Proceedings of the 9th International Symposium on
Dynamics and Control of Process Systems (DYCOPS 2010),
Leuven, Belgium, July 5-7, 2010
Mayuresh Kothare, Moses Tade, Alain Vande Wouwer, Ilse Smets (Eds.)

TuAT4.2

Investigating Bayesian Robust Experimental Design with Principles of Global Sensitivity Analysis

Fei He*, Hong Yue*, Martin Brown**

** Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow,
G1 1QE, UK (Tel: (+44) 141-548-2098; e-mail: fei.he@ eee.strath.ac.uk; hong.yue@eee.strath.ac.uk)
** School of Electrical and Electronic Engineering, University of Manchester, Manchester M60 1QD,
UK (e-mail: martin.brown@manchester.ac.uk)*

Abstract: The purpose of model-based experimental design is to maximise the information gathered for quantitative model identification. Instead of the commonly used optimal experimental design, robust experimental design aims to address parametric uncertainties in the design process. In this paper, the Bayesian robust experimental design is investigated, where both a Monte Carlo sampling strategy and local sensitivity evaluation at each sampling point are employed to achieve the robust solution. The link between global sensitivity analysis (GSA) and the Bayesian robust experimental design is established. It is revealed that a lattice sampling based GSA strategy, the Morris method, can be explicitly interpreted as the Bayesian A-optimal design for the uniform hypercube type uncertainties.

Keywords: Robust experimental design, Bayesian method, global sensitivity analysis, lattice sampling, parameter estimation, systems biology

1. INTRODUCTION

Experimental design is an important part of any system identification processes, especially when the models are complex and the data are sparse and relatively noisy. The purpose of model-based experimental design is to maximise the information gathered for quantitative model identification whilst minimising the experimental efforts. This is of particular interests to modelling and parameter identification for complex biological systems, where data are usually obtained from expensive and time-consuming *in vitro* or *in vivo* experiments. A comprehensive review of model-based experimental design methods for biological and chemical modelling was recently given by Franceschini and Macchietto (2008). Experimental design problems in biochemical network modelling has been reported on design of initial molecular concentrations, external cellular signals, sampling time, measurement variable selection, etc. (Balsa-Canto et al., 2008; Cho et al., 2003; Faller et al., 2003; Kutalik et al., 2004).

Almost all these works are based on the standard optimal experimental design (OED) (Atkinson and Donev, 1992) that aims at maximizing a measure of the Fisher information matrix (FIM), which is functionally dependant on the nominal parameter values (Baltes et al., 1994; Körkel et al., 2004). However, as the model parameters are only approximately estimated *a prior*, the OED results can be unreliable when only poorly estimated parameters are available. Sequential experimental design (Körkel et al., 1999; Silvey, 1980) is a conventional way to deal with the parametric uncertainty, in which the experimental design and parameter estimation are implemented in an iterative manner. In practice, however, the cost of cellular experiments often limits the number of iterations. In the past decade, the maximin and Bayesian robust strategies have been introduced to dynamic model based experimental design (Han and Chaloner, 2004; Körkel et al., 2004; Rojas et al., 2007; Yue et al., 2008). In our recent work, we applied these two robust design approaches into the state measurement set selection problem for biological systems and compared them under different uncertainty levels (He et al., 2010). It shows that how to solve the NP-hard optimization problem in the robust experimental design, and compromise the robustness and computational efficiency are very challenging tasks.

Model-based experimental design is formulated on measures of FIM, which is by definition dependant on the parametric sensitivity coefficients of local sensitivity analysis (LSA). The accurate calculation of parametric sensitivities is therefore crucial for performing an optimal experimental design, as well as evaluating/representing the parametric uncertainty for a robust experimental design. There is little work on the role of sensitivity calculation in experimental design, especially in robust design with model uncertainties. The computation of local sensitivity coefficients is relatively straightforward; however, the coefficient values are often affected by the magnitude of the states or parameters of the system. Also, the interactions between parameters are not considered in LSA. Several formulations of scaled and normalised sensitivities have been proposed (Saltelli et al., 2000) to eliminate the scale effects. This will influence the outcome of experimental design. Thomaseth and Cobelli (1999) defined a new "generalised sensitivity" which takes account of the interactions between parameters.

