Insights into the behaviour of systems biology models from dynamic sensitivity and identifiability analysis: a case study of an NF-κB signalling pathway†

Hong Yue,*[ab] Martin Brown,[c] Joshua Knowles,[ab] Hong Wang,[c] David S. Broomhead[bd] and Douglas B. Kell[ab]

Received 3rd July 2006, Accepted 26th September 2006
First published as an Advance Article on the web 23rd October 2006
DOI: 10.1039/b609442b

Mathematical modelling offers a variety of useful techniques to help in understanding the intrinsic behaviour of complex signal transduction networks. From the system engineering point of view, the dynamics of metabolic and signal transduction models can always be described by nonlinear ordinary differential equations (ODEs) following mass balance principles. Based on the state-space formulation, many methods from the area of automatic control can conveniently be applied to the modelling, analysis and design of cell networks. In the present study, dynamic sensitivity analysis is performed on a model of the IκB–NF-κB signal pathway system. Univariate analysis of the Euclidean-form overall sensitivities shows that only 8 out of the 64 parameters in the model have major influence on the nuclear NF-κB oscillations. The sensitivity matrix is then used to address correlation analysis, identifiability assessment and measurement set selection within the framework of least squares estimation and multivariate analysis. It is shown that certain pairs of parameters are exactly or highly correlated to each other in terms of their effects on the measured variables. The experimental design strategy provides guidance on which proteins should best be considered for measurement such that the unknown parameters can be estimated with the best statistical precision. The whole analysis scheme we describe provides efficient parameter estimation techniques for complex cell networks.

1. Introduction

Sensitivity analysis is an important tool in studies of the dependence of systems on their parameters. It is normally used to analyze how sensitive a system is with respect to the change of parameters[1] and is perhaps best known in systems biology *via* the formalism of metabolic control analysis.[2–4] The study of sensitivity analysis helps to identify those parameters that have more impacts on the system output and capture the essential characteristics of the system.[5] It is particularly useful for complex biological networks that involve a large number of variables and parameters. Sensitivity coefficients, which are the partial derivatives of the model states with respect to the model parameters, play an important role in experimental design, parameter estimation, uncertainty analysis, model discrimination and reduction, *etc.* for biological systems.[6] In a recent work on model reduction of complex metabolism models, time-varying local sensitivity analysis has been performed to compose the matrix of normalized sensitivity coefficients, based on which, different methods were used to discard parameters that have less influence on the model dynamics.[7] Using the Monte Carlo method, Cho *et al.* employed multi-parametric global sensitivity analysis on the TNFα-mediated NF-κB signal transduction pathway for experimental design.[8] Schwacke and Voit presented a Taylor integration method for the efficient computation of time-dependent sensitivities for generalized mass action systems, then investigated the effects of different initial species concentrations on the system dynamics.[9]

Sensitivity analysis methods can be classified into two main categories: (1) local sensitivities that provide information on the effect of a small change in each input parameter individually; and (2) global sensitivities that instead describe the effect of simultaneous 'arbitrary' variations of multiple parameters on the dependent variables.[5,10,11] Global sensitivity analysis should be peformed when some parameters are most likely deviated far from the true value or for a rather nonlinear system. Local sensitivity analysis is still the most commonly used method in the area of systems biology when the system parameters are reliably provided by experiments or computation. In this paper, local sensitivities have been calculated and studied to analyze a model of a signalling pathway system. Local sensitivity analysis can be performed in a static way or a dynamic way. Static sensitivity analysis is based on the steady-state response to constant changes in parameters. It can provide adequate description of system behaviour for mechanisms that are under homeostatic control and tend to exhibit uniform behaviour with insignificant transients, in particular

[a]*School of Chemistry, University of Manchester, Sackville St., Manchester, UK M60 1QD. E-mail: h.yue@manchester.ac.uk; dbk@manchester.ac.uk; Fax: +44 (0)1613064556; Tel: +44 (0)1613064492*
[b]*The Manchester Interdisciplinary Biocentre, The University of Manchester, 131 Princess St., Manchester, UK M1 7ND*
[c]*School of Electrical and Electronic Engineering, University of Manchester, Manchester, UK M60 1QD*
[d]*School of Mathematics, University of Manchester, Manchester, UK M60 1QD*
† Electronic supplementary information (ESI) available: Appendices. See DOI: 10.1039/b609442b

systems that are asymptotically stable. However, for general systems with time-varying nominal behaviour, such as signal transduction and regulatory systems, dynamic sensitivity analysis is of primary interest. More emphasis should then be put on the dynamics rather than the instantaneous concentrations of the components.[9,12] Investigation of the transient behaviour in signal transduction networks or any phenomena involving limit cycle oscillations requires a dynamic analysis.[13]

A number of methods have been developed for the computation of local sensitivities such as the finite difference method (FDM), the direct differential method (DDM), the Green's function method (GFM), the analytically integrated Magnus method (AIM), *etc.*[5,11,14–16] These algorithms can be lumped into two general categories: sequential methods (*e.g.* GFM and AIM) and simultaneous methods (DDM).[17] In the most commonly used direct differential method,[18] the differential equations of the system model and the sensitivities are combined into a coupled system and are solved simultaneously. In most cases, sensitivity calculation is a problem of solving stiff ODEs (*i.e.* ODEs where the eigenvalues vary greatly) as biological systems normally involve a large number of reactions and the parameters can span several orders of magnitude. For this reason, continuous efforts have been made to develop efficient and robust integration algorithms for solving sensitivity analysis problems in biological systems.[9,17,19–25]

As sensitivity analysis describes the importance of the model parameters to the measurement variables, it plays an important role in biological model development in an iterative cycle between data analysis, identifiability assessment, measurement set selection, parameter estimation, model validation and experimental design. One crucial issue relating to parameter estimation is identifiability analysis, and this is closely related to parametric sensitivity analysis. Several techniques have been developed for identifiability analysis based on the sensitivity coefficient matrix[26–28] and applied to biological and chemical systems to assist in parameter estimation.[29–31] Another issue is measurement set selection. The optimal measurement set should consist of variables that have maximum information/benefit for parameter identification. This issue assumes significance in the modelling of biological networks because only a limited number of molecules can be tagged with fluorescent proteins to allow their detection. In some recent work on the modelling of biological networks, the Fisher Information Matrix was used to determine the measurement set in order to optimise the quality of parameter estimation in a certain statistical sense.[32,33]

Signal transduction pathways enable cells to receive, process and respond to biochemical stimuli (information). The components of a pathway interact not only with each other but also with components of other pathways, leading to complex cell networks. The nuclear factor κB (NF-κB) signalling pathway (see Fig. 1) is an important cellular signalling pathway, of which protein phosphorylation is a major factor controlling the activation of further downstream events.[34,35] The NF-κB proteins are a group of mainly dimeric nuclear transcription factors involved in a range of cell responses including immune and inflammatory reactions as well as the regulation of apoptosis.[36] NF-κB is normally held inactive in the cytoplasm by being bound to IκB (inhibitory κB) isoforms. In response to extracellular signals such as tumor necrosis, IKK (IκB kinase) is transformed from its neutral form into its active form, a form capable of phosphorylation and degradation of IκBα. Degradation of IκBα releases the main activator NF-κB, which then translocates to the nucleus and triggers transcription of numerous genes including IκB.[37] NF-κB regulation of IκBα transcription represents a delayed negative feedback loop that drives oscillations in NF-κB translocation.[38] Understanding of this system is required if we are to explore the therapeutic potential of NF-κB as a drug target for chronic inflammatory diseases, cancer, infections, chemotherapy, the immune system, *etc.*[39–41]

