Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

A Bayesian Robust Observation Design Approach for Systems with (Large) Parametric Uncertainties

Hui Yu * **Hong Yue** ** **Xian Wei** * **Xiaoke Su** **

* *Fujian Institute of Research on the Structure of Matter, Chinese Academy of Science, China (e-mail: yuhui@fjirsm.ac.cn; xian.wei@fjirsm.ac.cn)*
** *Wind Energy and Control Centre, Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow G1 1XW, UK (e-mail: hong.yue@strath.ac.uk; xiaoke.su@strath.ac.uk)*

Abstract: Classical optimal experimental design (OED) methods have not been fully exploited in modeling of complex systems, due to the brittle design results generated based on prior models and computational burden in the optimization scheme. In this work, a novel method for robust experimental design (RED) of combined measurement set selection and sampling time scheduling has been proposed for systems with large parameter uncertainties. A Bayesian design framework is employed, involving Gaussian quadrature formula (GQF) approximation of the expected performance of the posterior distribution over uncertain parameter domain. The robust Bayesian experimental design (BED) has been relaxed to a semi-definite programming (SDP) problem which can be solved as a convex optimization problem. The proposed method has been examined by simulation studies on a lab-scale enzymatic biodiesel production system, with results compared to OED and uniform sampling under two design scenarios.

Keywords: Robust experimental design (RED), Bayesian experimental design (BED), semi-definite programming (SDP), Gaussian quadrature formula (GQF), observation strategy.

1. INTRODUCTION

The selection of optimal experimental conditions is crucial to maximize the information content in the collected data, especially for experiments that are costly and/or time consuming to conduct. Optimal experimental design (OED) techniques have been developed to systematically choose necessary experimental strategies that generate the most informative data, from which the parameters can be estimated with a higher accuracy. As a powerful tool assisting data-based modeling process, OED methods have been used in a wide range of applications (Vanlier et al., 2012; Murphy et al., 2004; Dette et al., 2005; He et al., 2008; Chen et al., 2011).

Typical OED tasks include stimulation signal design, sampling (time) profile optimization, measurement set selection, among others. Experimental design has been well studied in both theory and practice (Gil et al., 2014; Yue et al., 2013; Paquet-Durand et al., 2015). In a standard model-based framework, OEDs are usually derived using optimality criteria formulated from the expected Fisher information matrix (FIM). Many useful methods have been developed for classical experimental design to linear models. There is a lack of OED theory and methods for modeling of nonlinear systems using experimental data.

In engineering practice, nonlinear models are normally linearized around selected operating points with nominal parameters provided for each linearized model. Such a linearized model works well mainly in a nearby region of the operating point, inevitably having parametric uncertainties when the operation is in a wider range. In addition, real experiments contain measurement noise and may subject to unexpected process variations. These uncertainties affect the reliability of standard OED and, hence, require robust experimental design (RED) solutions.

Several information theoretic criteria have been proposed for Bayesian experimental design (BED) of systems with prior distribution of model parameters. Lindley suggests using expected gain in Shannon information, from prior to posterior, as a measure of the information provided by an experiment; the same objective can be justified from a decision theoretic perspective Lindley (1956). Another experimental design criterion is proposed to maximize Shannon entropy of the marginal distribution of the data, which can be treated as a special case of Lindley's criterion (Sebastiani and Wynn, 2000). A few other utility functions have been developed for experimental design of specific issues. Interested readers can refer to (Chaloner and Verdinelli, 1995) for a review of utility functions applied in decision theoretic approaches.

⋆ We acknowledge sponsorship of UK Engineering and Physical Sciences Research Council (EPSRC) through grant EP/R007497/1. This work was partially supported by CAS Pioneer Hundred Talents Program (Type C) under Grant No.2017-122, National Science Found for Young Scholars under Grant No. 61806186.

The development of BED depends on the prior model, the criterion used in design, and the numerical procedures required to solve the optimization problem. The design task is not trivial especially for models that contain a large number of design variables and model parameters (Lin and Runolfsson, 2012; Long et al., 2013; Huan and Marzouk, 2014). Most Bayesian optimal design methods in the literature have been restricted to low dimensional designs, i.e., simple models and a small number of design variables. For complex dynamic systems, the computational load involved in the optimization design becomes a main cause that hinders the practical application of Bayesian design methods. Development of efficient optimization approaches for BED is expected. In this work, robust observation design is addressed for models with large parametric uncertainties through a Bayesian design approach. A Bayesian D-optimal design is proposed to determine the observation strategy, including sampling time scheduling and measurement set selection. The multi-dimensional Gaussian quadrature formula (GQF) is applied to approximate the Bayesian integration over the parameter uncertainty space.