Compared with LSA, the global sensitivity analysis (GSA) has the advantages of incorporating large parametric variations and also the correlations between parameters. GSA has been used to pre-screen potential models and identify important model parameters, but never for the purpose of experimental design problems (Franceschini and Macchietto, 2008). As we know, GSA represents 'lumped' sensitivity information over the parameter space rather than at a single

point as that in LSA. Therefore it is not used in formulating FIM for optimal experimental design. However, we argue that when considering large model uncertainties, GSA has the similar design principle as the Bayesian robust experimental design. In both cases, the domain of possible inputs (parameters) is defined first, and then the inputs are generated randomly either from a specified probability distribution or from a designed sampling strategy. Computations are performed using each set of inputs, and the results of the individual computations are aggregated into the final result.

The aim of this work is to investigate the links between the lattice-sampling based GSA and the Bayesian robust experimental design, and therefore explore the potential of applying GSA in robust experimental design. In Section 2, the Bayesian robust experimental design approach is briefly introduced, which handles parametric uncertainty by integrating/averaging a prior distribution of parameters in the design process. In Section 3, the principle of lattice sampling GSA and its link with the Bayesian experimental design is investigated. To verify the main result, in Section 4, the Morris GSA is compared with several Bayesian robust designs for a measurement set selection problem of a signal pathway system model. Conclusions are given in Section 5.

2. BAYESIAN ROBUST EXPERIMENTAL DESIGN

For a single biochemical system with n reaction species and m reactions, denote $X = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^T$ as the state vector and $\theta = \begin{bmatrix} \theta_1 & \theta_2 & \cdots & \theta_m \end{bmatrix}^T$ as the vector of parameters. The system model can be represented as:

$$\dot{X} = f(X,t,\theta), \quad X(t_0) = X_0$$
$$y = g(X,t,\theta) + \xi \quad (1)$$

where $f(\cdot)$ is the nonlinear state transition function, X_0 is the initial states vector at t_0, $g(\cdot)$ is the measurement function, $y \in \mathbb{R}^{n_y}$ is the measurement output vector with n_y being the number of measurement output. ξ is assumed to be a zero-mean, Gaussian additive noise vector. Parameter estimation for system (1) can usually be obtained from a least squares criterion.

The FIM quantifies the information content of the experimental data. For a nonlinear dynamic system, the FIM is a nonlinear function of the estimated parameters under the assumption that the measurement noise is independent and identically distributed with a zero-mean Gaussian distribution. Assuming the number of identifiable parameters is n_I, computationally the FIM for experimental design can be expressed as:

$$\text{FIM}(\theta,\omega,t) = \sum_{l=1}^{N} \sum_{i=1}^{n_y} \omega_i S_i^T(t_l,\theta) S_i(t_l,\theta) \quad (2)$$

where S is the n_y by n_I local sensitivity matrix of the measurable states with respect to the identifiable parameters. The (i,j)-th component in S is defined as $s_{ij} = \partial x_i / \partial \theta_j$. S_i in (2) stands for the i-th row in S. N is the number of sampling

points in time axis. Originally, $\omega = \begin{bmatrix} \omega_1 & \cdots & \omega_{n_y} \end{bmatrix}^T$ in FIM is the weighting vector related to the variances of the observation errors (Fisher 1912). It can also be interpreted and used as experimental design variables (Boyd 2004) for design purpose. For example, in the measurement set selection problem, ω_i is taken as the weight assigned to the i-th measurement. To simplify the representation, the time dimension (t) in the FIM is neglected in the following text.

In Bayesian robust experimental design, the parametric uncertainty is represented with a prior distribution of the parameters, $P(\theta)$. The uncertainty effects are integrated or averaged in the admissible parameter space. Since the samples have not been observed yet before the experiments, a general Bayesian design criterion over the parameter space applies (Lindley, 1972).

$$\hat{\omega} = \arg \max_{\omega \in \Omega} \, \underset{\theta \in \Theta}{\text{E}} \left\{ \Phi\left(\text{FIM}(\theta,\omega)\right) \right\}$$
$$= \arg \max_{\omega \in \Omega} \int_{\Theta} \Phi\left(\text{FIM}(\theta,\omega)\right) P(\theta) \, d\theta \quad (3)$$

where $\text{E}\{\cdot\}$ denotes the expectation operator, Ω is the design space for ω. Θ represents the set of admissible parameters. $\Phi(\cdot)$ indicates the widely used "alphabet" experimental design criteria that are normally scalar functions of FIM, such as: A-optimal: $\text{trace}(\text{FIM})$; D-optimal: $\det(\text{FIM})$; E-optimal: $\lambda_{\min}(\text{FIM})$, etc. Different Φ-optimality design criteria correspond to different utility functions employed in Bayesian decision theory (Chaloner and Verdinelli, 1995). For example, D-optimality corresponds to the expected gain in Shannon information, and A-optimality can be derived from a quadratic loss function.

