31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Parameter estimation in rational models of molecular biological systems

Fang-Xiang Wu, *Member, IEEE*, and Lei Mu

Abstract—**Based on statistical thermodynamics or Michaelis -Menten kinetics, molecular biological systems can be modeled by a system of nonlinear differential equations. The nonlinearity in the model stems from rational reaction rates whose numerator and denominator are linear in parameters. It is a nonlinear problem to estimate the parameters in such rational models of molecular biological systems. In principle, any nonlinear optimization methods such as Newton-Gauss method and its variants can be used to estimate parameters in the rational models. However, these methods may converge to a local minimum and be sensitive to the initial values. In this study, we propose a new method to estimate the parameters in the rational models of molecular biological systems. In the proposed method, the cost function in all parameters is first reduced to a cost function only in the parameters in the denominator by a separable theorem. Then the parameters in the denominator are estimated by minimizing this cost function using our proposed new iteration method. Finally, the parameters in the numerator are estimated by a well defined linear least squares formula. A simple gene regulatory system is used as an example to illustrate the performance of the proposed method. Simulation results show that the proposed method performs better than the general nonlinear optimization methods in terms of the running time, robustness (insensitivity) to the initial values, and the accuracy of estimates.**

Keywords: Parameter estimation, nonlinear optimization, molecular biological systems, rational model.

I. INTRODUCTION

As modern molecular biology moves towards the mechanism of biological systems, the modeling and simulation become very important tools. The dynamics of molecular biological systems are commonly modeled in terms of systems of ordinary differential equations that involve parameters corresponding to kinetic constants. Most, if not all, models for molecular biological systems are nonlinear in both parameters and system state variables. Estimation of parameters in these models is a nonlinear estimation problem. In particular, if they are derived on basis of statistical thermodynamics [1, 2] or Michaelis-Menten kinetics [2-4], nonlinear functions in the resultant models are rational functions whose numerator and denominator are linear in parameters. Parameters in such rational molecular biological systems are typically reaction constants of interest. Estimation of these parameters is crucial to construct the whole molecular biological systems [5]. In general, all nonlinear optimization programs can be used to estimate parameters in the rational models of biological systems, for example, Gauss-Newton iteration method and its variants such as Box-Kanemasu interpolation method, Levenberg damped least squares methods, and Marquardt's method [6]. However, these iteration methods are sensitive to initial values. Another main shortcoming is that these methods may converge to the local minimum of the least squares cost function, and thus cannot find the real values of the parameters.

In general, a rational model contains a rational (linear fractional) function in the following format:

$$\eta(\mathbf{X},\boldsymbol{\beta}) = \frac{N_0(\mathbf{X}) + \sum_{i=1}^{p_N} N_i(\mathbf{X})\beta_{Ni}}{D_0(\mathbf{X}) + \sum_{j=1}^{p_D} D_j(\mathbf{X})\beta_{Dj}} \quad (1)$$

where the vector $\mathbf{X}$ consists of the independent observation variables, the $p-$ dimensional vector $\boldsymbol{\beta}$ consists of all parameters in the rational function, which can naturally be divided into two groups: those in the numerator, β_{Ni} ($i = 1,\cdots,p_N$), and those in the denominator β_{Dj} ($j = 1,\cdots,p_D$), where we have that $p_D + p_N = p$. The coefficient functions $N_i(\mathbf{X})$ ($i = 0,1,\cdots,p_N$) and $D_i(\boldsymbol{X})$ ($j = 0,1,\cdots,p_D$) are the known functions of the independent variables and do not contain any unknown parameters. Either $N_0(\mathbf{X})$ or $D_0(\boldsymbol{X})$ must be nonzero, and otherwise from identifiability analysis [6] the parameters in model (1) cannot be uniquely identified.

In this study, we take use of the special structure of rational model (1): numerator and denominator are linear in parameters, and propose a new method to estimate the parameters in the rational models of molecular biological systems. In the proposed method, the cost function in all parameters is first reduced to a cost function only in the parameters in the denominator by a separable theorem [7]. Then the parameters in the denominator are estimated by minimizing this cost function by our developed new iteration method. Finally, the parameters in the numerator are estimated by a well defined least squares formula. Briefly, the reminder of paper is organized as follows. Section II introduces the proposed method. Section III provides an illustrative example to show the performance of the proposed method, comparing with the nonlinear optimization algorithm. Finally we give conclusions and future work in Section IV.

Manuscript received April 7 2009. This research is supported by Natural Sciences and Engineering Research Council of Canada (NSERC).

