

Integrated Time Sampling Design and Measurement Set Selection for Biochemical Systems

Hui Yu, Hening Yu, Hong Yue

Department of Electronic and Electrical Engineering
University of Strathclyde
Glasgow, G1 1XW, UK

Email: hui.yu.100@strath.ac.uk; hening.yu@uni.strath.ac.uk;
hong.yue@strath.ac.uk

Jinglin Zhou

College of Information Science and Technology
Beijing University of Chemical Technology
Beijing 100029, China
jinglinzhou@mail.buct.edu.cn

Abstract—The optimal experimental design (OED) for observation strategy is investigated in this paper to collect the most informative experimental data for parameter estimation. The aim is to determine the best sampling time points and also select the most valuable measurement state variables through OED. The two design objectives are integrated together as a single-objective optimisation problem in which the variables and their sampling times are weighted in an expanded time sampling framework. Three optimisation methods, i.e., the Powell’s method, the sequential selection method, and the sequential quadratic programming method, are employed to solve the optimisation problem. Their computation efficiencies are compared using a biodiesel case study system simulation. Simulation results demonstrate the effectiveness of the proposed method in reducing parameter estimation uncertainties as well as reducing parameter correlations. It can also be observed that the integrated OED doesn’t cost extra computation efforts when variable selection is considered together with the time sampling task.

I. INTRODUCTION

Mathematical modelling is an essential tool to help understand the behaviour of biochemical and biological processes. The capability of a model to describe a system largely depends on model parameters, the values of which need to be estimated by comparing model prediction to real experimental data [1]. However, due to the complexity in system dynamics and the limitations in experiment techniques, the measurement data is usually sparse and heavily contaminated with noise. This apparently hinders parameter estimation precision and adversely affects further analysis of the established models. Therefore, it is important to design necessary experiments so that the most informative data can be obtained in order to facilitate parameter estimation. The optimal experimental design (OED) technique is one of the most widely used methods during system identification [2]. The main purpose of OED is to find the optimal experimental condition based on the optimisation of scalar measure of Fisher information matrix (FIM), which will lead to the most informative experimental data that can reduce parameter estimation errors.

OED tasks generally include the following: input design, sampling design and measurement set selection. The input design is to decide the type, level and duration of input signals [3, 4, 5]; the sampling design is to determine when and how the samples should be collected [6, 7]; and the measurement set

design is to find the most valuable measurement state variables [8, 9, 10].

In this work, the observation design of a biodiesel production system is investigated because this system includes series of parallel, consecutive, and competitive reversible reactions. It contains considerable dynamic complexity and represents a typical class of biochemical processes. In previous work, researchers have mentioned the difficulty of determining distinctive parameter estimation values due to high correlations between parameter pairs caused by parallel reactions [11]. An OED method has been proposed aiming to reduce parameter correlations [12]. However, this might not effectively reduce the parameter estimation uncertainties. Our work will mainly focus on reduction of parameter uncertainties through the best observation strategy. Here the observation strategy design includes both design targets on finding the best sampling time points and determining the most useful measurement variables among the available set. In addition, we will also examine the effect of our method on reducing parameter correlations. The rest of this paper is organized as follows. In Section 2, relevant preliminaries on OED are briefly introduced. The observation design formulation and the numerical optimization methods are described in detail in Section 3. Then the optimal observation design for a biodiesel production system and the analysis of simulation results are presented in Section 4. Finally, the conclusions and future work are discussed in Section 5.

II. PRELIMINARIES

A. Local Sensitivity Analysis

In order to analyse the underlying dynamic behaviour of biochemical systems, mathematical modelling is a basic and essential method to describe dynamic systems. A general model of biochemical system can be developed based on energy and mass balance laws, as a set of ordinary differential equations

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

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

where $\mathbf{f}(\cdot)$ is a set of nonlinear functions of states transition which refer to the reaction mechanisms. $\mathbf{h}(\cdot)$ is the vector

of measurement function. $\mathbf{X} = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$ is the vector of n state variables with initial condition \mathbf{X}_0 . In biochemical reactions, the components in \mathbf{X} denoted as x_i normally represent the concentration of reactants. $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_p]^T \in \mathbb{R}^p$ represents the vector of p model parameters, the components of which denote kinetic reaction rates. $\mathbf{Y} \in \mathbb{R}^m$ stands for the measurement output vector with m ($m \leq n$) measurable variables. $\boldsymbol{\xi}$ is the measurement noise vector which is assumed to follow a zero-mean Gaussian distribution in this work.