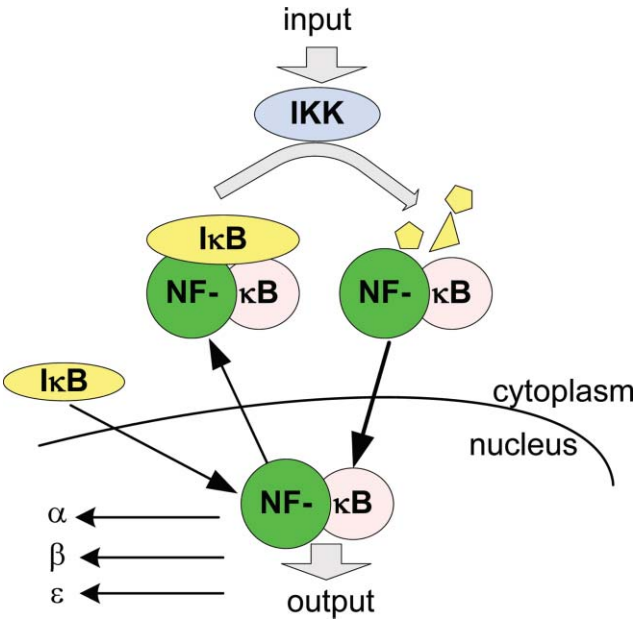

Fig. 1 IκB–NF-κB signal pathway module.[45]

The IκB–NF-κB system is the central signalling module of the NF-κB pathway. It acts to transduce all the NF-κB response from the activation of Inhibitor-κB kinase (IKK) to the transport rates into and out of the nucleus of each of the components. We have analysed the static, local sensitivities and the effects of dual modulation of critical parameters on features selected from the concentration profile of NF-κB in the nucleus ($NF\text{-}\kappa B_n$).[8,42,43] This type of feature sensitivity analysis is mostly used for oscillating reactions. However, the interpretation of feature sensitivities is not straightforward in general. It refers only to the importance of a parameter with respect to the specific features.[11] In this work, we implement dynamic sensitivity analysis to the NF-κB signal transduction pathway by considering the effects of parameters on the dynamic responses of multiple variables.

This paper is organized as follows. The nonlinear states model of the IκB–NF-κB signal pathway is given in section 2. Section 3 briefly introduces the preliminaries on dynamic sensitivity analysis and its relationship with parameter estimation. Implementation of sensitivity analysis for single and multiple variables is performed in section 4. A group of sensitive parameters have been identified with the univariate

analysis. In section 5, correlation analysis and identifiability analysis are investigated based on the relative sensitivity matrix and the best measurement set is decided by a forward selection algorithm using the modified E-optimal criteria. Finally, conclusions are made in section 6. The detailed IκB–NF-κB reaction mechanism, the concentration profiles of the reaction species and the orthogonal procedure used in identifiability analysis are provided in the appendix.

2. Signal pathway state-space model

A state-space model is a convenient way to describe a nonlinear system in terms of first-order differential equations only.[44] The mechanism of a simplified IκB– NF-κB signal pathway is described by Hoffmann *et al.*,[45] Nelson *et al.*[38] and (slightly differently) by Lipniacki *et al.*,[37] in which there are 26 reaction species participating in 64 reactions. Out of the 26 reaction species, 24 species are changing dynamically and their concentrations are defined as the state variables in Table 1 (i stands for the ith reaction species).

Details of the 64 kinetic reactions can be found in Appendix A1, which are summarized from the published works.[42,45] The reaction rate is denoted as $k_j (j = 1, \ldots 64)$. Following the mass balance principle, 24 ordinary differential equations can be written for the concentration dynamics of the reaction species.

$$
\begin{aligned}
\dot{x}_1 &= -(k_{37} + k_{38})x_1 + k_2x_3 + k_{35}x_8 + k_{39}x_{16} + k_{36}x_{22} - k_1x_1x_2 - k_{34}x_1x_{10}\\
\dot{x}_2 &= -k_{19}x_2 + (k_2 + k_{16})x_3 + (k_4 + k_{17})x_5 + (k_6 + k_{18})x_7 + (k_8 + k_9)x_9 + (k_{11} + k_{12})x_{12} + (k_{14} + k_{15})x_{14} + k_{20}x_{15} - k_1x_1x_2 - k_3x_2x_4 - k_5x_2x_6 - k_7x_2x_8 - k_{10}x_2x_{11} - k_{13}x_2x_{13}\\
\dot{x}_3 &= -(k_2 + k_{16})x_3 + k_{53}x_9 + k_{54}x_{17} + k_1x_1x_2 - k_{52}x_3x_{10}\\
\dot{x}_4 &= -(k_{43} + k_{44})x_4 + k_4x_5 + k_{41}x_{11} + k_{45}x_{18} + k_{42}x_{23} - k_3x_2x_4 - k_{40}x_4x_{10}\\
\dot{x}_5 &= -(k_4 + k_{17})x_5 + k_{56}x_{12} + k_{57}x_{19} + k_3x_2x_4 - k_{55}x_5x_{10}\\
\dot{x}_6 &= -(k_{49} + k_{50})x_6 + k_6x_7 + k_{47}x_{13} + k_{51}x_{20} + k_{48}x_{24} - k_5x_2x_6 - k_{46}x_6x_{10}\\
\dot{x}_7 &= -(k_6 + k_{18})x_7 + k_{59}x_{14} + k_{60}x_{21} + k_5x_2x_6 - k_{58}x_7x_{10}\\
\dot{x}_8 &= -(k_{35} + k_{62})x_8 + k_8x_9 - k_7x_2x_8 + k_{34}x_1x_{10}\\
\dot{x}_9 &= -(k_8 + k_9 + k_{53})x_9 + k_7x_2x_8 + k_{52}x_3x_{10}\\
\dot{x}_{10} &= (k_{35} + k_{62})x_8 + (k_9 + k_{53})x_9 - k_{61}x_{10} + (k_{41} + k_{63})x_{11} + (k_{56} + k_{12})x_{12} + (k_{47} + k_{64})x_{13} + (k_{15} + k_{59})x_{14} - k_{34}x_1x_{10} - k_{52}x_3x_{10} - k_{40}x_4x_{10} - k_{55}x_5x_{10} - k_{46}x_6x_{10} - k_{58}x_7x_{10}\\
\dot{x}_{11} &= -(k_{41} + k_{63})x_{11} + k_{11}x_{12} - k_{10}x_2x_{11} + k_{40}x_4x_{10}\\
\dot{x}_{12} &= -(k_{11} + k_{12} + k_{56})x_{12} + k_{10}x_2x_{11} + k_{55}x_5x_{10}\\
\dot{x}_{13} &= -(k_{47} + k_{64})x_{13} + k_{14}x_{14} - k_{13}x_2x_{13} + k_{46}x_6x_{10}\\
\dot{x}_{14} &= -(k_{14} + k_{15} + k_{59})x_{14} + k_{13}x_2x_{13} + k_{58}x_7x_{10}\\
\dot{x}_{15} &= k_{19}x_2 - k_{20}x_{15} + k_{22}x_{17} + k_{24}x_{19} + k_{26}x_{21} - k_{21}x_{15}x_{16} - k_{23}x_{15}x_{18} - k_{25}x_{15}x_{20}\\
\dot{x}_{16} &= k_{38}x_1 - k_{39}x_{16} + k_{22}x_{17} - k_{21}x_{15}x_{16}\\
\dot{x}_{17} &= -(k_{22} + k_{54})x_{17} + k_{21}x_{15}x_{16}\\
\dot{x}_{18} &= k_{44}x_4 - k_{45}x_{18} + k_{24}x_{19} - k_{23}x_{15}x_{18}\\
\dot{x}_{19} &= -(k_{24} + k_{57})x_{19} + k_{23}x_{15}x_{18}\\
\dot{x}_{20} &= k_{50}x_6 - k_{51}x_{20} + k_{26}x_{21} - k_{25}x_{15}x_{20}\\
\dot{x}_{21} &= -(k_{26} + k_{60})x_{21} + k_{25}x_{15}x_{20}\\
\dot{x}_{22} &= k_{27}S - k_{29}x_{22} + k_{28}x_{15}^2\\
\dot{x}_{23} &= k_{30}S - k_{31}x_{23}\\
\dot{x}_{24} &= k_{32}S - k_{33}x_{24}
\end{aligned}
\tag{1}
$$