The remaining of the paper is organized as follows. Section 2 presents preliminaries on classic FIM-based OED and a global sensitivity analysis method. In Section 3, the BED framework is introduced, in which GQF is adopted to approximate the performance function integration required in a Bayesian design problem, and the combined observation design problem is relaxed to a semi-definite programming (SDP) problem. In Section 4, the proposed method is simulated using a lab-scale biodiesel production system, and the results are discussed for various sampling options. Conclusions are given in Section 5.

2. PRELIMINARIES ON OED AND GLOBAL SENSITIVITY ANALYSIS

Consider a general nonlinear dynamic model with n state variables, p parameters and m output variables, the state and output description can be written in the form of ordinary differential equations (ODEs) and algebraic equations as follows

$$\dot{\mathbf{X}}(t) = \mathbf{f}\left(\mathbf{X}(t), \boldsymbol{\theta}\right), \mathbf{X}(t_0) = \mathbf{X}_0 \tag{1}$$

$$\mathbf{Y}(t) = \mathbf{h}\left(\mathbf{X}(t), \boldsymbol{\theta}\right) + \boldsymbol{\xi}(t) \tag{2}$$

where $\mathbf{f}(\cdot)$ consists of a set of state transition functions of the system dynamics which are assumed to be continuous and first-order derivative; $\mathbf{X} = [x_1, x_2, \cdots, x_n]^T \in \mathbb{R}^n$ denotes the vector of n state variables with initial condition $\mathbf{X}_0$; $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_p]^T \in \mathbb{R}^p$ is the vector of p model parameters with uncertainties within the closed domain of $\boldsymbol{\Theta}$; $\mathbf{Y} = [y_1, y_2, \cdots, y_m] \in \mathbb{R}^m$ is the measurement output vector with $m (m \leq n)$ measurable variables, and $\mathbf{h}(\cdot)$ is the measurement function, normally used for selecting which variables to be measured. $\boldsymbol{\xi}$ is the vector of measurement errors which can be classified into systematic errors and random errors. An experiment should be designed to eliminate the systematic errors, however, the random errors that contaminate the observations always exist. In this work, the measurement errors are assumed to be zero mean, Gaussian noises.

2.1 Classic Optimal Experimental Design

Fisher information matrix (FIM) is widely used in model-based OED as the basis to quantify information content in data for parameter estimation. When the design factors that characterize the experiment are denoted as $\boldsymbol{\phi}$, the FIM can be expressed as (Ljung, 1998)

$$\mathbf{FIM}\left(\boldsymbol{\theta}, \boldsymbol{\phi}\right) = \mathbf{S}\left(\boldsymbol{\theta}, \boldsymbol{\phi}\right)^T \mathbf{W} \mathbf{S}\left(\boldsymbol{\theta}, \boldsymbol{\phi}\right) \tag{3}$$

where the weighting matrix, $\mathbf{W}$, quantifies the reliability of measurement data, which is normally taken as the inverse of the measurement error covariance matrix. $\mathbf{S} = \partial \mathbf{X} / \partial \boldsymbol{\theta}$ is the local parametric sensitivity matrix representing the local effects of parameters on model outputs. Following the Cramer-Rao inequality, the FIM is approximately equal to the inverse of the parameter estimation error covariance matrix, thus providing the local lower bound of the variance for parameter estimates. An OED can be expressed as the optimization of a proper measure of parameter error covariance matrix, i.e.,

$$\boldsymbol{\phi}^* = \underset{\boldsymbol{\phi} \in \boldsymbol{\Phi}}{\arg\min}\, v\left(\left(\mathbf{FIM}^{-1}\left(\boldsymbol{\theta}, \boldsymbol{\phi}\right)\right)\right) \tag{4}$$

where $\boldsymbol{\Phi}$ is the admissible space of the design factors, $v(\cdot)$ represents a scalar function of the inverse of FIM. Common design criteria in OED include A-, D- and E-optimality that scalarize different features of FIM. For example, when errors are independent and normally distributed, the D-optimality criterion minimizes the volume of the confidence ellipsoid of the model parameters. With A-optimality, the (squared) diagonal of the bounding box of the confidence ellipsoid is minimized. For E-optimality, the squared in-ball radius is minimized geometrically by maximizing the minimum eigenvalue of FIM.

