithm makes in estimating θ is given by the following

$$\Delta\theta = \left\{ \left(\frac{\partial\theta}{\partial A_0} \epsilon_{A_0} \right)^2 + \left(\frac{\partial\theta}{\partial B_0} \epsilon_{B_0} \right)^2 + \left(\frac{\partial\theta}{\partial A_f} \epsilon_{A_f} \right)^2 + \left(\frac{\partial\theta}{\partial B_f} \epsilon_{B_f} \right)^2 \right\}^{\frac{1}{2}} \quad (19)$$

where

$$\frac{\partial\theta}{\partial A_0} = \frac{1}{t_f - t_0} \left\{ \frac{1}{(A_0 - B_0)^2} \ln \frac{A_f B_0}{A_0 B_f} - \frac{1}{A_0 - B_0} \frac{B_f}{A_0} \right\} \quad (20)$$

$$\frac{\partial\theta}{\partial B_0} = \frac{1}{t_f - t_0} \left\{ -\frac{1}{(A_0 - B_0)^2} \ln \frac{A_f B_0}{A_0 B_f} + \frac{1}{A_0 - B_0} \frac{A_f}{B_0} \right\} \quad (21)$$

$$\frac{\partial \theta}{\partial A_f} = \frac{1}{t_f - t_0} \frac{1}{A_0 - B_0} \frac{B_0}{A_f} \quad (22)$$

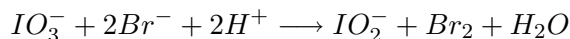
$$\frac{\partial \theta}{\partial B_f} = \frac{1}{t_f - t_0} \frac{1}{B_0 - A_0} \frac{A_0}{B_f} \quad (23)$$

4 Case studies

All the mathematical operations described above has been automated through the implementation of an algorithm in C++. We tested our algorithm on some real case-study, whose rate coefficients were experimentally determined and well documented in literature for simple chemical reactions and for more complex biochemical interactions. We show the ability of our algorithm of obtaining reasonable estimates for the rate coefficients in the following cases (see Table 1):

R_1 Bromine molecule formation $2Br \longrightarrow Br_2$. The data referred to this reaction are the bromine concentration time-points after a flash photolysis of a mixture of bromine and SF_6 with $[Br_2]/[SF_6] = 3.2 \times 10^{-2}$ (experiment of Graff and Lang [11]).

R_2 . In presence of an acid solution of phenol, the ion IO_3^- is reduced to IO_2^- by Br^- accordingly to the reaction



With $[IO_3^-] = 5 \times 10^{-3} M$ and $[Br^-] = 1 \times 10^{-2} M$, Sharma and Gupta obtained the following data for a solution at 35 C with $[C_6H_5OH] = 2 \times 10^{-2} M$ [11].

R_3 . the first-order decomposition in azomethane $CH_3N_2CH_3 \longrightarrow CH_3CH_3 + N_2$ ([13]);

R_4 . the alkaline hydrolysis of ethyl nitrobenzoate ([13]);

R_5 . the glutathione S-transferase-catalysed dehalogenation of haloaromatic compounds with aromatic substrates ([15, 10]);

R_6 . degradation of I κ Ba-UB in breast cancer cells (refer to [3] for a comparison);

R_7 . Nuclear factor (NF) κ B translocation from cytoplasm into the nucleus in breast cancer cells (r. t. [3] for a comparison);

R_8 . the glucose consumption rate in astrocytes [9].

Reaction	k (experimental)	k (estimate)	σ (estimate)
R_1	$(2.75 \pm 1.00) \times 10^8 M^{-1} s^{-1}$	$(1.93 \pm 1.28) \times 10^8 M^{-1} s^{-1}$	3.2×10^{-6}
R_2	$(13.8 \pm 1.00) M^{-2} s^{-2}$	$(15.92 \pm 6.98) M^{-2} s^{-1}$	3.5×10^{-1}
R_3	$(3.60 \pm 2.16) \times 10^{-4} s^{-1}$	$(0.80 \pm 0.45) \times 10^{-4} s^{-1}$	5.36×10^{-3}
R_4	$(8.1 \pm 2.31) \times 10^{-2} s^{-1}$	$(5.9 \pm 1.5) \times 10^{-2} s^{-1}$	3.3×10^{-3}
R_5	$(7.4 \pm 0.5) \times 10^{-4} s^{-1}$	$(2.60 \pm 1.17) \times 10^{-4} s^{-1}$	4.1×10^{-3}
R_6	$(5.56 \pm 0.11) \times 10^{-4} s^{-1}$	$(3.1 \pm 2.2) \times 10^{-4} s^{-1}$	5×10^{-3}
R_7	$(3.85 \pm 2.67) \times 10^{-4} s^{-1}$	$(1.3 \pm 0.4) \times 10^{-4} s^{-1}$	8.5×10^{-3}
R_8	$(1.33 \pm 0.04) \times 10^{-2} s^{-1}$	$(1.24 \pm 0.78) \times 10^{-2} s^{-1}$	6×10^{-4}

Table 1: The estimates of the rate coefficients compared with the experimental values.