F. X. Wu is with Department of Mechanical Engineering and associated with Division of Biomedical Engineering in the University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada. (phone: 1(306)966-5280; fax: 1(306)966-5427; email: faw341@mail.usask.ca).

L.Mu is with Department of Mechanical Engineering in the University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada

II. ALGORITHM DESCRIPTION

Let us introduce some notation first. R^n denotes the set of all n-dimensional column (or row) vectors; $R^{n\times m}$ denotes the set of all $n\times m$ matrices. The superscript T denotes the matrix transpose. The Euclidian norm of an n-dimensional vector $\mathbf{a}=[a_1 \quad a_2 \quad \cdots \quad a_n]\in R^n$ is defined by $\|a\|=\sqrt{\sum_{i=1}^{n}a_i^2}$. $diag\ [a_1 \quad a_2 \quad \cdots \quad a_n]$ (or $diag\ [a_1 \quad a_2 \quad \cdots \quad a_n]^T$) is a diagonal matrix whose diagonal elements are a_i $(i=1,2,\cdots,n)$ sorted in this order.

Suppose that at a series of time points we obtain a sequence of measurements (observations) of dependent variable: y_t $(t=1,2,\cdots,n)$, which can be represented by a rational function of independent variables and parameters. In practice, any measurements can be contaminated by some random noises. Assume that measurement errors are additive. Thus we have the relationship

$$y_t=\eta(\boldsymbol{X}_t,\boldsymbol{\beta})+\varepsilon_t=\eta_t(\boldsymbol{\beta})+\varepsilon_t,\quad t=1,2,\cdots,n \tag{2}$$

where ε_t $(t=1,2,\cdots,n)$ stand for the measurement errors at time point t, and $\mathbf{X}_t$ $(t=1,2,\cdots,n)$ stand for the measured or known values of independent variables at time point t. Assume that independent variables $\mathbf{X}$ and parameters $\boldsymbol{\beta}$ are non-random variables. Further, without the loss of the generality, assume that the measurement errors ε_t $(t=1,2,\cdots,n)$ have the mean of zeros.

Define the two parameter vectors $\boldsymbol{\beta}_N$ and $\boldsymbol{\beta}_D$ for parameters in the numerator and in the denominator, respectively,

$$\boldsymbol{\beta}_N=[\beta_{N1},\quad \beta_{N2},\quad \cdots\quad ,\beta_{Np_N}]^T\in R^{p_N}$$
$$\boldsymbol{\beta}_D=[\beta_{D1},\quad \beta_{D2},\quad \cdots\quad ,\beta_{Dp_D}]^T\in R^{p_D}$$

Define the following vectors and matrices

$$\boldsymbol{Y}=[y(1),\quad y(2),\quad \cdots,\quad y(n)]^T\in R^n$$
$$\boldsymbol{\eta}(\boldsymbol{\beta})=[\eta_1(\boldsymbol{\beta}),\quad \eta_2(\boldsymbol{\beta}),\quad \cdots,\quad \eta_n(\boldsymbol{\beta})]^T\in R^n$$
$$\varphi_N(\boldsymbol{X}_t)=[N_1(\boldsymbol{X}_t),\quad N_2(\boldsymbol{X}_t),\quad \cdots\quad ,N_{p_N}(\boldsymbol{X}_t)]\in R^{p_N}$$
$$\varphi_D(\boldsymbol{X}_t)=[D_1(\boldsymbol{X}_t),\quad D_2(\boldsymbol{X}_t),\quad \cdots\quad ,D_{p_D}(\boldsymbol{X}_t)]\in R^{p_D}$$

$$\Phi_{N_i}=\begin{bmatrix}N_i(\boldsymbol{X}_1)\\N_i(\boldsymbol{X}_2)\\\vdots\\N_i(\boldsymbol{X}_n)\end{bmatrix}\in R^n,\ \text{for } i=0,\ 1,\ \ldots,\ p_N$$

$$\Phi_N=[\Phi_{N_1}\quad \cdots\quad \Phi_{N_{P_N}}]\in R^{n\times p_N}$$

$$\Phi_{D_i}=\begin{bmatrix}D_i(\boldsymbol{X}_1)\\D_i(\boldsymbol{X}_2)\\\vdots\\D_i(\boldsymbol{X}_n)\end{bmatrix}\in R^n,\ \text{for } i=0,\ 1,\ \ldots,\ p_D$$

$$\Phi_D=[\Phi_{D_1}\quad \cdots\quad \Phi_{D_{P_D}}]\in R^{n\times p_D}$$
$$\Psi(\boldsymbol{\beta}_D)=diag\left[\Phi_{D_0}+\Phi_D\boldsymbol{\beta}_D\right]\in R^{n\times n}$$