From the computational point of view, the integration operator in (3) can also be placed into the design criterion function Φ to form an alternative representation for the Bayesian design, i.e.

$$\max_{\omega \in \Omega} \Phi\left(\int_{\Theta} \text{FIM}(\theta,\omega) P(\theta) \, d\theta \right) \quad (4)$$

where the FIM is integrated over a prior region in parameter space. Assume that $P(\theta)$ is a uniform distribution function, the FIM integration can be approximated as

$$\text{E}\{\text{FIM}(\theta,\omega)\} = \int_{\Theta} \text{FIM}(\theta,\omega) P(\theta) \, d\theta$$
$$\approx \frac{1}{K} \sum_{r=1}^{K} \left\{ \text{FIM}\left(\theta^{(r)},\omega\right) \right\} \quad (5)$$

where the superscript (r) indicates the r-th sampling set for θ and K is the number of samplings. (5) is known as the Bayesian information matrix (BIM) (Merlé and Mentré, 1995) or expected FIM (Asprey and Macchietto, 2002). The integration of FIM can be calculated by an approximation as in (5), or by numerical quadrature, Laplace approximation or Monte Carlo integration within the uncertainty region. This is also named as Pseudo Φ-optimal Bayesian experimental design in Atkinson's terminology (Atkinson and Donev,

1992). Taking the measurement set selection problem as an example, the BIM can be expressed explicitly as:

$$
\begin{aligned}
\text{BIM}(\theta,\omega) &= \frac{1}{K}\sum_{r=1}^{K}\sum_{l=1}^{N}\sum_{i=1}^{n_y}\left(\text{FIM}(t_l,\theta^{(r)},\omega_i)\right) \\
&= \frac{1}{K}\sum_{r=1}^{K}\sum_{l=1}^{N}\sum_{i=1}^{n_y}\omega_i S_i^T\left(t_l,\theta^{(r)}\right)S_i\left(t_l,\theta^{(r)}\right)
\end{aligned} \quad (6)
$$

Accordingly, a Bayesian E-optimal design can be formed as a semidefinite programming (SDP) problem based on BIM:

$$
\begin{aligned}
&\max_{\omega}\ v \\
&s.t\quad \frac{1}{K}\sum_{r=1}^{K}\sum_{l=1}^{N}\sum_{i=1}^{n_y}\omega_i S_i^T(t_l,\theta^{(r)})S_i(t_l,\theta^{(r)}) \geq v\mathbf{I}_{n_I} \\
&\qquad \sum_{i=1}^{n_y}\omega_i = 1,\quad \omega_i \geq 0
\end{aligned} \quad (7)
$$

$\mathbf{I}_{n_I}$ is an n_I-dimension identity matrix.

3. BAYESIAN ROBUST EXPERIMENTAL DESIGN AND GLOBAL SENSITIVITY ANALYSIS

It can be seen from Section 2 that both the Monte Carlo sampling and the local sensitivity evaluations at each sampling point are applied in the Bayesian experimental design. On the other hand, GSA is meant to consider the interactive variations between all parameters in a large uncertainty domain, where the design result is also produced by aggregating individual local results. It would be interesting to investigate the links between the sampling-based Bayesian experimental design with the lattice-sampling based GSA approaches.

3.1 Lattice Prior Sampling

Prior elicitation is an important step in Bayesian experimental design and performance analysis afterwards. Prior information can be obtained from earlier experiments or from conjectures that motivate the investigation (Clyde, 2001). However, for biochemical systems, prior distribution on kinetic parameters may not always be easily available. Instead, it is more likely to be attained from previous parameter estimations or using historical data from previous experiments to construct a hierarchical normal linear model (Lindley and Smith, 1972). In simulation studies, the prior distribution $P(\theta)$ is often obtained via Monte-Carlo sampling in the most likely parametric uncertainty region. Prior sampling distribution can be uniform across a hypercube or Gaussian across a hyper-ellipse in parameter space. The latter is generally assumed in the Bayesian design literature (Chaloner and Verdinelli, 1995; Clyde, 2001). The common sampling strategy is the random uniform sampling, or more efficient orthogonal sampling, such as Latin square sampling (Iman et al., 1981) and Taguchi orthogonal arrays (Taguchi and Yokoyama, 1993). In this work, a lattice sampling scheme used in the Morris global sensitivity analysis is analyzed for the purpose of studying the Bayesian experimental design.