Sensitivity analysis is used to investigate how model outputs depend on model parameters and initial conditions. The local parametric sensitivity matrix is defined as $\mathbf{S} = \frac{\partial \mathbf{X}}{\partial \boldsymbol{\theta}}$, and its dynamics can be expressed as $\dot{\mathbf{S}} = \mathbf{J}\mathbf{S} + \mathbf{F}$, where $\mathbf{J} = \frac{\partial \mathbf{f}}{\partial \mathbf{X}}$ is the Jacobian matrix and $\mathbf{F} = \frac{\partial \mathbf{f}}{\partial \boldsymbol{\theta}}$ is the parameter Jacobian matrix. For most biochemical systems with several model outputs and a large number of model parameters, there are usually large differences of sensitivity values among parameters across outputs. Performing sensitivity analysis can help to find which parameters have large impacts on output variables, and thus the parameter estimation only needs to be focused on those selected key parameters. This will not only reduce the computational load but will also increase the identifiability level.

B. Fisher Information Matrix and Optimal Experimental Design Criteria

In OED, FIM is used to measure the amount of data information contained in the experimental data. It can be built based on parametric sensitivity matrix and measurement error covariance matrix as follows

$$\mathbf{FIM} = \mathbf{S}^T \mathbf{Q}^{-1} \mathbf{S} \quad (3)$$

where \mathbf{Q} represents measurement error covariance matrix. The Cramer-Rao bound states that the inverse of the FIM provides a lower bound of parameter estimation variance-covariance matrix which is the fundamental support for FIM-based OED [13].

The target of OED is to optimize a scalar measure of FIM using different criteria. The objective function of OED can be expressed as

$$\boldsymbol{\omega}^* = \arg \min_{\boldsymbol{\omega} \in \Omega} \Phi \left((\mathbf{FIM}(\boldsymbol{\theta}, \boldsymbol{\omega}))^{-1} \right) \quad (4)$$

where $\boldsymbol{\omega}$ is a vector representing design factors and Ω is the possible space for $\boldsymbol{\omega}$. \mathbf{FIM} is characterized by both the model parameters $\boldsymbol{\theta}$ and the design factors $\boldsymbol{\omega}$. $\Phi(\cdot)$ denotes the design criteria such as A-, D- and E- optimal design that will get scalar features from FIM [14]. The choice of design criterion is case dependent. In this work, the D-optimal and E-optimal design will be compared in the simulation in order to find the best design criterion for this biodiesel production model.

III. OPTIMAL OBSERVATION DESIGN METHOD

A. Problem Definition

An OED problem can be formulated as the optimization of certain measure of FIM as represented in equation (4). The

design factors may include input levels, initial conditions and observation strategies. Our interest in this work is to design the best observation strategy. Observation design is to determine the best sampling time strategy and the most valuable measurement state variables that can provide the most informative experimental data for parameter estimation. For continuous time dynamic systems, the time sampling design problem is hard to solve since it is an infinite dimensional non-convex dynamic optimization problem. Thus, the observation design problem is generally transformed into discrete optimization problem by adding weighting factors to all available sampling points for each measurable state variable.

$$\boldsymbol{\xi} = \begin{Bmatrix} t_1 & \cdots & t_{N \times n} \\ \omega_1 & \cdots & \omega_{N \times n} \end{Bmatrix}$$

$$\boldsymbol{\xi}^* = \arg \min_{\boldsymbol{\omega} \in \Omega} \Phi \left(\left(\sigma^2 \sum_{l=1}^{N \times n} \omega_l \mathbf{S}(t_l)^T \mathbf{S}(t_l) \right)^{-1} \right) \quad (5)$$