2.2 Global Sensitivity Analysis

Global (parametric) sensitivity analysis (GSA) are used to analyze impacts of parameters on outputs of interest, when model parameters have large uncertainties or interactions among them. From the experimental design point of view, the main advantage of GSA over a local sensitivity analysis (LSA) method is that multiple parameters can be varied simultaneously, rather than individually, over their entire uncertainty ranges instead of the close region around the nominal values. GSA aims at apportioning the output uncertainty to the uncertainty in model parameters over a possible wide range.

In this work, Sobol's method (Sobol, 2001), a well accepted GSA method, is employed to identify key parameters that have large impacts on model outputs. For $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_p]^T \in \mathbb{R}^p$, the integrable output function, written as $g(\boldsymbol{\theta})$, can be decomposed into a summation of 2^p terms with increasing dimensionality:

$$\begin{aligned} g(\boldsymbol{\theta}) =& g_0 + \sum_{i=1}^{p} g_i(\theta_i) + \sum_{i=1}^{p} \sum_{j=i+1}^{p} g_{i,j}(\theta_i, \theta_j) + \cdots \\ &+ g_{1,2,\cdots,p}(\theta_1, \cdots, \theta_p) \end{aligned} \tag{5}$$

where $g_0 = \int_{\Omega^p} g(\boldsymbol{\theta}) d\boldsymbol{\theta}$, Ω^p is the p-dimensional hypercube space of model parameters. The total variance can then be determined as

$$V = \int_{\Omega^p} g^2(\boldsymbol{\theta}) d\boldsymbol{\theta} - g_0^2 \tag{6}$$

The partial variances, which are the components of the total variance decomposition, are computed from each of the terms in (5) as

$$V_{i_1,\cdots,i_k} = \int_{i_1} \cdots \int_{i_k} g_{i_1,\cdots,i_k}(\theta_{i_1},\cdots,\theta_{i_k}) d\theta_{i_1} \cdots d\theta_{i_k} \tag{7}$$

where $1 \leq i_1 \leq \cdots \leq i_k \leq p$. With the assumption that parameters are mutually orthogonal, the variance of outputs to parameters can be decomposed as

$$V = \sum_{i=1}^{p} V_i + \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} V_{i,j} + \cdots + V_{1,2,\cdots,p} \tag{8}$$

In this way, the variance contributions to the total output variance from individual parameters and the interaction among parameters can be determined. The variance contributions are characterized by the ratio of the partial variance to the total variance, also called the Sobol sensitivity indices (SI), written as follows:

First order SI: $S_i = \frac{V_i}{V}$

Second order SI: $S_{i,j} = \frac{V_{i,j}}{V}$

Total order SI: $S_{T_i} = S_i + \sum_{j \neq i} S_{i,j} + \cdots$

The first order index is a measure of the variance contribution from a individual parameter, θ_i, to the total output variance, which is taken as the main effect. The total order index, S_{T_i}, includes the result of the main effect from θ_i and from all its interactions with other parameters. Normally only interaction terms between two parameters are considered, higher-order interaction terms are ignored.

3. BAYESIAN OBSERVATION DESIGN

3.1 Bayesian Experimental Design

Bayesian experimental design methods have been developed for systems with parameters that can be characterized by probability distribution functions (Telen et al., 2013; Dette et al., 2007; Flaherty et al., 2006; Huan and Marzouk, 2013). The designs are developed based around prior distribution of parameter estimates other than on the chosen single-point values (nominal values), thus more information on parametric uncertainties are taken into account. In a Bayesian optimal design, a utility function, $U(\boldsymbol{\phi}, \boldsymbol{\theta}, \mathbf{Y})$, needs to be defined that describes the value of choosing the experimental design factors in $\boldsymbol{\phi}$, from the admissible design space $\boldsymbol{\Phi}$, yielding data $\mathbf{Y}$ for parameter estimation of $\boldsymbol{\theta}$. The utility function can be chosen as a function of the posterior distribution of model parameters, $p(\boldsymbol{\theta})$. In this case, the BED is developed to maximize the expected utility function $U(\boldsymbol{\phi}, \boldsymbol{\theta}, \mathbf{Y})$ over the uncertain parameter range, $\boldsymbol{\Theta}$, i.e.,