From the above definitions, we have

$$y_t=\eta(\boldsymbol{X}_t,\boldsymbol{\beta})=\frac{N_0(\boldsymbol{X}_t)+\varphi_N(\boldsymbol{X}_t)\boldsymbol{\beta}_N}{D_0(\boldsymbol{X}_t)+\varphi_D(\boldsymbol{X}_t)\boldsymbol{\beta}_D}+\varepsilon_t \tag{3}$$

Form a sum of squared errors (the cost function)

$$J(\boldsymbol{\beta})=J(\boldsymbol{\beta}_N,\boldsymbol{\beta}_D)=[\boldsymbol{Y}-\boldsymbol{\eta}(\beta)]^T[\boldsymbol{Y}-\boldsymbol{\eta}(\beta)] \tag{4}$$

Minimizing $J(\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}$ can give the least squares estimation of parameters $\boldsymbol{\beta}_N$ and $\boldsymbol{\beta}_D$.

As parameters $\boldsymbol{\beta}$ are nonlinear in the rational function, Gauss-Newton iteration method and its variants [6] can typically be applied to estimation of these parameters by minimizing the cost function (4). However, it is well known that Gauss-Newton method may fall into a local minimum and thus cannot find the estimates of the parameters. We have observed that the parameters $\boldsymbol{\beta}_N$ are linear in the model $\eta(\mathbf{X},\boldsymbol{\beta})$. Let $F(\boldsymbol{\beta}_D)=[\Psi(\boldsymbol{\beta}_D)]^{-1}\Phi_N$ and $G(\boldsymbol{\beta}_D)=[\Psi(\boldsymbol{\beta}_D)]^{-1}\Phi_{N_0}$, the cost function (4) becomes

$$\begin{aligned}&J(\boldsymbol{\beta})=J(\boldsymbol{\beta}_N,\boldsymbol{\beta}_D)\\&=[\boldsymbol{Y}-G(\boldsymbol{\beta}_D)-F(\boldsymbol{\beta}_D)\boldsymbol{\beta}_N]^T[\boldsymbol{Y}-G(\boldsymbol{\beta}_D)-F(\boldsymbol{\beta}_D)\boldsymbol{\beta}_N]\end{aligned} \tag{5}$$

From Theorem 2.1 in reference [7], if there exists an (n-k) × n matrix $\boldsymbol{X}_1(\boldsymbol{\beta}_D)$ having the rank of n-k and satisfying

$$\boldsymbol{X}_1(\boldsymbol{\beta}_D)F(\boldsymbol{\beta}_D)=0 \tag{6}$$

Then

$$\min_{\boldsymbol{\beta}_N,\boldsymbol{\beta}_D}J(\boldsymbol{\beta}_N,\boldsymbol{\beta}_D)=\min_{\boldsymbol{\beta}_D}K(\boldsymbol{\beta}_D) \tag{7}$$

where the cost function $K(\boldsymbol{\beta}_D)$

$$\begin{aligned}&K(\boldsymbol{\beta}_D)=[\boldsymbol{X}_1(\boldsymbol{\beta}_D)(\boldsymbol{Y}-G(\boldsymbol{\beta}_D))]^T\times\\&[\boldsymbol{X}_1(\boldsymbol{\beta}_D)\boldsymbol{X}_1(\boldsymbol{\beta}_D)^T]^{-1}\boldsymbol{X}_1(\boldsymbol{\beta}_D)(\boldsymbol{Y}-G(\boldsymbol{\beta}_D))\end{aligned} \tag{8}$$

is independent of parameters in the numerator $\boldsymbol{\beta}_N$. Let $\hat{\boldsymbol{\beta}}_D$ be the optimizer of (7) and (8), then $\hat{\boldsymbol{\beta}}_D$ and $\hat{\boldsymbol{\beta}}_N$ minimize the cost function (5), where $\hat{\boldsymbol{\beta}}_N$ is calculated as follows:

$$\hat{\boldsymbol{\beta}}_N=\left[F^T(\hat{\boldsymbol{\beta}}_D)F(\hat{\boldsymbol{\beta}}_D)\right]^{-1}F^T(\hat{\boldsymbol{\beta}}_D)[\boldsymbol{Y}-G(\hat{\boldsymbol{\beta}}_D)] \tag{9}$$