$$s.t. \quad \omega_l \in \{0, 1\}, \mathbf{1}^T \boldsymbol{\omega} = N_{sp}$$

where N is the total number of available sampling time points of each state variable. With this formulation, when the weighting factor ω_l takes value of 1, it means the related sampling time point should be selected; those sampling time points with weighting factors of value of 0 will be ignored during the measurement. Therefore, the optimization problem can be considered as determining the best schedule of weighting factors for all available sampling points across all variables. The observation design problem can thus be transferred into an integer optimization problem. When a biochemical system contains large number of available sampling time points, equation (5) is not easy to solve. An easy way is to relax weighting factors in equation (5) to continuous variables and the relaxed optimization problem can be described as:

$$\boldsymbol{\xi}^* = \arg \min_{\boldsymbol{\omega} \in \Omega} \Phi \left(\left(\sigma^2 \sum_{l=1}^{N \times n} \omega_l \mathbf{S}(t_l)^T \mathbf{S}(t_l) \right)^{-1} \right) \quad (6)$$

$$s.t. \quad \omega_l \in [0, 1], \sum_{l=1}^{N \times n} \omega_l = 1, \forall l$$

The solution of equation (6) will give all available sampling points small weighting values between [0,1] rather than selecting required number of sampling points which is what (5) does. Within this solution, larger weighting factors indicate the corresponding sampling time points are relatively more important which should be selected in measurement and sampling time points with small weightings will be ignored.

B. Numerical Optimization Methods

Different optimization methods have been attempted to solve the above optimal observation design problem.

- Powells method

This method is to try different combinations of sampling points to find the best sampling strategy which is used to solve problem (5) Firstly, the required number of sampling points are randomly selected. Then, in each step, one sampling point will be removed and replaced by a non-selected measurement point. This new combination will be tested and the measurement point which leads to the optimal objective value will be selected. This process will be iterated through all selected sampling points until no improvement can be found through the substitution of non-selected sampling points.

- Sequential selection method

The basic idea of this method is to select the best sampling points in a sequential manner until the required number is achieved. This method is also suited for the integer optimization problem (5). Firstly, select the measurement point which contains the most data information about key parameters. Then, select the next measurement point which can make the best objective value with the combination of the selected points. Repeat the step until the required number of samples is selected.

- Sequential quadratic programming (SQP) method

This method is used to solve the continuous optimization problem (6). The MATLAB tool *fmincon* with sequential quadratic programming algorithm is applied here to solve this problem. The solution of (6) can provide a lower bound of the solution of (5). The optimal solution gives every sampling points a small weighting to show their importance. Based on the required number of sampling time points N_{sp} , the time sampling points with the top N_{sp} largest weighting values will be selected as the measurement points.

IV. OBSERVATION DESIGN OF BIODIESEL PRODUCTION SYSTEM

A. Parameter Significance Analysis

The proposed observation design is investigated using a biodiesel production model proposed by Nouredini and Zhu [15]. A brief introduction of the reaction mechanism and mathematical model is given in the Appendix. This model is chosen for simulation study mainly because there are OED discussions on the same model in a previous work [16]. There are 6 reactants and 6 parameters in this non-linear model.

In order to make the comparison properly, the initial conditions and the experiment set-up used in this work are chosen to be the same as in the previous work [16]. The initial concentrations of the reactants, the kinetic parameter values, the time-varying input density and the feeding rate of the time varying input are listed in Table IV and V in the Appendix.

Fig. 1. shows the time profile of the 6 state variables (concentrations of reactants).

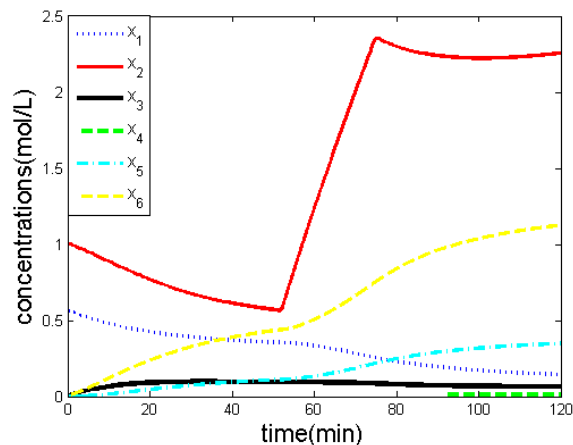


Fig. 1. Time profile of output state variables

It can be seen that, during the reaction process, the amount of oil (x_1) continuously decreases while that of ester (x_6) is increased. The concentration of Methanol (x_2) changes significantly at 52 minute and 75 minute, respectively. Then, sensitivity analysis is applied to identify the most important parameters. In this system, ester is the main product and the time profile of parametric sensitivities for x_6 is shown in Fig. 2.