These ODEs include linear and bilinear terms of the state variables. Denoting

$$X = [x_1\ x_2\ \ldots\ x_{24}]^T \tag{2}$$

Table 1 NF-κB reaction species and the states

i	Participant species	i	Participant species
1	IκBα, x_1	14	IKKIκBε–NF-κB, x_{14}
2	NF-κB, x_2	15	NF-κB$_n$, x_{15}
3	IκBα–NF-κB, x_3	16	IκBα$_n$, x_{16}
4	IκBβ, x_4	17	IκBα$_n$–NF-κB$_n$, x_{17}
5	IκBβ–NF-κB, x_5	18	IκBβ$_n$, x_{18}
6	IκBε, x_6	19	IκBβ$_n$–NF-κB$_n$, x_{19}
7	IκBε–NF-κB, x_7	20	IκBε$_n$, x_{20}
8	IKKIκBα, x_8	21	IκBε$_n$–NF-κB$_n$, x_{21}
9	IKKIκBα–NF-κB, x_9	**	Source ($S = 1$)
10	IKK, x_{10}	22	IκBα–t, x_{22}
11	IKKIκBβ, x_{11}	**	Sink (sink = 0)
12	IKKIκBβ–NF-κB, x_{12}	23	IκBβ–t, x_{23}
13	IKKIκBε, x_{13}	24	IκBε–t, x_{24}

as the state vector, and

$$\theta = [k_1\ k_2\ \ldots\ k_{64}]^T \tag{3}$$

as the parameter vector, model (1) can be represented as

$$\dot{X} = f(X,\theta,t),\ X(t_0) = X_0 \tag{4}$$

where $f(\cdot)$ is a nonlinear function, X_0 is the initial states vector at t_0. Model (4) can be described in a more general case with $X \in \mathbb{R}^n$, $\theta \in \mathbb{R}^m$, where n is the number of states and m is the number of parameters. The parameters in θ may include rate coefficients, Michaelis–Menten parameters, Arrhenius parameters, *etc.*

For this system (eqn (1)), the first-order derivatives of state variables are linearly linked with the parameter vector, therefore, an alternative formulation is

$$\dot{X} = g(X)\theta \tag{5}$$

$g(X) \in \mathbb{R}^{n\times m}$ is the nonlinear function matrix reflecting the structure of system ODEs.

The concentration profiles of the 24 reaction species are illustrated in Appendix A2.† The oscillatory behaviour can be clearly observed from the time response curves of several variables. In this case study, measurement data used are pseudo-experimental, *i.e.*, generated by computer simulation of the model.

3. Preliminaries

3.1 Dynamic sensitivities

Denoting x_i as the ith state in X, θ_j the jth parameter in θ, the effect of the parameter change in θ_j to the state of a species of interest, x_i, can be expressed by a Taylor series expansion:

$$
\begin{aligned}
x_i(\theta_j + \Delta\theta_j, t) = x_i(\theta_j, t) &+ \sum_{j=1}^{m} \frac{\partial x_i}{\partial \theta_j}\Delta\theta_j\\
&+ \frac{1}{2}\sum_{l=1}^{m}\sum_{j=1}^{m} \frac{\partial^2 x_i}{\partial\theta_l\partial\theta_j}\Delta\theta_l\Delta\theta_j + \cdots
\end{aligned}
\tag{6}
$$

In eqn (6), the partial derivatives $\partial x_i/\partial\theta_j$ are called the first-order local concentration sensitivity coefficients, while $\partial^2 x_i/\partial\theta_l\partial\theta_j$ are the second-order local concentration sensitivity coefficients, *etc.*[11] Normally only the first-order sensitivity

coefficients are considered. The absolute sensitivity matrix is defined as

$$S = \frac{\partial X}{\partial \theta} = \begin{bmatrix} s_{1,1} & s_{1,2} & \cdots & s_{1,m} \\ s_{2,1} & s_{2,2} & \cdots & s_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n,1} & s_{n,2} & \cdots & s_{n,m} \end{bmatrix} \quad (7)$$

where $s_{i,j} = \partial x_i/\partial \theta_j$.

Matrix S can be obtained conveniently by differentiation if the analytical solution of the ODEs in eqn (4) is available. Unfortunately, this is very rare for cell network systems whose dynamics are described by complex nonlinear ODEs. Therefore, numerical methods have to be applied to obtain S at each sample time. The two most commonly used numerical methods are FDM and DDM. The finite difference approximation is used in FDM, in which the sensitivity coefficient $s_{i,j}$ is calculated from the difference of the nominal and perturbed solutions

$$s_{i,j}(t) = \frac{\partial x_i(t)}{\partial \theta_j} = \frac{x_i(\theta_j + \Delta\theta_j, t) - x_i(\theta_j, t)}{\Delta \theta_j} \quad (8)$$

This method is straightforward in that only the calculation of x_i is required with nominal and perturbed parameters. However, the numerical values obtained may vary significantly with $\Delta\theta_j$, and repeated solution of the model is required at least once for each parameter. It also implies inherent discontinuities with respect to the initial state parameter.

In this work, we use DDM to calculate the local sensitivities as a function of time. Taking the partial derivative of eqn (4) with respect to θ_j yields the following set of absolute sensitivity differential equations

$$\frac{d}{dt}\frac{\partial X}{\partial \theta_j} = \frac{\partial f}{\partial X}\frac{\partial X}{\partial \theta_j} + \frac{\partial f}{\partial \theta_j} = J \cdot S_j + F_j \quad (9)$$

where

$$J = \frac{\partial f}{\partial X} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}, \quad F_j = \frac{\partial f}{\partial \theta_j} = \begin{bmatrix} \frac{\partial f_1}{\partial \theta_j} \\ \frac{\partial f_2}{\partial \theta_j} \\ \vdots \\ \frac{\partial f_n}{\partial \theta_j} \end{bmatrix}$$

are referred to as the Jacobian matrix, the parametric Jacobian matrix, and

$$S_j = \frac{\partial X}{\partial \theta_j} = \begin{bmatrix} s_{1,j} \\ s_{2,j} \\ \vdots \\ s_{n,j} \end{bmatrix} \quad (10)$$

is the column sensitivity vector with respect to the jth parameter. The initial conditions of $s_{i,j}$ can be obtained by differentiation of the initial condition of $X(t_0)$ in eqn (4) as follows:

$$s_{i,j}(t_0) = \delta(\theta_j - x_i^0) \quad (11)$$

where δ is the Kronecker delta function, X_i^0 is the initial value of the ith species. By solving the n equations in eqn (4) together with the n equations in eqn (9) as a set of differential equations, *i.e.*,

$$\begin{cases} \dot{X} = f(X, \theta, t), & X(t_0) = X_0 \\ \dot{S}_j = J \cdot S_j + F_j, & S_j(t_0) = S_0 \end{cases} \quad (12)$$

both $X(t)$ and $\partial X(t)/\partial \theta_j$ can be determined simultaneously. In real systems analysis, the following relative sensitivities are normally used instead of $s_{i,j}$ to allow direct comparison of responses at different states or across different parameters.

$$\bar{s}_{i,j} = \frac{\partial x_i/x_i}{\partial \theta_j/\theta_j} = \frac{\partial x_i}{\partial \theta_j} \cdot \frac{\theta_j}{x_i} \quad (13)$$

3.2 Least squares estimation and dynamic sensitivities

Given the model structure and a set of experiment data of the measured variables, the target of parameter estimation is to determine unkonwn parameters so as to match the measured data with the best statistical quality. This can be achieved by minimizing a cost function that measures the distance between the measured and predcited data profiles. Under the assumption that the measurement errors are uncorrelated and normally distributed with zero mean and constant variance, the weighted least squares criterion can be used for parameter estimation.

$$J(\theta) = \frac{1}{2}\sum_k \sum_i \omega_i (\tilde{x}_i(k) - x_i(k, \theta))^2 \quad (14)$$