$$\begin{aligned} \boldsymbol{\phi}^* &= \arg\max_{\boldsymbol{\phi} \in \boldsymbol{\Phi}} \mathbb{E}_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \left(U(\boldsymbol{\phi}, \boldsymbol{\theta}, \mathbf{Y}) \right) \\ &= \arg\max_{\boldsymbol{\phi} \in \boldsymbol{\Phi}} \int_{\mathbf{Y}} \int_{\boldsymbol{\Theta}} U(\boldsymbol{\phi}, \boldsymbol{\theta}, \mathbf{Y}) p(\boldsymbol{\theta}, \mathbf{Y} | \boldsymbol{\phi}) d\boldsymbol{\theta} d\mathbf{Y} \end{aligned} \tag{9}$$

The optimization with utility function in (10) can be reduced to maximizing a D-optimal function under the Bayesian design (Chaloner and Verdinelli, 1995). The D-optimal BED can be formulated as

$$\boldsymbol{\phi}_D^* = \arg\max_{\boldsymbol{\phi} \in \boldsymbol{\Phi}} \int_{\boldsymbol{\Theta}} \left(\det \left(\mathbf{FIM}(\boldsymbol{\theta}, \boldsymbol{\phi}) \right) \right) p(\boldsymbol{\theta}) d\boldsymbol{\theta} \tag{10}$$

The optimization problem in (10) does not normally have a closed form solution, thus numerical approximations are required to solve the maximization problem that includes a multi-dimensional integration.

3.2 Integration Approximation with Gaussian Quadrature Formula

Gaussian quadrature formulas are a class of methods that take appropriate weights and nodes to numerically approximate the integration of a function $f(x)$ to a high degree of accuracy (Duarte and Wong, 2015). The basic formula for a one-dimensional integral over a compact interval $[a, b]$ can be expressed as follows:

$$\int_a^b \omega(x) f(x) dx = \sum_{i=1}^{\mu} \omega_{i,\mu} f(x_{i,\mu}) \tag{11}$$

where $\omega(x)$ is a weighting function, and μ is the number of designated nodes. The accuracy of the approximation depends on the number of nodes, their locations, and the selected weights at the designated nodes. The main advantage of GQF is that it requires much less function evaluations to calculate the integration in the Bayesian design problem, compared to those traditional approximation methods such as Markov chain Monte Carlo (MCMC) and importance sampling. Multi-dimensional regular domain based integrals can be determined by implementing one GQF in each dimension.

3.3 Semi-definite Programming for Bayesian Observation Design

RED of a combined observation strategy, including selected measurement set and sampling time profiles, is investigated in a Bayesian framework. Consider a general nonlinear model in (1), where there are m measurable state variables in output and the sampling profile of each measurable state can be independent from each other. Denoting the total sampling size required in RED as N_{sp}, these N_{sp} sampling points will be allocated among the m measurable state variables through the design. For each state variable, assume K measurement points are made available for selection; they are also called supporting points for design ϕ. If all states have the same number of supporting points which is common in practice, then the total number of points for overall sampling selection is $m * K$, out of which N_{sp} points will need to be selected by observation design.

GQF is employed to approximate the expectation of the performance index over the discretized parameter space $\boldsymbol{\Theta}$. Suppose every Legendre polynomial used to approximate the integral, in each dimension of $\boldsymbol{\Theta}$, has a degree of $L-1$, the total number of discretization points is then L^p

for p parameters, and they are obtained from combining roots of the $(L-1)$-th order Legendre polynomial across p dimensions in $\mathbf{\Theta}$. Each discretization point is a Cartesian product of the set containing GQF points taken from each dimension of $\mathbf{\Theta}$. To this end, the weight at the l-th point, β_l, is given by

$$\beta_l = \prod_{j=1}^{p} \omega_{l,j} \frac{\theta_j^U - \theta_j^L}{2}, l = 1, \cdots, L^p \tag{12}$$

where $\omega_{l,j}$ is the weights vector of the Legendre polynomials for the j-th parameter on the interval $[-1, 1]$; θ_j^U and θ_j^L are the upper and lower bounds of the j-th model parameter.