Let $\Phi_N^{\perp}$ be an (n-k) × n matrix with the rank of n-k orthogonal to matrix Φ_N, then we can construct

$$\boldsymbol{X}_1(\boldsymbol{\beta}_D)=\Phi_N^{\perp}\Psi(\boldsymbol{\beta}_D) \tag{10}$$

From the definition of $\Psi(\boldsymbol{\beta}_D)$ we have

$$\begin{aligned}&\Psi(\boldsymbol{\beta}_D)(\boldsymbol{Y}-G(\boldsymbol{\beta}_D))\\&=diag[Y]\Phi_{D_0}-\Phi_{N_0}+diag[Y]\Phi_D\boldsymbol{\beta}_D=\boldsymbol{b}+\boldsymbol{A}\boldsymbol{\beta}_D\end{aligned} \tag{11}$$

where constant vector $\boldsymbol{b} = diag[\boldsymbol{Y}]\Phi_{D_0} - \Phi_{N_0} \in R^n$, and constant matrix $\boldsymbol{A} = diag[\boldsymbol{Y}]\Phi_D \in R^{n \times p_D}$ are independent of estimated parameters

Substituting (10) and (11) into (8) yields to

$$K(\boldsymbol{\beta}_D) = [\boldsymbol{b} + \boldsymbol{A}\boldsymbol{\beta}_D]^T \Phi_N^{\perp T} \boldsymbol{M}(\boldsymbol{\beta}_D)^{-1} \Phi_N^{\perp} [\boldsymbol{b} + \boldsymbol{A}\boldsymbol{\beta}_D] \quad (12)$$

where matrix $\boldsymbol{M}(\boldsymbol{\beta}_D) = \Phi_N^{\perp} \Psi(\boldsymbol{\beta}_D) \Psi(\boldsymbol{\beta}_D)^T \Phi_N^{\perp T}$. The necessary condition for minimizing $K(\boldsymbol{\beta}_D)$ with respect to $\boldsymbol{\beta}_D$ is that $\partial K(\boldsymbol{\beta}_D)/\partial \boldsymbol{\beta}_D = 0$, which gives

$$\begin{aligned} &\boldsymbol{A}^T \Phi_N^{\perp T} \boldsymbol{M}(\boldsymbol{\beta}_D)^{-1} \Phi_N^{\perp} \boldsymbol{b} + \boldsymbol{A}^T \Phi_N^{\perp T} \boldsymbol{M}(\boldsymbol{\beta}_D)^{-1} \Phi_N^{\perp} \boldsymbol{A}\boldsymbol{\beta}_D \\ &- \boldsymbol{U}(\boldsymbol{\beta}_D)^T \boldsymbol{u}_0(\boldsymbol{\beta}_D) - \boldsymbol{U}(\boldsymbol{\beta}_D)^T \boldsymbol{U}(\boldsymbol{\beta}_D) \boldsymbol{\beta}_D = 0 \end{aligned} \quad (13)$$

where matrix $\boldsymbol{U} = [\boldsymbol{u}_1 \quad \cdots \quad \boldsymbol{u}_{p_D}]$ and its i-th column vector is defined as

$$\begin{aligned} &\boldsymbol{u}_i = diag[\Phi_{D_i}] \Phi_N^{\perp T} \boldsymbol{M}(\boldsymbol{\beta}_D)^{-1} \Phi_N^{\perp} [\boldsymbol{b} + \boldsymbol{A}\boldsymbol{\beta}_D] \in R^n \\ &\text{for i} = 0, 1, \ldots, \text{p}_\text{D} \end{aligned} \quad (14)$$

From (13), we propose an iteration formula to solve optimization problem (7) as follows:

$$\begin{aligned} \boldsymbol{\beta}_D^{k+1} = &[\boldsymbol{A}^T \Phi_N^{\perp T} \boldsymbol{M}(\boldsymbol{\beta}_D^k)^{-1} \Phi_N^{\perp} \boldsymbol{A}]^{-1} [\boldsymbol{U}(\boldsymbol{\beta}_D^k)^T \boldsymbol{U}(\boldsymbol{\beta}_D^k) \boldsymbol{\beta}_D^k \\ &- \boldsymbol{A}^T \Phi_N^{\perp T} \boldsymbol{M}(\boldsymbol{\beta}_D^k)^{-1} \Phi_N^{\perp} \boldsymbol{b} - \boldsymbol{U}(\boldsymbol{\beta}_D^k)^T \boldsymbol{u}_0(\boldsymbol{\beta}_D^k)] \end{aligned} \quad (15)$$