Here $\tilde{x}_i(k)$ and $x_i(k,\theta)$ are the measured and predicted values at sample time k, respectively, ω_i are the weights to normalize the contributions of different state variables and can be taken as

$$\omega_i = \left(\frac{1}{\max_k(\tilde{x}_i(k))} \right)^2 \quad (15)$$

The gradient of $J(\theta)$ with respect to the jth parameter θ_j is expressed as

$$g = \frac{\partial J}{\partial \theta_j} = -\sum_k \sum_i \omega_i r_i(k) \frac{\partial x_i(k,\theta)}{\partial \theta_j} = -\sum_k \sum_i r_i(k) \omega_i s_{i,j}(k) \quad (16)$$

where $r_i(k) = \tilde{x}_i(k) - x_i(k,\theta)$ is defined as the residual at time k. The curvature of $J(\theta)$ (Hessian matrix) can be found by calculating the second-order derivative as

$$\frac{\partial^2 J}{\partial \theta_j \partial \theta_l} = \sum_k \sum_i \omega_i s_{i,j}(k) s_{i,l}(k) - \sum_k \sum_i r_i(k) \omega_i \frac{\partial s_{i,j}(k)}{\partial \theta_l} \quad (17)$$

The second term in (17) can be neglected when the residuals are small. The element of the Hessian matrix at (j,l) is then approximated by

$$H(j,l) = \sum_k \sum_i \omega_i s_{i,j}(k) s_{i,l}(k) \quad (18)$$

Therefore, the Hessian matrix should be represented as

$$H = \tilde{S}^T \tilde{S} \tag{19}$$

where

$$\tilde{S} = \begin{bmatrix} \tilde{S}(1) \\ \tilde{S}(2) \\ \vdots \\ \tilde{S}(N) \end{bmatrix}, \quad \tilde{S}(k) = \left[s_{i,j}(k) \cdot \sqrt{\omega_i} \right] \in \mathbb{R}^{n \times m} \tag{20}$$

$k = 1$ to N are the sample time. It can be seen that the gradient and Hessian matrix are closely related to the sensitivity coefficients. They can be used in any gradient-based, first, second or quasi-second order algorithms for parameter estimation. As such, sensitivity analysis provides crucial information for general least squares parameter estimation.

4 Dynamic sensitivity analysis

4.1 Initial conditions

The initial conditions of the state variables are taken from the equilibrium states, which is run from $t = 0$ to $t = 2000$ min. At $t = 0$, all the initial values in X are set to be zeros except for that of NF-κB. The concentration of NF-κB at $t = 0$ is taken to be 0.1. The equilibrium states are used as the initial conditions of X_0 at t_0 with the exception of the species IKK. IKK is treated as the activator and its initial value is set to be 0.1 at t_0. The initial conditions of the dynamic sensitivities are taken from formulation in eqn (11).

For all the calculations relating to dynamic sensitivities and other analysis in this work, the dynamic time length is 400 min and the sample frequency is 1 per min, *i.e.*, $N = 400$ in eqn (20). Also, as the change in concentration profiles of x_{23} and x_{24} are too small compared with other variables, only the first 22 variables in Table 1 are considered for analysis.

4.2 Dynamic sensitivity analysis with a single variable

For the dynamic sensitivity analysis performed on the basis of one variable, the following L^2-norm performance is used to measure the relative coefficients alone the time axis

$$RS_{i,j} = \frac{1}{N} \sqrt{\sum_{k=1}^{N} \left| \bar{s}_{i,j}(k) \right|^2} \tag{21}$$

Here k is the time instance and N is the total number of sampling points. Fig. 2 illustrates the value distribution of $RS_{i,j}$.

Taking the 15th state, NF-κB$_n$, as the variable (Fig. 3 shows the time series of its concentration), parameters can be ranked by $RS_{15,j}$ in descending order as denoted in vector ***S1***.

S1 = [k_{29} k_{36} k_{28} k_{38} k_{52} k_{61} k_9 k_{62} k_{37} k_{21} k_{34} k_1 k_{19} k_{10} k_7 k_{54} k_{27} k_{53} k_{39} k_{55} k_2 k_{20} k_{35} k_3 k_{12} k_8 k_{44} k_{30} k_{42} k_{58} k_5 k_{31} k_{23} k_{15} k_4 k_{33} k_{32} k_{16} k_{48} k_{56} k_{50} k_{57} k_{22} k_6 k_{60} k_{25} k_{63} k_{45} k_{43} k_{64} k_{11} k_{40} k_{51} k_{59} k_{46} k_{49} k_{41} k_{14} k_{13} k_{24} k_{47} k_{17} k_{18} k_{26}]

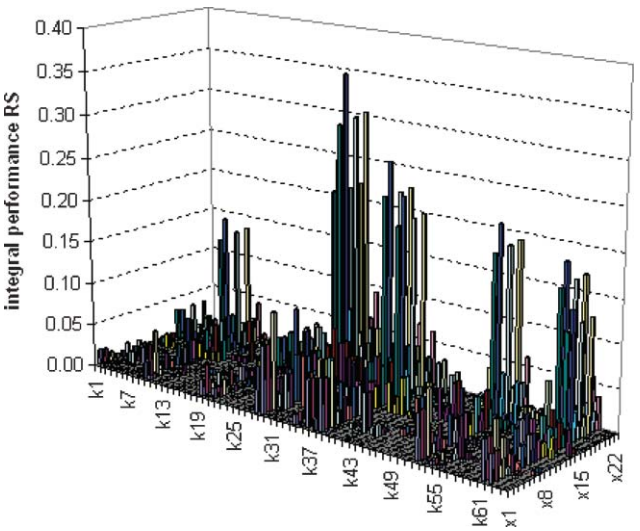

Fig. 2 Sensitivity matrix of integral performance $RS_{i,j}$.

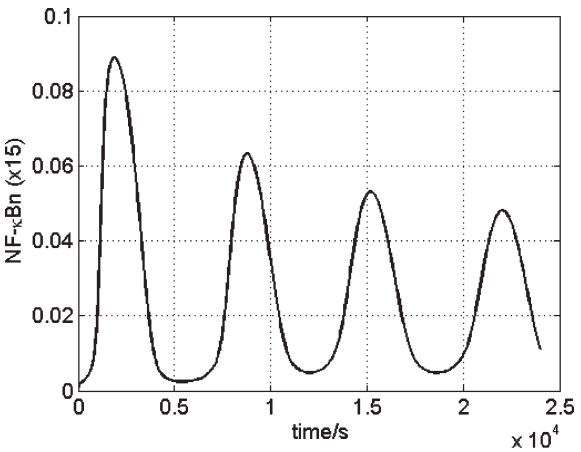

Fig. 3 Time response curve of NF-κB$_n$.

This result can be further illustrated in Fig. 4, which shows all the values of $RS_{15,j}$ from $j = 1$ to $j = 64$. For this group of results, the following 8 parameters can be classified as the most sensitive parameters (in descending order): k_{29}, k_{36}, k_{28}, k_{38}, k_{52}, k_{61}, k_9, k_{62}.

The time courses of the top 8 relative sensitivity coefficients are presented in Fig. 5. It is interesting to note that the dynamic sensitivity profiles of k_{28} and k_{36} are very similar. This suggests that the effects of changes in some parameters are very similar to those of other parameters in this model.

The sensitivity analysis results may be different if another variable is chosen for analysis. Taking the species IκBε–NF-κB (x_7) as an example, the concentration profile of IκBε − NF-κB and the relative sensitivity coefficient $RS_{7,j}$ are shown in Fig. 6 and Fig. 7. The 64 parameters are ranked by $RS_{7,j}$ in the descending order as denoted in vector ***S2***.