Now we can extend the SDP formulation to construct the BED. For a typical D-optimal design, the Bayesian formulation can be written as (Boyd and Vandenberghe, 2004)

$$\begin{aligned} &\arg\max_{\mathbf{\Phi}} \sum_{i}^{m} \sum_{k=1}^{K} \sum_{l=1}^{L^p} \det(\mathbf{FIM}(\phi_{i,k}, \theta_l)) p(\theta_l) \beta_l \quad (13) \\ &s.t. \;\; \phi_{i,k} \geq 0, \;\; k = 1, ..., K; \;\; i = 1, \ldots, m \\ &\sum_{i}^{m} \sum_{k=1}^{K} \phi_{i,k} = N_{sp} \end{aligned}$$

Here $\phi_{i,k}$ are the weighting coefficients for generated supporting points, for which binary values of 0 and 1 are taken to start with (1 means the point is selected and 0 not selected). This integer optimization problem is then recast to a continuous optimization problem by relaxing the weightings to a continuous value between [0, 1], which can be solved by convex optimization tools such as SeDuMi (Sturm, 1999).

4. CASE STUDY: BED FOR ENZYMATIC BIODIESEL PRODUCTION SYSTEM MODEL

4.1 System Settings and Global Sensitivity Analysis

A lab-scale enzymatic biodiesel production system model is used to examine the effectiveness of the proposed BED method. The kinetic reaction mechanism and the full model equations can be found in a previous work (Yu et al., 2015). The nominal values of the kinetic parameters and the initial condition of input variables are given in the Appendix. The lower and upper bounds for uncertainty ranges of model parameters are set to be 50% and 150% of their nominal values, respectively, and the parameters are assumed to follow uniform distribution within the bounded ranges.

For this system, five state variables, T, D, M, BD, and FFA, are measurable ($m = 5$). The Sobol's sensitivity analysis method is applied to identify key parameters that are most influential on these five states. The Latin hypercube sampling strategy is employed where 2,000 samples for model parameters are selected for the analysis. The first order sensitivity measures and the total sensitivity measures by considering five model outputs, $\{T, D, M, BD, FFA\}$, are shown in Fig. 1. It can be observed that, for state variable BD, $k_2, k_8, k_{-8}, k_9, k_{-9}$ are found to be more important than other parameters (note that in the total effect some parameters have negative effects on the output BD). Similar results can be observed for state FFA, where the most important parameters identified are k_8, k_{-8}, k_9 and k_{-9}, but k_2 has shown little effect on FFA. Not surprisingly, k_2 is the only parameter that has significant effect on state variables T, D and M (only the effect on state T is shown in Fig.1, similar results are obtained for the other two states D and M). This is because k_2 dominates the decomposition of state variables T, D to generate acyl enzyme complex, and it implies that the transesterification process is mainly decided by the starting reaction rate k_2. Therefore, from the Sobol's sensitivity analysis, we can find that the following five model parameters, $\{k_2, k_8, k_{-8}, k_9, k_{-9}\}$, are the most important parameters for the output of interests.

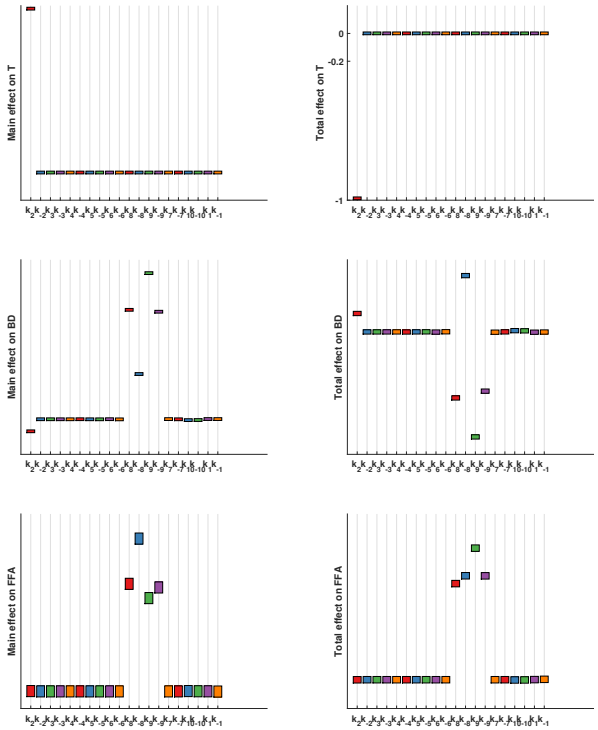

Fig. 1. Parameter effects to selected model outputs based on the analysis of variances, including the main effect and the total effect