Consider the model (1) with m parameters. Without loss of generality, a simple lattice sampling scheme can be assumed that each θ_j is scaled in the interval $\left[\theta_j^{lb},\theta_j^{ub}\right]$ with θ_j^{lb} and θ_j^{ub} being the lower and upper bounds. Denote $\mu_j = \theta_j^{ub} - \theta_j^{lb}$ as the length of the interval. Within this interval, p discrete values are uniformly taken as $\left\{\theta_j^{lb},\theta_j^{lb}+\mu_j/(p-1),\cdots,\theta_j^{lb}+(p-2)\mu_j/(p-1),\theta_j^{ub}\right\}$. The parametric sample space, Θ, is then an m-dimensional, p-level grid. A random selection of a value for each θ_j ($j=1,\cdots,m$) from the grid forms a random set ($\theta^{(r)}$) in the parameter space, and the local sensitivity is calculated at each $\theta^{(r)}$. This calculation is repeated K times. Considering $P(\theta_j)$ as a uniform distribution for θ_j, the corresponding Bayesian information matrix is formulated as in (6).

Lattice sampling is efficient and easy to implement for the Bayesian prior approximation. Its efficiency can be further improved when combined with orthogonal sampling strategies (e.g. Latin hypercube). However, the p-level grid sampling strategy is usually only applicable for the uniform hypercube type uncertainties, whereas asymptotic Gaussian distributions are more widely used in Bayesian design for nonlinear models. In this context, the uniform lattice sampling scheme sometimes provides conservative design results.

3.2 Bayesian Experimental Design and Morris GSA

Similar to the Bayesian experimental design, some GSA approaches are also based on the lattice sampling scheme. This is exactly the sampling strategy used in the so-called Morris GSA method (Morris, 1991). The Morris method is a screening GSA method. It is based on the estimation of mean and variance of the functional parametric sensitivity, termed as elementary effect (EE), through a pre-defined random sampling within a hypercube in the parameter space.

Based on the original Morris design, we introduced a few changes in the formulation to adapt to the problem under discussion. Similar to the model formulation in Section 3.1, each parameter, θ_j, is scaled in the interval $\left[\theta_j^{lb},\theta_j^{ub}\right]$ and may take a value from the p discrete points within this interval. The whole parameter space of interest is an m-dimensional, p-level grid. At each sampling set in the parameter space, $\theta^{(r)} \in \Theta$, the elementary effect of the i-th state with respect to the j-th parameter is defined as:

$$
F_{i,j}\left(t,\theta^{(r)}\right) = \left(\frac{x_i(t,\theta^{(r)}+\delta_j e_j) - x_i(t,\theta^{(r)})}{\delta_j}\right)^2 \quad (8)
$$

For the j-th parameter, $\delta_j = \left(\theta_j^{ub}-\theta_j^{lb}\right)\Delta$, Δ is a predetermined multiple of $1/(p-1)$ and is taken to be $p/\left[2(p-1)\right]$ in this work. e_j is a standard Cartesian basis

vector with the j-th component being 1 and all the others zeros. In all the sampling cases, there is $\theta_j \leq \theta_j^{ub} - \delta_j$. Producing one value for the elementary effect, $F_{i,j}(t, \theta^{(r)})$, requires a random selection of m values for all the m parameters from the grid and the evaluation of x_i twice, one at the selected point $\theta^{(r)}$, the other at the point by increasing the j-th parameter in $\theta^{(r)}$ with δ_j. The calculation is repeated K times and the mean value of $F_{i,j}(t, \theta^{(r)})$ is taken as follows:

$$\tilde{s}_{ij} = \frac{1}{N \cdot K} \sum_{r=1}^{K} \sum_{t=1}^{N} F_{i,j}(t, \theta^{(r)}) \qquad (9)$$

The Morris global sensitivity indices are obtained for each variable with respect to each individual parameter in the model. The combination of all these sensitivity measures corresponds to the sum of the diagonal elements of the FIM or equivalently its trace. As we know the A-optimal experimental design optimizes the trace of FIM. In the following, the equivalence between the Morris global sensitivity measure and a Bayesian A-optimal information measure is established. This understanding is useful for experimental design in that it explains what information is contained in the global sensitivity measure and it also provides a computationally simple way to evaluate the Bayesian design using a lattice-based sampling scheme.