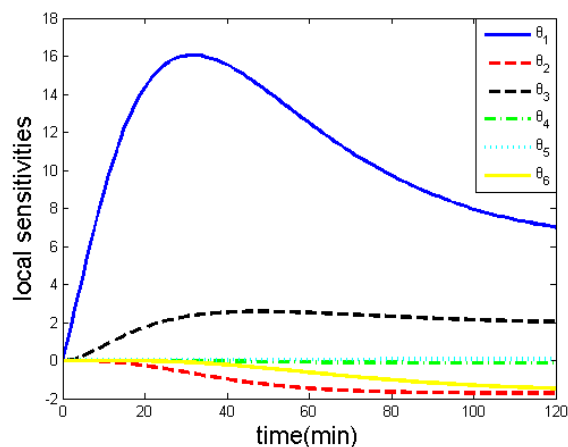


Fig. 2. Local parameter sensitivities of state variable x_6

B. Optimal Observation Design

Three parameters, θ_1 , θ_3 and θ_2 , have been identified as crucial parameters through sensitivity analysis, the Fisher information matrix can then be built based on those key parameters for all available sampling points applied to all 6 state variables. Then the OED problem can be formulated as the choice of optimal sampling points from the available measurement data. Three different numerical optimization methods introduced in Section 3 have been implemented to this OED problem. The problem of how to select a proper

design criterion for this model is considered first. Table I provides parameter pair correlations under nominal condition where some parameters are highly correlated. Thus the A-optimal design is not suitable for this case as it can only handle OED problem with relatively low parameter correlations. The modified E-optimal is mainly focused on the reduction of parameter correlation which might not efficiently improve parameter estimation precision. Therefore, D-optimal and E-optimal design criteria are considered in this work. The confidence intervals of parameter pair $[\theta_1, \theta_2]$ with two different design criteria by using Powell's method for sampling design are shown in Fig. 3.

TABLE I
PARAMETER PAIR CORRELATIONS UNDER NON-DESIGNED CONDITION

parameters	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6
θ_1	1	0.6835	-0.4679	-0.3145	-0.0691	0.1074
θ_2		1	-0.3170	-0.0825	0.0271	0.2457
θ_3			1	0.8043	0.1037	0.0558
θ_4				1	0.3000	0.0800
θ_5					1	0.8857
θ_6						1

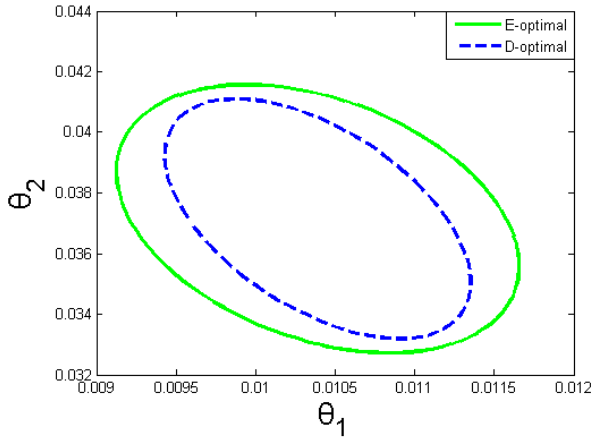


Fig. 3. Comparison of D-optimal and E-optimal

It can be observed that for the considered parameter pairs, when using the Powell's method to find the optimal solution, the D-optimal design provides smaller confidence intervals over the E-optimal design. Consistent results can be found by using the other two numerical optimization methods. The D-optimal results for sampling time design with different numerical optimization methods are compared in Fig. 4. The simulation results from the three optimization methods are consistent in general. The simulation result based on Powell's method and SQP method is slightly better than the result via sequential selection method.