III. ILLUSTRATIVE EXAMPLE

The expression of a gene is regulated by regulatory proteins and/or RNA polymerase (RNAP) which are binding to gene's regulatory binding site [1]. The regulatory binding site of a gene is a short piece of DNA sequence closed to it. One gene can have a number of binding sites. The binding sites for regulatory proteins are called operators while those for RNAP are called promoters. A gene regulatory network is a collection of genes that regulate one another's expression rates through their encoded proteins which serve as regulatory proteins. To illustrate the proposed algorithm, this section will consider the parameter identification of one simple gene regulatory network with one gene, two operators and one promoter as shown in Figure 1.

Based on the statistical thermodynamic theory and biochemical kinetics, the model of this network can be expressed as follows [1]:

$$\dot{x} = \frac{a_0 + a_1 x + a_2 x^2}{1 + b_1 x + b_2 x^2} - \lambda x \quad (16)$$

where x is the concentration of protein encoding by the gene, a_i (i = 0,1,2) and b_i (i = 1,2) are positive constants related to the biochemical kinetics and λ is a positive constant representing the protein degradation rate. One can see that model (16) is a rational model with positive parameters λ, a_i and b_i. Note that model (16) is slightly different from the one in reference [1]. To uniquely identify parameters in model (16), we have rescaled the parameters such that the constant term in the denominator is 1.

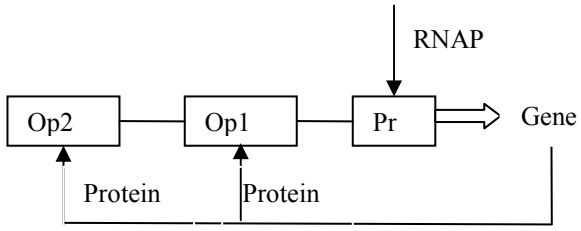

Figure 1. A gene regulatory network with one gene, two operators (Op1 and Op2) and one promoter (Pr).

In this study, a group of artificial data is generated from the model of gene regulatory system (16), with nominal parameter values and initial states provided. The nominal values of parameters are set as: a_0=0.4, a_1=2.8, a_2=0.24, b_1=0.5, b_2=1.4, λ=0.4. In this example, we use the nominal values to generate the trajectory of $x(t)$ shown as in Figure 2, The time starts at t= 0s. From Figure 2, system (16) is stable at its steady state x^* = 2.18 after 10s. Therefore, we don't use the simulated data after 10 seconds.

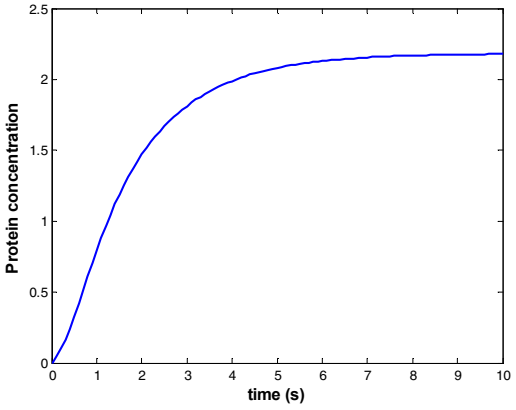

Figure 2. Trajectory of system (16)

There is no noise added on the artificial data in the simulation, so they can be considered as noise-free measurements. Nevertheless, unreasonable noises can be introduced in numerically calculating the derivatives by finite difference formulas. In general, the higher the sampling frequency and more data points are used, the more accurate the numerical derivatives are. On the other hand, in practice we may not obtain data with high frequency because of experimental limitations. In this study, the sampling frequency is 100Hz. In numerically calculating the concentration change rate (derivative $\dot{x}(t)$) at each time point from concentration x, we adopt the five-point central finite difference formula as follows.

$$\dot{x}(t_n) = \frac{1}{12\Delta t}[x(t_{n-2}) - 8x(t_{n-1}) + 8x(t_{n+1}) - x(t_{n+2})] \quad (17)$$

After obtaining the derivatives at different sampling points, we can apply the proposed method to estimating parameters in these models. The relative estimation error is

employed to measure the performance of estimation methods. The relative estimation error (REE) is defined as:

$$REE = \frac{\|estimate - true\_value\|}{\|true\_value\|} \quad (19)$$

In system (16), the right-handed side is a linear fractional function plus a function linear in one parameter, which is not the same format as in (1). For the purpose of parameter estimation, we can transform system (16) into the following model

$$\dot{x} = \frac{a_0 + (a_1 - \lambda)x + (a_2 - b_1\lambda)x^2 - b_2\lambda x^3}{1 + b_1 x + b_2 x^2} \quad (20)$$

although the numerator is not linear in 6 original parameters any more in model (16). However, if we view a single coefficient as a new parameter, the numerator in model (20) is linear in the new parameters. In addition, as there are 6 parameters in 6 coefficients in model (20), using our method the 6 original parameters can be uniquely identified.