S2 = [k_{31} k_{29} k_{58} k_{48} k_{33} k_{32} k_{61} k_{36} k_{38} k_{28} k_{52} k_9 k_{62} k_{34} k_5 k_{46} k_{37} k_{15} k_{21} k_1 k_{19} k_{49} k_{50} k_{59} k_{25} k_6 k_{47} k_{64} k_{60} k_{10} k_{54} k_7 k_{35} k_{53} k_{27} k_{55} k_{12} k_2 k_3 k_{39} k_{20} k_8 k_{44} k_{23} k_{18} k_{14} k_{30} k_{42} k_{13} k_4 k_{51} k_{16} k_{56} k_{57} k_{26} k_{22} k_{40} k_{63} k_{45} k_{41} k_{11} k_{43} k_{24} k_{17}]

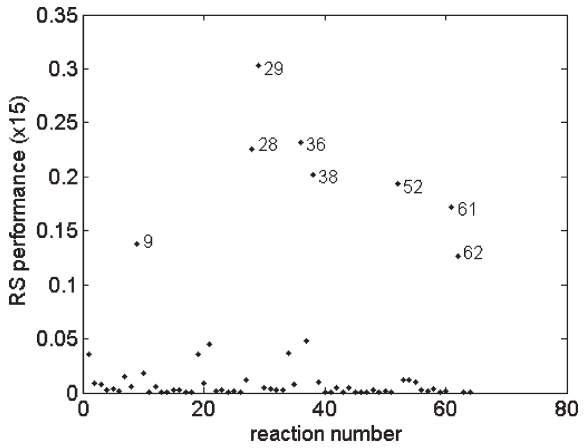

Fig. 4 Performance index $RS_{15,j}$ (variable: NF-κB$_n$).

Comparing ***S1*** with ***S2***, it is clearly seen that for a multi-variable system, the conclusions of dynamic sensitivity analysis depend heavily on the variable chosen. The question then arises as to whether there are measures of sensitivity that contain useful information on all the variables, or at least the most important ones.

4.3 Dynamic sensitivity analysis with multiple variables

A simple way to consider the overall effect of a parameter change to all (or multiple) species is to use the Euclidean-norm

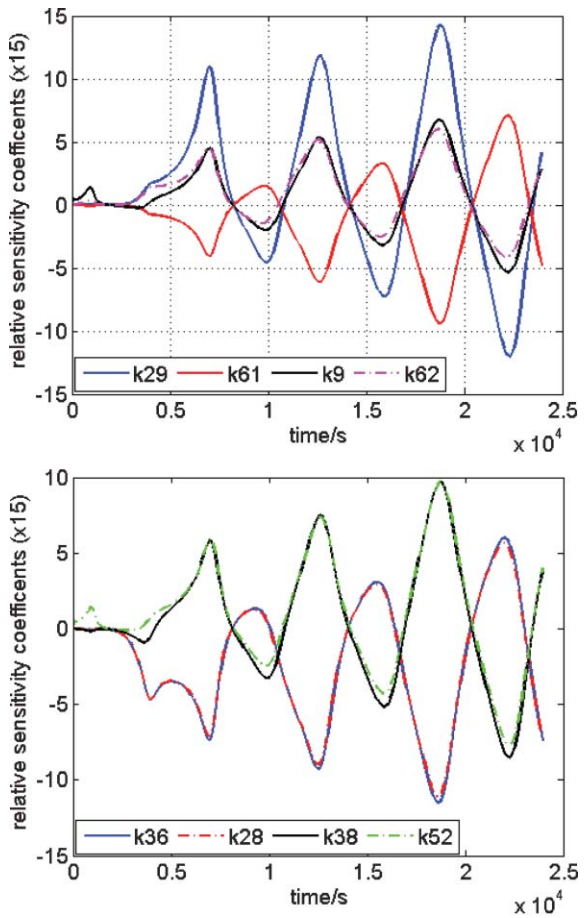

Fig. 5 Profiles of the top 8 sensitivity coefficients for NF-κB$_n$.

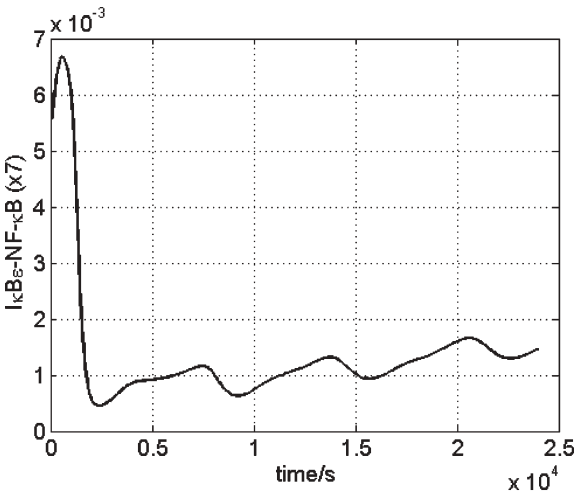

Fig. 6 Time response curve of IκBε –NF-κB.

for sensitivity coefficients

$$OS_j = \frac{1}{N}\sqrt{\sum_{k=1}^{N}\sum_{i=1}^{p}\left|\bar{s}_{i,j}(k)\right|^2} \tag{22}$$

OS_j groups the overall impact of the involved reaction species with respect to the jth parameter. This performance is also termed the overall sensitivity in some literature. The calculation results of OS_j with 22 variables for the 64 parameters are presented in Fig. 8.

The 64 parameters are ranked in the following descending order denoted by ***S3***.

S3 = [k_{29} k_{36} k_{28} k_{38} k_{52} k_{61} k_9 k_{62} k_{31} k_{21} k_{48} k_{37} k_{42} k_{33} k_{32} k_{30} k_{19} k_{34} k_1 k_{54} k_{55} k_{12} k_{58} k_{50} k_{44} k_{60} k_{57} k_{25} k_{15} k_{23} k_3 k_5 k_{10} k_{46} k_{40} k_7 k_{64} k_4 k_{27} k_{53} k_{63} k_{41} k_{39} k_{35} k_6 k_{43} k_{20} k_{47} k_2 k_{49} k_8 k_{56} k_{11} k_{59} k_{14} k_{13} k_{16} k_{22} k_{26} k_{24} k_{45} k_{51} k_{17} k_{18}]

Based on the calculation results of OS_j, the following 8 parameters are considered to be the most sensitive parameters for the IκB–NF-κB signal pathway: k_{29}, k_{36}, k_{28}, k_{38}, k_{52}, k_{61}, k_9, k_{62}. They are defined in the following reactions.

k_{29}: IκBα$_{-t}$ → sink

k_{36}: IκBα$_{-t}$ → IκBα + IκBα$_{-t}$

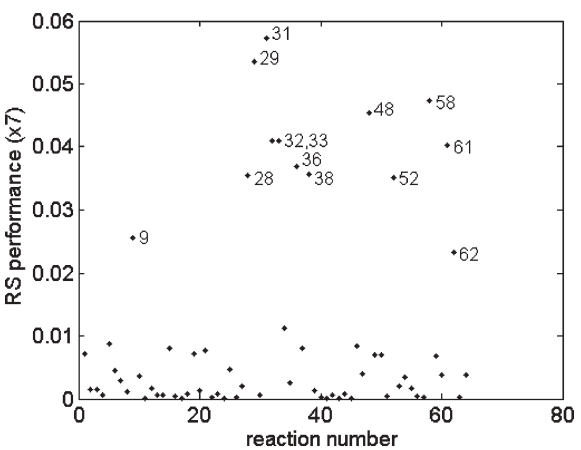

Fig. 7 Performance index $RS_{7,j}$ (IκBε–NF-κB).