5 Conclusion

The ability of this algorithm to estimate the expected value of the kinetic rate constant significantly depends on the experimental error on the data. If the time-course measurements of the reactant concentration are affected by a high level of noise, this technique estimate a rate constant very close to zero, indicating that there is no dynamic in the systems. Therefore the observed variations in the concentration values are attributed exclusively to the noise. For an accurate estimate of the rate constants the procedure requires high-quality and high-resolution data, that at the time of writing are relatively difficult to obtain, In the meantime experimental biology evolves to make such data more readily available, the development of inference techniques allow identification of experimental design issues that need to be addressed to get more accurate data for more reliable estimate of rate constants.

References

- [1] J. M. Bower, H. Bolouri (Eds.), *Computational Modeling of Genetic and Biochemical Networks*, Cambridge, Massachusetts, MIT Press, 2001.
- [2] E. . Crampin, P. E. McSharry, and S. Schnell, *Extracting Biochemical Reaction Kinetics from Time Series Data*, M. GH. Negoita et al. (Eds): KES 2004, LNAI 3214, p. 329-336, Springer-Verlag Berlin Heidelberg 2004.
- [3] J. W. Fuseler, D. M. Merril, J. A. Rogers, M. B. Grisham and R. E. Wolf, *Analysis and quatitation of NF-kB nuclear translocation in tumor necrosis factor alpha (TNF- α) activated vascular endothelial cells*, Microsc. Microanal. **12**, pp. 269-276, 2006.

- [4] A. Golightly and D. J. Wilkinson, *Bayesian sequential inference for stochastic kinetic biochemical network inference*, Journal of Computational Biology vol. 13, no. 3 pp. 838-851, 2006.
- [5] A. Golightly and D. J. Wilkinson, *Bayesian inference for stochastic kinetic models using a diffusion approximation*, Biometrics, vol. 61, no. 3, pp. 781-788, 2005.
- [6] T. B. L. Kirkwood, R. J. Boys, C. S. Gillespie, C. J. Proctor, D. P. Shanley, and D. J. Wilkinson, *Toward an e-biology of ageing: integrating theory and data*, Nature Reviews Molecular Cell Biology 4, pp. 243-249, 2003.
- [7] H. Kitano (Ed.), *Foundations of System Biology*, Cambridge, Massachusetts, MITOress, 2001.
- [8] N. Lalam and C. A. J. Klassen, *Pseudo maximum likelihood estimation for differential equations*,
- [9] P. Lecca and M. Lecca, *Molecular mechanism of glutamate-triggered brain glucose metabolism: a parametric model from FDG PET-scans*, In Procs of BVAI 2007, To appear in LNCS, 2007.
- [10] I. Mancini, G. Guella, G. Chiasera and F. Pietra, *Glutathione S-transferase catalysed dehalogenation of haloaromatic compounds which lack nitro groups near the reaction centre*, Tetrahedron Letters, **39**, pp. 1611-1614, 1998.
- [11] C. R. Metz, *Chimica fisica*, McGraw-Hill, 1994.
- [12] I. Nachman, A. Regev and N. Friedman, *Inferring quantitative models of regulatory networks from expression data*, Bioinformatics, vol. 20, no. 1, pp. 248-256, 2004.
- [13] V. Nummert and M. Piirsalu, *Kinetic study of hydrolysis of benzoates*, J. Chem. Soc., Perkin Trans. **2**, pp. 583-594, 2000.
- [14] S. Reinker, R. M. Altman and J. Timmer, *Parameter estimation in stochastic biochemical reactions*, IEE Proc. Syst. Biol, vol. 153, no. 4, July 2006.
- [15] B. Rossi, I. Mancini and G. Guella, *Kinetic studies of glutathione S-transferase-catalysed process through on-line liquid chromatography-electrospray mass spectrometry*, Philosophical Magazine, **84**(13-16), pp. 1373-1382, 2004.
- [16] J. Yu, V. A. Smith, P. P. Wang, a. J. Hartemink and E. D. Jarvis, *Advances to Bayesian network inference for generating causal networks*

from observational biological data, Bioinformatics vol 20, no. 18, pp. 3594 - 3603, 2004.

- [17] R. J. Boys, D. J. Wilkinson and T. B. Kirkwood, Bayesian inference for a discretely observed stochastic kinetic model, currently in submission (available at <http://www.mas.ncl.ac.uk/~ndjw1/docs/BWK04.pdf>).