Parameters in model (16) are estimated by using three methods: 1) directly minimizing the objective function (4) by nonlinear optimization method; 2) minimizing the objective functions (7) and (8) by nonlinear optimization method; and 3) minimizing the objective functions (7) and (8) by the proposed iteration formula (15). These three methods are all iterative-type, and need the initial values to start-up. In this study, initial values are chosen as true values plus a relative Gaussian noise, i.e.

$$Initial\ value = true\_values \cdot (1 + \sigma \bullet \varepsilon) \quad (21)$$

where ε follows the standard normal distribution and σ is the standard deviation. These methods are implemented in Matlab version 7.01, MS Windows XP Professional SP2, Pentium [R] D CPU2.80GHZ and 2.00 GB of RAM. In Methods 1) and 2), the Matlab embedded function *fminsearch()* is directly called.

When the initial values of all 6 parameters are initialized by formula (21) with σ=1, method 1) converged to the values far from the true value in 20 out of 20 (=100%) runs. When the initial values of 2 parameters in denominator are initialized by formula (21) with σ=1 and others are taken the true values, method 1) converged to the values far from the true value in 18 out of 20 (=90%) runs. This indicates that the nonlinear optimization is not effective to estimate the parameters in the rational model when the initial values moderately deviate from the true values.

In the following we mainly compare the proposed method with method 2) to minimize the objective function (7) and (8). These two methods only need the initial values of parameters in the denominator. The initial values of 2 parameters in the denominator are initialized by formula (21) with σ=1. The results are listed in Table 1. REE is for the minimum REE over 20 runs. The CPU time is average running time over 20 runs. Robustness is the percentage of runs converging with the minimum REE.

Table 1. Comparison between proposed method and nonlinear optimization method

	CPU Time (s)	REE	Robustness
Proposed method	3.3657	0.0236	90%
Traditional method	26.3656	0.0236	80%

From Table 1, the proposed method shows the same estimation accuracy as the nonlinear optimization method if both methods converge in terms of the relative estimation error. However, the proposed method uses much less CPU time to converge and is more robust (insensitive) to the initial values than the nonlinear optimization method.

IV. CONCLUSION

In this paper, we have developed a method for estimating parameters in the rational models of molecular biological systems. The results from the illustrative example have shown that the proposed method consumes much less CPU time to converge and is more robust to the initial values than the nonlinear optimization method. In this study, we do not consider the noises in the data except those introduced by numerical derivatives. One direction of future work is to investigate the robustness of the proposed method to noises in the data. In addition, low sampling frequency is expected in practice, particularly for biological systems. Another direction of future work is to investigate the performance of the proposed method with low sampling frequency.

REFERENCES

[1] D.M. Wolf and F. H. Eeckman, "On the relationship between genomic regulatory element organization and gene regulatory dynamics," *Journal of Theoretical Biology*, 195: 167-186, 1998.

[2] M. Fussenegger, J.E. Bailey, and J. Varner, "A mathematical model of caspase function in apoptosis," *Nature Biotechnology*, 18: 768-774, 2000.

[3] J. Nielsen, J. Villadsen, and G. Liden, "Bioreaction Engineering Principles," 2nd edition, New York: Kluwer Academic/Plenum Publishers, 2003.

[4] G. N. Stephanopoulos, A.A. Aristidou, and J. Nielsen, "Metabolic Engineering: Principles and Methodologies," San Diego: Academic Press, 1998.

[5] K.G. Gadkar, R. Gunawan, F.J. Doyle III, "Iterative approach to model identification of biological networks," *BMC Bioinformatics*, 6:155, 2005.

[6] J.V. Beck and K.J. Arnold, "Parameter Estimation in Engineering and Science," New York: John Wiley & Sons, 1977.

[7] M.R. Osborne, "Some special nonlinear least squares problems", *SIAM Journal on Numerical Analysis*, 12(4): 571-592, 1975.