Theorem

The Morris global sensitivity measure is equivalent to a (sparse) Bayesian A-optimal information measure, where the prior probability is a function of both the sample space lattice partition and the form of the state functions.

Proof

This result is established firstly by considering what the elementary effects represent in Morris GSA, secondly by deriving the prior parameter distribution, and finally by analyzing the random averaging or expectation operation over the parameter space. Consider a single elementary effect as defined in (8). Giving that the states are continuous functions of the parameters, there exists at least one parameter $\theta^{(r)*}$ such that

$$\frac{x_i(t, \theta^{(r)} + \delta_j e_j) - x_i(t, \theta^{(r)})}{\delta_j} = \frac{\partial x_i(t, \theta^{(r)*})}{\partial \theta_j} \qquad (10)$$

where $\theta^{(r)*} \in \left[\theta^{(r)}, \theta^{(r)} + \delta_j e_j \right]$. This is directly inferred from the intermediate value theorem. This means for any of the random sampling points on the parameter lattice, there exists a nearby point, which lies on the lattice for all but the j-th coordinate, such that the squared gradient at this point is equal to the elementary effect. As the number of all possible lattice points is finite, there exists a corresponding finite set of points, at which the equivalent squared gradient can be evaluated. This set must lie strictly within Θ. Assuming all the sampling sets have the equal opportunities to be taken, the

parameter distribution can be represented by the probability distribution function:

$$P(\theta) = \frac{1}{K} \sum_{r=1}^{K} \delta\left(\theta - \theta^{(r)*}\right) \qquad (11)$$

where δ is the impulse response function. Taking an equivalent random sampling scheme as used in the Morris global sensitivity measure, the overall global sensitivity measure can be expressed as

$$\mathrm{E}_{\theta,j}\left(\mathrm{E}_{i,t}\left(\left(\frac{x_i\left(t, \theta^{(r)} + \delta_j e_j\right) - x_i\left(t, \theta^{(r)}\right)}{\delta_j} \right)^2 \right) \right)$$

$$= \mathrm{E}_{\theta,j}\left(\mathrm{E}_{i,t}\left(\left(\frac{\partial x_i}{\partial \theta_j}\left(t, \theta_j^{(r)*}\right) \right)^2 \right) \right)$$

$$\approx \frac{1}{K} \int_{\Theta} \sum_{i,j,t} \sum_{r=1}^{K} s_{ij}^2\left(t, \theta_j^{(r)*}\right) \delta\left(\theta - \theta^{(r)*}\right) \mathrm{d}\theta \qquad (12)$$

$$= \frac{1}{K} \int_{\Theta} \sum_{r=1}^{K} \mathrm{tr}\left(S^T\left(\theta^{(r)}\right) S\left(\theta^{(r)}\right) \right) \delta\left(\theta - \theta^{(r)*}\right) \mathrm{d}\theta$$

$$= \int_{\Theta} \mathrm{tr}\left(\Sigma^{-1}(\theta) \right) P(\theta) \mathrm{d}\theta$$

It is shown in (12) that the aggregated elementary effects in Morris GSA is equivalent to the Bayesian A-optimal design with a uniform distribution $P(\theta)$ on θ. $\qquad \square$

There are several issues worth being discussed on the equivalence of Morris GSA with Bayesian A-optimal design. (i) While the Morris method considers a uniform partition of the parameter space using the lattice, the induced prior probability function $P(\theta)$ is unlikely to be uniform because of the non-linear dependence of the states on the parameters. This results in the equivalent gradient sample points, $\theta^{(r)*}$, occurring at irregular spaces between the lattice nodes. (ii) Since the gradient is squared in (8), it is difficult to apply this simple scheme to a more general D- or E-optimal design criterion, where the covariance terms in FIM would have to be estimated. To consider the covariance terms, the square operation in (8) should be removed. (iii) A sufficiently large number of lattice samples need to be generated for calculation. The global sensitivity measure is evaluated at each sampling point for each state, time, and parameter, in order to mimic a similar averaging effect in constructing the Bayesian information matrix.