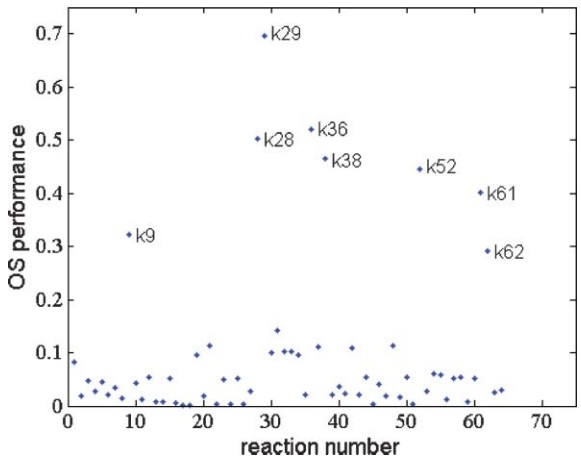

Fig. 8 Overall integral performance OS_j in natural order.

k_{28}: NF-κB$_n$ + NF-κB$_n$ → IκBα$_{-t}$ + NF-κB$_n$ + NF-κB$_n$
k_{38}: IκBα → IκBα$_n$ (Import)
k_{52}: IKK + IκBα–NF-κB → IKKIκBα–NF-κB
k_{61}: IKK → sink
k_9: IKKIκBα–NF-κB → IKK + NF-κB
k_{62}: IKKIκBα → IKK

Three reaction species, NF-κB, IKK, IκBα and their compounds are of most interest.

This result is similar to the static sensitivity analysis,[42] where it is claimed that the above 8 parameters and also k_{34} are the most sensitive parameters. It can be observed from the above analysis that three reaction species, *viz.* free IKK, IκBα and NF-κB$_n$, dominate the oscillation behaviour of this signal pathway. The analysis in ref. 42 was based on several oscillating features abstracted from the concentration profile of NF–κB$_n$. The similar results here show, importantly, that the oscillatory features of the NF–κB$_n$ concentration can be regarded as the dominant features for this particular system.

It should be noted that conclusions obtained so far are taken from univariate analysis because the correlation effects between parameters are not considered.

5. Identifiability analysis and measurement set selection *via* multivariate analysis

5.1 Identifiability analysis

In identifiability one is concerned with the question of the theoretical uniqueness of solutions for a given model and experiment.[26] A nonlinear system is said to be structurally identifiable if each set of parameter values yields unique output trajectories.[46] This *a priori* structural identifiability is a necessary condition but obviously not sufficient for successful parameter estimation from real data, as they are normally sparse and noisy. Two additional problems are commonly encountered in practice. (1) A parameter has a weak effect on the measured output. Estimation of such a parameter is difficult because its effect cannot be accurately quantified. (2) The effects of certain parameters on the measured outputs are nearly linearly dependent, resulting in parameter estimations that are highly correlated.[27,33]

To determine the correlations between parameters, the correlation matrix was calculated as follows:

$$M_c = correlation(\bar{S}) \tag{23}$$

where

$$\bar{S} = \begin{bmatrix} \bar{S}(1) \\ \bar{S}(2) \\ \vdots \\ \bar{S}(N) \end{bmatrix}, \quad \bar{S}(k) = \left[\bar{s}_{i,j}(k)\right] \in \mathbb{R}^{n\times m} \tag{24}$$

In the correlation matrix M_c, parameters that are in this sense highly correlated to other parameters have correlation values close to +1 or −1. For the IκB–NF-κB signal pathway, following the calculation of eqn (23) and eqn (24), it turns out that parameters in the pairs of (k_{31}, k_{32}), (k_{31}, k_{33}) and (k_{32}, k_{33}) have exact linear dependence on each other, *i.e.*, the correlation values for each pair are exactly +1 or −1. When a threshold of 0.99 is used, the following parameters are regarded as highly correlated: (k_7, k_8), (k_{16}, k_{37}), (k_{21}, k_{22}), (k_{28}, k_{36}), (k_{34}, k_{35}), (k_{40}, k_{41}), (k_{46}, k_{47}), (k_{52}, k_{53}), (k_{55}, k_{56}), (k_{58}, k_{59}). This correlation analysis supports the results of dynamic sensitivity calculation in section 4.2, where it shows that parameter changes of k_{28} and k_{36} have the similar impacts on the concentration dynamics of NF-κB$_n$.

The correlation behaviour of parameters may cause identifiability difficulties using least-squares estimation techniques. This can be illustrated by Fig. 9 and Fig. 10, in which the cost functions (residual $= \frac{1}{N}\sum_{k=1}^{N}(\tilde{x}_{15}(k) - x_{15}(k,\theta))^2$) with respect to the change of two parameters are shown when one variable x_{13} (IKKIκBε) is considered for the measurement. In Fig. 9, it is difficult to find a unique pair of (k_{28}, k_{36}) corresponding to the minimum cost function, because the sensitivities of this pair of parameters are highly correlated and yield a straight line basin in the cost function; however for the pair of parameters in Fig. 10, a good estimation can more obviously be made by searching for the global miminum point of the cost function.

The orthogonal method developed in ref. 28 takes into account both parameters' effects on model predictions and

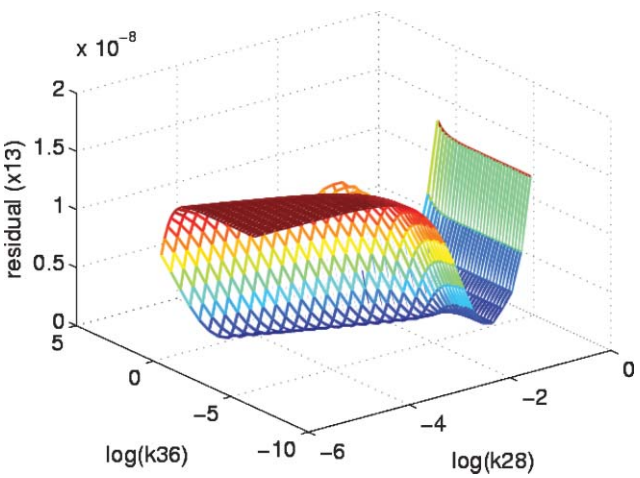

Fig. 9 Cost function wrt (k_{28}, k_{36}).

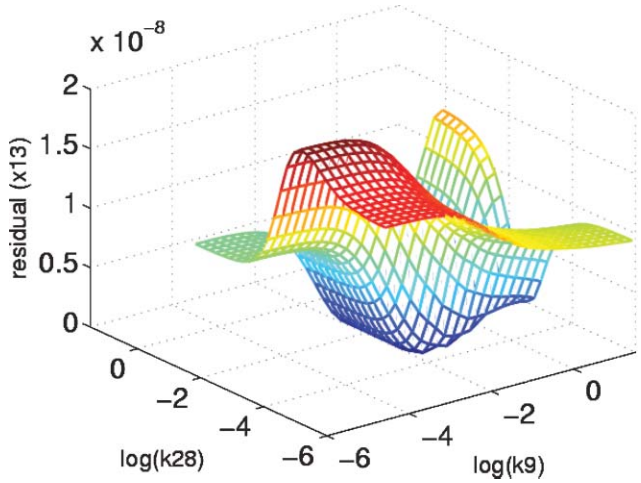

Fig. 10 Cost function wrt (k_9, k_{28}).

correlations between parameters. It can accommodate dynamic models wherein some responses are available at irregular sampling times.[30] The effect of indivial parameters can be determined by examining the magnitude of each column of the relative sensitivity coefficient matrix $\bar{S}$, which corresponds to a particular vector. A large value indicates a large effect of that parameter on the model predictions. The correlation feature is examined by checking whether the columns of the sensitivity coefficient matrix corresponding to the set of estimable parameters are correlated with each other. Implementation of this algorithm can be found in appendix A3. The algorithm can be interpreted as a forward selection procedure where the parameter being selected is the one with the highest t-ratio.