All the simulation work are conducted in MATLAB with additional SeDuMi optimization package.

4.2 Bayesian Observation Design of Combined Sampling Time Scheduling and Measurement Set Selection

The five most important parameters identified from GSA in Section 4.1 are included for experimental design. The proposed Bayesain observation design is applied to find the best experimental strategy. The experimental length is set to be 25 hours (1,500 minutes). A sampling rate of 5 minutes is used to generate the data set, therefore the number of supporting points for experimental design in $K = 301$ for each state. In previous experimentation studies of this lab system, all five states take 28 measurements at the same time points equally spaced during the experimental process, which gives a total number of 140

sampling points. We refer this as non-designed strategy in the later comparison.

In this experimental design simulation, the target is to find a total number of 140 optimal samples ($N_{sp} = 140$) from the five measurable states along the experimental time horizon. We will consider two separate design scenarios.

(i) Same sampling time profiles for all states

In the first one, all five measurable state variables take the same sampling time profile, each state has 28 sampling points during the experimental length. The results from three methods are compared in Fig. 2, i.e., the non-designed strategy, the standard D-optimal design (OED) without considering parameter uncertainties, and the proposed Bayesian D-optimal design for uncertain system (BED). It can be observed from Fig. 2 that the sampling time schedule generated from BED is between the non-designed and local D-optimal design results. In both OED and BED, most sampling time points are taken in three regions where local parametric sensitivities are relatively high (the local sensitive analysis results are not included in this paper due to page limit). For BED, a few sampling points are also selected loosely among the parametric non-sensitive regions. This is the result from incorporating parameter uncertainties in the design. According to real lab experiences, equally spaced sampling strategy can intuitively provide useful data information. When the model contains large parameter uncertainties, the BED result is more close to the equally spaced sampling strategy rather than the OED result.

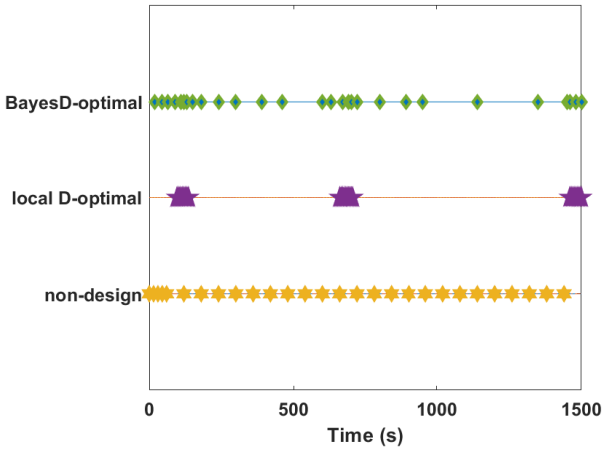

Fig. 2. Comparison of three different sampling strategies which include non-designed, local D-optimal and Bayesian D-optimal designs

(ii) Different sampling time profiles for measurable states

In the second scenario, the sampling of each state measurements is independent from sampling of other states. The total number of sampling points is still fixed to $N = 140$.

The Bayesian observation design result is shown in Fig. 3. It can bee seen that, among the five available state variables, only T, BD and FFA are selected to be included in the measurement set. The sampling time points for T are mostly selected between 0-600 minutes. This is consistent with the real process operation since the catalysed process