4. CASE STUDY ON MEASUREMENT SELECTION OF A SIGNAL PATHWAY MODEL

In this section, different Bayesian Φ-optimal designs and the Morris GSA are compared for a measurement set selection problem. The simulation study is designed to verify the link between the Morris GSA and the Bayesian A-optimal design. We consider a simplified IκBα-NF-κB signal transduction pathway network as the model for simulation. The model details can be found in the appendix. In the simulation study, the following 5 parameters, $\theta_5, \theta_{12}, \theta_{13}, \theta_{16}, \theta_{18}$, are taken for

parameter estimation since they are regarded as the important parameters from our previous work on sensitivity analysis. The objective of experimental design is to select the most informative state measurements from all the 10 states so as to best facilitate the estimation of these 5 parameters. To address the model uncertainties, a 50% uniform uncertainty around the nominal values is considered for the five parameters to be estimated. A random uniform sampling is used in the Bayesian design and a lattice sampling is employed in Morris GSA design. The state rankings calculated by different approaches are compared in Table I, and the corresponding 95% confidence ellipses are illustrated with two parameters (θ_5 and θ_{12}) in Fig. 1.
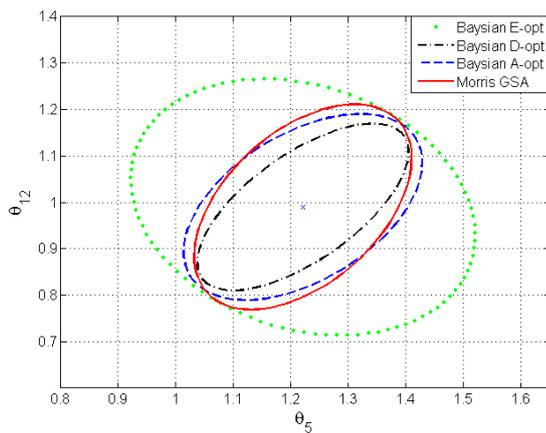
(a) p=5 in Morris GSA design
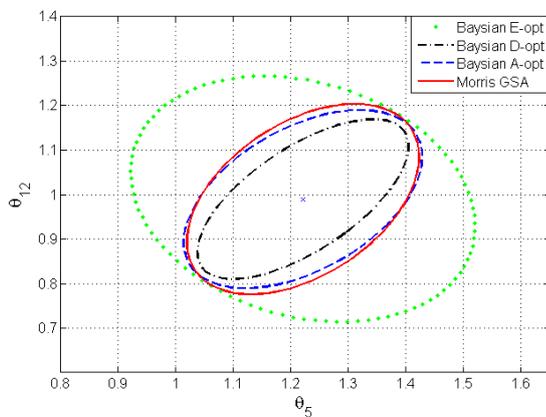
(b) p=10 in Morris GSA design

Fig. 1 The 95% confidence ellipse for Bayesian A-, D-, E-designs and Morris GSA based design using two parameters for illustration

Table I. State measurement design results for different methods: (1) Bayesian A-optimal, (2) Bayesian D-optimal, (3) Bayesian E-optimal, (4) Morris GSA design

Methods	Ranking of state measurement weights (descending order)									
(1)	**8**	**5**	**7**	**1**	10	4	9	3	2	6
(2)	**1**	**5**	**8**	**4**	10	7	3	9	2	6
(3)	**5**	**8**	**7**	**10**	1	4	3	2	9	6
(4)	**5**	**8**	**1**	**7**	4	10	9	3	2	6

Comparing the top four selected state variables of different designs in Table I, it can be observed that the Morris GSA design selects the same top four state measurements as the Bayesian A-optimal design, although in a slightly different order, and three out of four of the Morris design are the same as those of the Bayesian D- and E-optimal designs. Comparison of the 95% confidence ellipse in Fig. 1 also demonstrates the closeness of the Bayesian A-optimal design with the Morris GSA design. This result verifies the theorem in Section 3.2 and highlights the link between the GSA design and the Bayesian experimental design. The slight difference between the Morris design and the Bayesian A-optimal design is caused by numerical computation since it is impractical to take *p* to be very large. A lattice sampling without a sufficiently large number of partitions may result in a poor representation of the uncertainty region. As shown in Fig.1, the larger the lattice partition number is, the closer the result is to the sampling-based Bayesian experimental design.