After calculation, the ranking result for all 64 parameters in the IκB–NF-κB model is as follows (see Fig. 11 for illustration):

Identifiability ranking: $\boldsymbol{I1} = [k_{29}\, k_{36}\, k_{31}\, k_{61}\, k_{42}\, k_{38}\, k_{52}\, k_{19}\, k_9\, k_{21}\, k_{50}\, k_{54}\, k_{44}\, k_{28}\, k_{58}\, k_{15}\, k_{55}\, k_{12}\, k_{25}\, k_{23}\, k_{34}\, k_{46}\, k_{40}\, k_{60}\, k_{57}\, k_5\, k_1\, k_3\, k_{48}\, k_{30}\, k_{62}\, k_{64}\, k_4\, k_6\, k_{39}\, k_{37}\, k_{10}\, k_{63}\, k_2\, k_{11}\, k_{27}\, k_{20}\, k_{13}\, k_{45}\, k_{51}\, k_{41}\, k_7\, k_{47}\, k_{43}\, k_{49}\, k_{14}\, k_{26}\, k_{35}\, k_{24}\, k_{56}\, k_8\, k_{59}\, k_{22}\, k_{53}\, k_{17}\, k_{18}\, k_{16}\, k_{32}\, k_{33}]$.

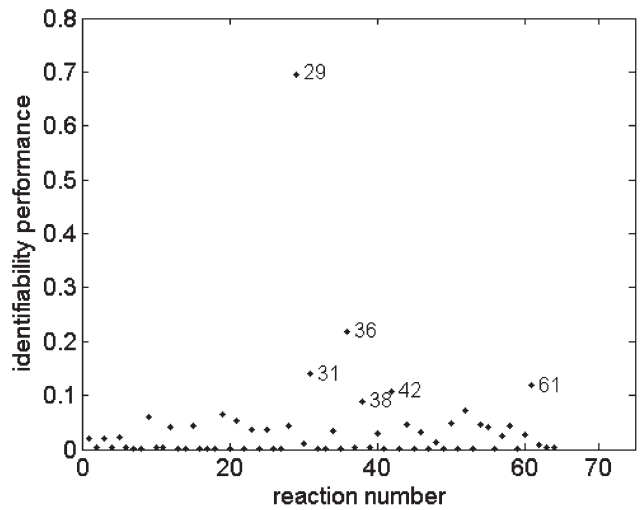

Fig. 11 Parameter identifiablity results by orthorgonal forward selection.

It suggests that k_{29} is the most identifiable parameter, k_{36} is the second most identifiable parameter and so forth. A cut-off threshold can be used to determine the subset of parameters which are practically estimable. The choice of this threshold is normally heuristic. One means of selection is based on the differences between estimability measures. If the measure of the Lth ranked parameter is much greater than that of the $(L + 1)$th ranked parameter, then the threshold can be set between these two parameters. Of course some problems do not allow such a clear distinction. In the work of Yao *et al.*, they suggested a value of 0.04 for the 2-norm measure, which means a 10% parameter change should make at least a 2% variable change for the purpose of parameter estimation.[28] An alternative method is to perform parameter estimation tests to determine the proper number of estimable parameters.[27]

If none of the parameters is correlated to any other, then the ranking order in the sensitivity ranking vector ***S3*** should show the same identifiability ranking of parameters from easy to difficult, *i.e.*, ***S3*** = ***I1***. For the IκB–NF-κB model, however, 13 pairs of parameters are either exactly correlated or highly correlated. Therefore, the descending order of the 64 parameters in the identifiability ranking ***I1*** is different from that of the sensitivity ranking ***S3***. Among the correlated pairs, at least one parameter is moved towards the direction of less estimable from ***S3*** to ***I1***. This clearly shows the difference between the univariate and multivariate analysis.

5.2 Measurement set selection

As performing experiments to obtain rich data for modelling is expensive and time-consuming, measurement set selection aims to find a necessary or a minimum set of variables for the experimental measurements such that the unknown parameters are estimated with the best statistical quality. For this purpose, the sensitivity analysis results can also be interpreted as an estimate of the "observability" of each parameter from the perspective of each variable. An interesting and important question to consider is, "Which variables are the most important, useful or discriminating for parameter estimation?" This question can be framed in a number of different ways, depending on what is meant by a good estimate. In this work, the estimation quality is assessed by the least-squares measure along the variable trajectories.

Under the scheme of the least squares parameter estimation described in section 3.2, the information content of measurements can be quantified by the Hessian matrix H in eqn (19). In general, the smaller the joint confidence intervals for the estimated parameters are, the more information is contained in the measurements.[47] The commonly used optimal design criteria are: the A-optimal design (min trace(H^{-1})), the D-optimal design (max det(H)), the E-optimal design (max $\lambda_{\min}(H)$) and the modified E-optimal design (min $\lambda_{\max}(H)/\lambda_{\min}(H)$). Here $\lambda_{\min}$ and $\lambda_{\max}$ are the minimum and maximum eigenvalues of H, while det indicates the determinant and tr the trace of H.

We use the modified E-optimal design for the measurement set selection. This design minimizes the ratio of the maximum value to the minimum value of the eigenvector, therefore the functional shape of the confidence intervals of the estimated

parameters is optimised. Similar to that of the identifiability analysis, forward selection procedures are performed to find the best sets of variables This algorithm can be performed as follows.

Step 1: Considering the case that one variable is used for parameter estimation ($p = 1$), formulate the corresponding matrix H with one variable at each run. Use the modified E-optimal design to find the first variable for the measurement set.

Step 2: Consider the case that one more variable is used for estimation, augment the matrix H with the new variable. Use the modified E-optimal design to find the next variable to be included in the measurement set.

Step 3: Increase the iteration number p by 1, go to step 2 until all the variables are checked.

The result of the measurement set selection is related to the parameter set to be estimated. When considering the top 8 identifiable parameters in ***II*** to be estimated, the calculation results is shown in Fig. 12. It can be seen that three or four variables will provide good enough estimation for the 8 parameters: k_{29}, k_{36}, k_{31}, k_{61}, k_{42}, k_{38}, k_{52}, k_{19}. In this case, x_{12} (IKKIκBβ–NF-κB), x_{21} (IκBε$_n$–NF-κB$_n$), x_{13} (IKKIκBε) and x_{19} (IκBβ$_n$–NF-κB$_n$) are the top 4 variables to be included in the measurement set. It should be noted that the result will be different if different parameters are to be estimated.

6. Conclusions

Based on the nonlinear state-space formulation, a general scheme of system analysis is provided for the purpose of least squares system identification. It includes dynamic sensitivity analysis, correlation analysis, model identifiability assessment and measurement set selection. Together these will allow efficient parameter estimation of any cellular network.

For the simplified IκB–NF-κB signal pathway system studied in this work, the univariate analysis of dynamic sensitivities shows that out of the 64 parameters in the model, the following 8 parameters have the main impacts on the oscillation behaviour of the nuclear NF-κB when they are varied individually: k_9, k_{28}, k_{29}, k_{36}, k_{38}, k_{52}, k_{61}, k_{62}. It suggests that three reaction species, *viz.* free IKK, IκBα and NF-κB$_n$, dominate the oscillation behaviour of this signal pathway. This conclusion is made without considering the correlations between parameters. Pairwise phenomena are observed from the 8 most sensitive parameters, especially for k_{28} and k_{36}. That is to say, the two parameters have very close effects on the model predictions. Further correlation analysis indicates more pairs of parameters that are in this sense, by their effects, highly correlated to each other, among which, k_{31}, k_{32}, and k_{33} are exactly linearly dependent on each other.

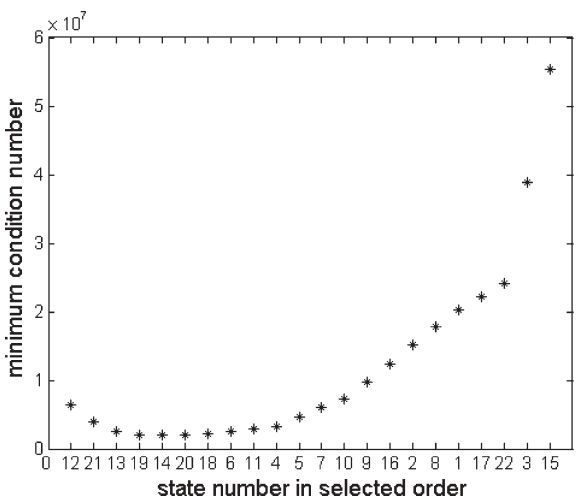

Fig. 12 States selection by the modified E-optimal design.