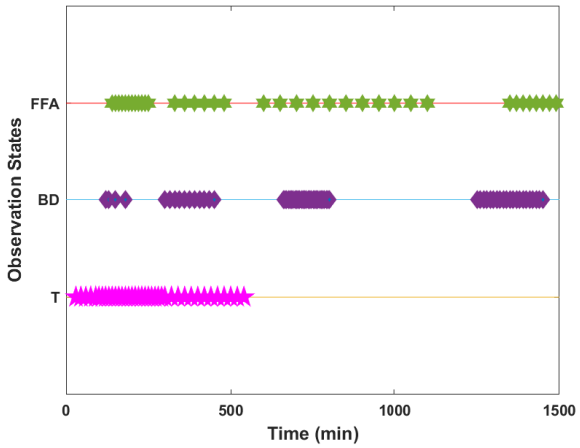

Fig. 3. Observation design results from Bayesian D-optimal design under the condition that measurement of each state is independent from each other

to generate acyl enzyme complex is more dominant in an early stage of reaction. The sampling time points for BD and FFA are selected covering a wide range of reaction process, again favouring regions of higher parametric sensitivities.

5. CONCLUSION

In this work, a BED strategy is proposed to find the optimal observation strategy when the initial model is subject to large parametric uncertainties. By taking the multidimensional GQF to approximate the expected performance of the posterior distribution over the bounded parameter domain, the Bayesian D-optimal design has been transformed into a SDP problem which can be solved as a convex optimization problem.

Using a lab-scale enzymatic biodiesel production system model, the Sobol's sensitivity analysis is implemented to determine crucial model parameters. The Sobol's method calculates the variance of the model outputs attributed to parameter variations. Considering uncertainty ranges of $\pm 50\%$ around the nominal values, the GSA result shows that five parameters, k_2, k_8, k_{-8}, k_9 and k_{-9}, have most overall influence on the measurable states. Taking these five crucial parameters into the design scheme, the proposed Bayesian observation design is implemented in order to find the optimal measurement set and sampling time schedule(s) at the same time. The Bayesian D-optimal design provides samplings that are taken mostly in parametric sensitive regions without completely ignoring other regions due to existence of large parameter uncertainties. This is a more balanced view in comparison to uniform sampling (non-designed) and standard D-optimal design (OED) without considering modeling uncertainties at all.

The computational cost of Bayesian RED is much higher than OED due to the integration of performance index over the uncertain parameter space. Further investigations on BED are still required to explore efficient numerical procedures. REDs that include multiple experimental factors will impose more complexity in design. One challenging task is to combine observation design and stimulation

design into a united optimization scheme. Most current experimental design studies are focused to improve modeling qualities; expanding this to include financial cost factors should make the results more appealing in applications.

REFERENCES

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press.

Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: A review. *Stat. Sci.*, 273–304.

Chen, B.S., Chang, C.H., Wang, Y.C., Wu, C.H., and Lee, H.C. (2011). Robust model matching design methodology for a stochastic synthetic gene network. *Mathematical Biosciences*, 230(1), 23–36.

Dette, H., Haines, L.M., and Imhof, L.A. (2007). Maximin and bayesian optimal designs for regression models. *Statistica Sinica*, 463–480.

Dette, H., Melas, V.B., Pepelyshev, A., and Strigul, N. (2005). Robust and efficient design of experiments for the monod model. *J. Theor. Biol.*, 234(4), 537–550.

Duarte, B.P.M. and Wong, W.K. (2015). Finding bayesian optimal designs for nonlinear models: A semidefinite programming-based approach. *Int. Stat. Rev.*, 83(2), 239–262.

Flaherty, P., Arkin, A., and Jordan, M.I. (2006). Robust design of biological experiments. In *Advances in Neural Information Processing Systems*, 363–370.

Gil, M.M., Miller, F.A., Silva, C.L., and Brandão, T.R. (2014). Application of optimal experimental design concept to improve the estimation of model parameters in microbial thermal inactivation kinetics. *J. Food Eng.*, 134, 59–66.

He, F., Brown, M., Yue, H., and Yeung, L.F. (2008). Robust experimental design and feature selection in signal transduction pathway modeling. In *Neural Networks, 2008. IJCNN 2008.(IEEE World Congress on Computational Intelligence). IEEE Int. Joint Conf.*, 1544–1551. IEEE.

Huan, X. and Marzouk, Y. (2014). Gradient-based stochastic optimization methods in bayesian experimental design. *Int. J. Uncertain Quantif.*, 4(6), 479–510.

Huan, X. and Marzouk, Y.M. (2013). Simulation-based optimal bayesian experimental design for nonlinear systems. *J. Comput. Phys.*, 232(1), 288–317.