5. CONCLUSIONS

Previous works on experimental design are mainly related to local sensitivity analysis via the formulation of FIM in the optimization. The possible contribution from GSA has rarely been discussed before. In this work, the explicit link between a lattice sampling based GSA, the Morris method, and the Bayesian A-optimal design is established based on a *prior* distribution information on the model uncertainties. This link suggests the potential use of GSA in improving robust experimental design with lattice sampling strategies.

REFERENCES

Asprey, S. P. and Macchietto, S. (2002). Designing robust optimal dynamic experiments. *Journal of Process Control,* 12(4)**,** 545-556.

Atkinson, A. C. and Donev, A. N. (1992). *Optimum Experimental Design.* Oxford University Press, Clarendon, Oxford.

Balsa-Canto, E., Alonso, A. A. and Banga, J. R. (2008). Computational procedures for optimal experimental design in biological systems. *IET Syst. Biol.,* 2(4)**,** 163-172.

Baltes, M., Schneider, R., Sturm, C. and Reuss, M. (1994). Optimal experimental design for parameter estimation in unstructured growth models. *Biotechnol. Prog. ,* 10(5)**,** 480-488.

Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: a review. *Statist. Sci.,* 10(3)**,** 273-304.

Cho, K.-H., Shin, S.-Y., Kolch, W. and Wolkenhauer, O. (2003). Experimental design in systems biology, based on parameter sensitivity analysis using a Monte Carlo method: a case study for the TNFa-Mediated NF-kB signal transduction pathway. *Simulation,* 79(12)**,** 726-739.

Clyde, M. A. (2001) Experimental design: a Bayesian perspective. *International Encyclopedia of the Social and Behavioral Sciences.* Elsevier Science.

Faller, D., Klingmuller, U. and Timmer, J. (2003). Simulation methods for optimal experimental design in system biology. *Simulation,* 79(12)**,** 717-725.

Franceschini, G. and Macchietto, S. (2008). Model-based design of experiments for parameter precision: State of the art. *Chemical Engineering Science,* 63(19)**,** 4846-4872.

Han, C. and Chaloner, K. (2004). Bayesian experimental design for nonlinear mixed-effects models with application to HIV dynamics. *Biometrics,* 60(1)**,** 25-33.

He, F., Brown, M. and Yue., H. (2010). Maximin and Bayesian robust experimental design for measurement set selection in modelling Biochemical regulatory systems. *Int. J. Robust and Nonlinear Contr.,* (in press).

Hoffmann, A., Levchenko, A., Scott, M. L. and Baltimore, D. (2002). The IκB-NF-κB signaling module: temporal control and selective gene activation. *Science,* 298**,** 1241-1245.

Iman, R. L., Helton, J. C. and Campbell, J. E. (1981). An approach to sensitivity analysis of computer models, Part 1. Introduction, input variable selection and preliminary variable assessment. *J. Quality Technology,* 13(3)**,** 174-183.

Körkel, S., Bauer, I., Bock, H. G. and Schlöder, J. P. (1999) A sequential approach for nonlinear optimum experimental design in DAE system. *Proceedings of the International Workshop on Scientific Computing in Chemical Engineering.* TU Hamburg-Harburg, Germany.

Körkel, S., Kostina, E., Bock, H. G. and Schlöder, J. (2004). Numerical methods for optimal control problems in design of robust optimal experiments for nonlinear dynamic processes. *Optimization Methods and Software,* 19(3-4)**,** 327-338.

Kutalik, Z., Cho, K.-H. and Wolkenhauer, O. (2004). Optimal sampling time selection for parameter estimation in dynamic pathway modeling. *BioSystems,* 75(1-3)**,** 43-55.

Lindley, D. V. (1972). *Bayesian Statistics - A Review.* Cambridge University Press, Cambridge, UK.

Lindley, D. V. and Smith, A. F. M. (1972). Bayes estimates for the linear model. *J. Royal Statistical Society, Series B,* 34(1)**,** 1-4.

Merlé, Y. and Mentré, F. (1995). Bayesian design criteria: Computation, comparison, and application to a pharmacokinetic and a pharmacodynamic model. *Journal of Pharmacokinetics and Biopharmaceutics,* 23(1)**,** 101-125.

Morris, M. D. (1991). Factorial Sampling Plans for Preliminary Computational Experiments *Technometrics,* 33**,** 161-174.