Using the dynamic sensitivity matrix, identifiability of the parameters is studied *via* a forward selection algorithm. This multivariate analysis shows which parameters are more identifiable and which are less when all the variables are included in the measurement set. Finally, the modified E-optimal design is introduced to select the measurement set for parameter estimation. It is encouraging to see that for the group of the most identifiable parameters, only a small number of variables are needed to be measured to provide satisfactory estimation.

Based on the results herein, different methods of nonlinear parameter estimation in signal pathway systems are now being investigated.

Acknowledgements

This project is supported by the UK Biotechnology and Biological Sciences Research Council (BBSRC). The authors would like to thank Drs Stefano Panzeri, Niklas Ludtke and Steve Wilkinson for helpful discussions.

References

1 B. P. Ingalls and H. M. Sauro, *J. Theor. Biol.*, 2003, **222**, 23–36.
2 D. B. Kell and H. Westerhoff, *FEMS Microbiol. Rev.*, 1986, **39**, 305–320.
3 D. A. Fell, *Understanding the control of metabolism*, Portland Press, London, 1996.
4 R. Heinrich and S. Schuster, *The regulation of cell systems*, Chapman & Hall, New York, 1996.
5 A. Varma, M. Morbidelli and H. Wu, *Parametric Sensitivity in Chemical Systems*, Cambridge University Press, Cambridge, United Kingdom, 1999.
6 D. B. Kell and J. D. Knowles, in *Systems modelling in cellular biology: from concept to nuts and bolts*, ed. Z. Szallasi, J. Stelling and V. Periwal, MIT Press, Cambridge, MA, 2006.
7 D. Degenring, C. Froemel, G. Dikta and R. Takors, *J. Process Control*, 2004, **14**, 729–745.
8 K.-H. Cho, S.-Y. Shin, W. Kolch and O. Wolkenhauer, *Simulation*, 2003, **79**, 726–739.
9 J. H. Schwacke and E. O. Voit, *J. Theor. Biol.*, 2005, **236**, 21–38.
10 P. Seferlis and A. N. Hrymak, *Comp. Chem. Eng.*, 1996, **20**, 1177–1200.
11 T. Turanyi, *J. Math. Chem.*, 1990, **5**, 203–248.
12 D. B. Kell, *Biochem. Soc. Trans.*, 2005, **33**, 520–524.
13 B. P. Ingalls, *J. Phys. Chem.*, 2004, **108**, 1143–1152.
14 M. A. Kramer, H. Rabitz, J. M. Calo and R. J. Kee, *Int. J. Chem. Kinet.*, 1984, **16**, 559–578.
15 H. Rabitz and D. Edelson, in *Oscillations and traveling waves in chemical systems*, ed. R. J. Field and M. Burger, John Wiley and Sons, Editon edn., 1985, pp. 193–222.
16 R. P. Dickinson and R. J. Gelinas, *J. Comp. Phys.*, 1976, **21**, 123–143.
17 J. R. Leis and M. A. Kramer, *Comp. Chem. Eng.*, 1985, **9**, 93–96.
18 R. W. Atherton, R. B. Schainker and E. R. Ducot, *AIChE J.*, 1975, **21**, 441–448.
19 R. Zou and A. Ghosh, *IEE Proc. Syst. Biol.*, 2006, **153**, 79–90.

20 M. R. Kristensen, J. B. Jorgensen, P. G. Thomsen and S. B. Jorgensen, *Comp. Chem. Eng.*, 2004, **28**, 2695–2707.
21 M. Caracotsios and W. E. Stewart, *Comp. Chem. Eng.*, 1985, **9**, 359–365.
22 W. F. Feehery, J. E. Tolsma and P. I. Barton, *Appl. Num. Math.*, 1997, **25**, 41–54.
23 T. Maly and L. R. Petzold, *Appl. Num. Math.*, 1996, **20**, 57–79.
24 J. R. Leis and M. A. Kramer, *ACM Trans. Math. Software*, 1988, **14**, 45–60.
25 D. H. Irvine and M. A. Savageau, *SIAM J. Num. Anal.*, 1990, **27**, 704–735.
26 J. Jacquez and T. Perry, *Am. J. Physiol. Endocrinol. Metab.*, 1990, **258**, E727–E736.
27 R. Li, M. Henson and M. J. Kurtz, *IEEE Trans. Cont. Syst. Technol.*, 2004, **12**, 402–412.
28 K. Z. Yao, B. M. Shaw, B. Kou, K. B. McAuley and D. W. Bacon, *Polym. React. Eng.*, 2003, **11**, 563–588.
29 K. G. Gadkar, J. Varner and F. J. Doyle, III, *IEE Proc. Syst. Biol.*, 2005, **2**, 17–30.
30 B. Kou, K. B. McAuley, C. C. Hsu, D. W. Bacon and K. Z. Yao, *Ind. Eng. Chem. Res.*, 2005, **44**, 2428–2442.
31 D. E. Zak, G. E. Gonye, J. S. Schwaber and F. J. Doyle, III, *Genome Res.*, 2003, **13**, 2396–2405.
32 K. G. Gadkar, R. Gunawan and F. J. Doyle, III, in *BMC Bioinformatics*, 2005.
33 M. Rodriguez-Fernandez, P. Mendes and J. R. Banga, *BioSystems*, 2006, **83**, 248–265.
34 N. D. Perkins and T. D. Gilmore, *Cell Death Differ.*, 2006, **13**, 759–772.
35 B. Tian, D. E. Nowak, M. Jamaluddin, S. Wang and A. R. Brasier, *J. Biol. Chem.*, 2005, **280**, 17435–17438.
36 G. Nelson, L. Paraoan, D. G. Spiller, G. J. C. Wilde, M. A. Browne, P. K. Djali, J. F. Unitt, E. Sullivan, E. Floettmann and M. R. H. White, *J. Cell Sci.*, 2002, **115**, 1137–1148.
37 T. Lipniacki, P. Paszek, A. R. Brasier, B. Luxon and M. Kimmel, *J. Theor. Biol.*, 2004, **228**, 195–215.
38 D. E. Nelson, A. E. C. Ihekwaba, M. Elliott, J. R. Johnson, C. A. Gibney, B. E. Foreman, G. Nelson, V. See, C. A. Horton, D. G. Spiller, S. W. Edwards, H. P. McDowell, J. F. Unitt, E. Sullivan, R. L. Grimley, N. Benson, D. S. Broomhead, D. B. Kell and M. R. H. White, *Science*, 2004, **306**, 704–708.
39 P. P. Tak and G. S. Firestein, *J. Clin. Invest.*, 2001, **107**, 7–11.
40 Y. Yamamoto and R. B. Gaynor, *J. Clin. Invest.*, 2001, **107**, 135–142.
41 M. Karin, Y. Yamamoto and Q. M. Wang, *Nat. Rev. Drug Discov.*, 2004, **3**, 17–26.
42 A. E. C. Ihekwaba, D. S. Broomhead, R. L. Grimley, N. Benson and D. B. Kell, *IEE Proc. Syst. Biol.*, 2004, **1**, 93–103.
43 A. E. C. Ihekwaba, D. S. Broomhead, R. L. Grimley, N. Benson, M. R. H. White and D. B. Kell, *IEE Proc. Syst. Biol.*, 2005, **152**, 153–160.
44 J. Durbin and S. J. Koopman, *Time Series Analysis by State Space Methods*, Oxford University Press, 2001.
45 A. Hoffmann, A. Levchenko, M. L. Scott and D. Baltimore, *Science*, 2002, **298**, 1241–1245.
46 J. A. Jacquez and P. Greif, *Math. Biosci.*, 1985, **77**, 201–227.
47 D. Faller, U. Klingmuller and J. Timmer, *Simulation*, 2003, **79**, 717–725.