Lin, C. and Runolfsson, T. (2012). Efficient optimal design of uncertain discrete time dynamical systems. *Automatica*, 48(10), 2544–2549.

Lindley, D. (1956). On a measure of the information provided by an experiment. *Ann. Math. Stat.*, 27(4), 986–1005.

Ljung, L. (1998). System identification. In *Signal Analysis and Prediction*, 163–173. Springer.

Long, Q., Scavino, M., Tempone, R., and Wang, S. (2013). Fast estimation of expected information gains for bayesian experimental designs based on laplace approximations. *Comput. Methods Appl. Mech. Eng.*, 259, 24–39.

Murphy, E.F., Gilmour, S.G., and Crabbe, M.J.C. (2004). Efficient and cost-effective experimental determination of kinetic constants and data: the success of a bayesian systematic approach to drug transport, receptor binding, continuous culture and cell transport kinetics. *FEBS Letters*, 556(1-3), 193–198.

Paquet-Durand, O., Zettel, V., and Hitzmann, B. (2015). Optimal experimental design for parameter estimation of the peleg model. *Chemom. Intell. Lab. Syst.*, 140, 36–42.

Sebastiani, P. and Wynn, H.P. (2000). Maximum entropy sampling and optimal bayesian experimental design. *J. R. Stat. Soc. Series B Stat. Methodol.*, 62(1), 145–157.

Sobol, I.M. (2001). Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Math. Comput. Simul.*, 55(1), 271–280.

Sturm, J.F. (1999). Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones. *Optim. Methods Softw.*, 11(1-4), 625–653.

Telen, D., Houska, B., Logist, F., Diehl, M., and Van Impe, J. (2013). Guaranteed robust optimal experiment design for nonlinear dynamic systems. In *2013 European Control Conference (ECC)*, 2939–2944. IEEE.

Vanlier, J., Tiemann, C.A., Hilbers, P.A., and van Riel, N.A. (2012). A bayesian approach to targeted experiment design. *Bioinformatics*, 28(8), 1136–1142.

Yu, H., Yue, H., and Halling, P. (2015). Optimal experimental design for an enzymatic biodiesel production system. *IFAC-PapersOnLine*, 48(8), 1258–1263.

Yue, H., Halling, P., and Yu, H. (2013). Model development and optimal experimental design of a kinetically controlled synthesis system. *IFAC Proceedings Volumes*, 46(31), 327–332.

Appendix A. NOMINAL PARAMETER VALUES AND INITIAL CONDITION FOR THE ENZYMATIC BIODIESEL PRODUCTION SYSTEM

Table A.1. Nominal parameter values for enzyme biodiesel production system

k_1	4.95e4	k_6	9.13e4
k_{-1}	6.60	k_{-6}	5.43e5
k_2	1.69e6	k_7	7.06e6
k_{-2}	1.11e4	k_{-7}	4.93
k_3	2.07e4	k_8	2.36e4
k_{-3}	2.20e7	k_{-8}	3.51e6
k_4	3.41e6	k_9	2.54e4
k_{-4}	1.33e7	k_{-9}	2.05e5
k_5	1.55e7	k_{10}	3.23e-2
k_{-5}	1.81e5	k_{-10}	4.39e-4

Table A.2. Initial input values and feeding rate of methanol

Species	Ini. cond. $(mol \cdot L^{-1})$	Species	Ini. cond. $(mol \cdot L^{-1})$
$T(x_1)$	0.95	$EX(x_{10})$	0
$D(x_2)$	0.02	$ET(x_{11})$	0
$M(x_3)$	1.4e-3	$ED(x_{12})$	0
$B(x_4)$	1e-4	$EM(x_{13})$	0
$FFA(x_5)$	2.24e-2	$ECH(x_{14})$	0
$G(x_6)$	1e-6	$Ef(x_{15})$	9.72e-6
$W(x_7)$	2.39	$Vp(x_{16})$	6.61e-2
$CH(x_8)$	0.59	$V(x_{17})$	1.54
$E(x_9)$	0		
Methanol feed rate [eq $\cdot h^{-1}$]	Initial dose methanol [eq]	water [wt.% oil]	Enzyme [wt.% oil]
0.185 first 2hrs; 0.06 thereafter	0.2	5	0.5