Rojas, C. R., S.Welsh, J., Goodwin, G. C. and Feuer, A. (2007). Robust optimal experiment design for system identification. *Automatica,* 43 993 - 1008.

Saltelli, A., Chan, K. and Scott, E. M. (Eds.) (2000) *Sensitivity Analysis,* John Wiley & Sons Ltd, West Sussex, England.

Silvey, S. D. (1980). *Optimal Design.* Chapman and Hall, London, UK.

Taguchi, G. and Yokoyama, Y. (1993). *Taguchi Methods: Design and Experiments.* ASI Press, Tokyo, Japan.

Thomaseth, K. and Cobelli, C. (1999). Generalized sensitivity functions in physiological system identification. *Ann Biomed Eng,* 27(5)**,** 607-616.

Yue, H., Brown, M., He, F., Jia, J. F. and Kell, D. B. (2008). Sensitivity analysis and robust experimental design of a signal transduction pathway system. *Int. J. Chem. Kinet.,* 40(11)**,** 730-741.

APPENDIX SIMULATION MODEL

This is a simplified IκBα-NF-κB computational model developed based on Hoffmann's model (Hoffmann et al., 2002) with IκBβ and IκBε knock out. Due to page limitation, the reactions are not listed here. Interested readers can check details from the original modelling paper.

Table A1. State definitions

Species	States	Species	States
IκBα	x_1	IKK	x_6
NF-κB	x_2	NF-κB$_n$	x_7
IκBα-NF-κB	x_3	IκBα$_n$	x_8
IKKIκBα	x_4	IκBα$_n$-NF-κB$_n$	x_9
IKK-IκBα-NF-κB	x_5	IκBα$_{-t}$	x_{10}

The model is described by a set of ODEs.

$$\dot{x}_1 = -(\theta_{17} + \theta_{18})x_1 + \theta_2 x_3 + \theta_{15}x_4 + \theta_{19}x_8 + \theta_{16}x_{10} - \theta_1 x_1 x_2 - \theta_{14}x_1 x_6$$

$$\dot{x}_2 = -\theta_7 x_2 + (\theta_2 + \theta_6)x_3 + (\theta_4 + \theta_5)x_5 + \theta_8 x_7 - \theta_1 x_1 x_2 - \theta_3 x_2 x_4$$

$$\dot{x}_3 = -(\theta_2 + \theta_6)x_3 + \theta_{21}x_5 + \theta_{22}x_9 + \theta_1 x_1 x_2 - \theta_{20}x_3 x_6$$

$$\dot{x}_4 = -(\theta_{15} + \theta_{24})x_4 + \theta_4 x_5 + \theta_{14}x_1 x_6 - \theta_3 x_2 x_4$$

$$\dot{x}_5 = -(\theta_4 + \theta_5 + \theta_{21})x_5 + \theta_3 x_2 x_4 + \theta_{20}x_3 x_6$$

$$\dot{x}_6 = (\theta_{15} + \theta_{24})x_4 + (\theta_5 + \theta_{21})x_5 - \theta_{23}x_6 - \theta_{14}x_1 x_6 - \theta_{20}x_3 x_6$$

$$\dot{x}_7 = \theta_7 x_2 - \theta_8 x_7 + \theta_{10}x_9 - \theta_9 x_7 x_8$$

$$\dot{x}_8 = \theta_{18}x_1 - \theta_{19}x_8 + \theta_{10}x_9 - \theta_9 x_7 x_8$$

$$\dot{x}_9 = -(\theta_{10} + \theta_{22})x_9 + \theta_9 x_7 x_8$$

$$\dot{x}_{10} = \theta_{11}S - \theta_{13}x_{10} + \theta_{12}x_7^2$$

Parameter values are listed in Table A2 with units of μM in concentration and minute in time.

Table A2 Parameter values in the IκBα-NF-κB model

θ_1	30	θ_9	30	θ_{17}	0.00678
θ_2	6e-5	θ_{10}	6e-5	θ_{18}	0.018
θ_3	30	θ_{11}	9.24e-5	θ_{19}	0.012
θ_4	6e-5	θ_{12}	0.99	θ_{20}	11.1
θ_5	1.221	θ_{13}	0.0168	θ_{21}	0.075
θ_6	6e-5	θ_{14}	1.35	θ_{22}	0.828
θ_7	5.4	θ_{15}	0.075	θ_{23}	0.0072
θ_8	0.0048	θ_{16}	0.2448	θ_{24}	0.2442