Next the observation design to this biodiesel production system is investigated using D-optimal design and Powell's method. In most biochemical experiments, the equivalent sampling strategy is applied to all measurement variables. In

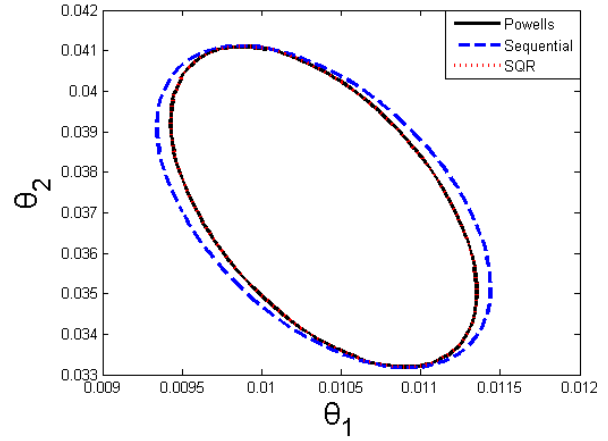


Fig. 4. Comparison of different numerical optimization methods

this work, we consider a more general situation that the time sampling selection of different state variables are independent to each other. This means the measurement data for different output state variables can be collected at different sampling points in time horizon. The observation design results is provided in Table II, from which it can be seen that only three state variables, x_1 , x_3 and x_5 contain selected sampling points, the other 3 state variables are not considered for sampling according to the result of the OED. For comparison purpose, the sampling strategy without OED design and the OED strategy from a previous work [16] are also provided in Table II in the last row, where in both cases all the 6 state variables are sampled at the same time points as listed in the table. The confidence intervals for parameter pair (θ_1, θ_3) under the proposed OED are compared with the non-designed result as shown in Fig. 5. It can be observed that, compared with the non-design result, the proposed observation design clearly improves the parameter estimation precision. The sampling strategy results also implies that the state variables x_1 , x_3 and x_5 are the most valuable measurement state variables as all measurement points are selected from these three state variables.

TABLE II
OPTIMAL SAMPLING STRATEGY FROM OBSERVATION DESIGN

state variables	sampling time points (unit: minute)
x_1	119.1, 119.2, 119.3, 119.4, 119.5, 119.6, 119.7, 119.8, 119.9, 120
x_3	4.1, 5.6, 5.8, 5.9, 6.2, 6.3, 6.4, 6.5, 6.6, 7.3, 7.4, 7.5, 51.3, 51.4, 51.5, 51.6, 51.7, 51.8, 51.9, 52
x_5	7.4, 7.5, 7.6, 7.7, 7.8, 7.9, 8, 8.1, 8.2, 8.3, 8.4, 8.5
non-design	5, 30, 60, 80, 90, 100, 110
OED in [17]	3, 7.4, 9.5, 53.8, 55.9, 118, 120

We then compare the proposed design based on D-optimal using Powell's method with the non-designed condition and also the result from a reference work [16], which is shown in Fig. 6. Among the three observation strategies, the proposed

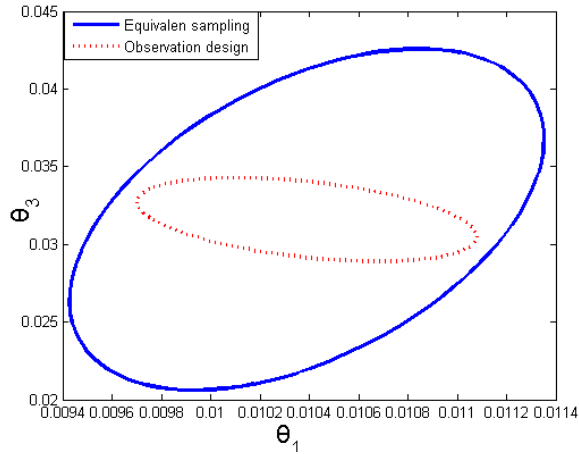


Fig. 5. Comparison of Independent sampling and traditional sampling strategy

method provides the minimum possible parameter estimation uncertainty. This is mainly because the proposed design combines both the time sampling design and the measurement set selection together while in other works the sampling design is assumed to apply to all variables in the same way.

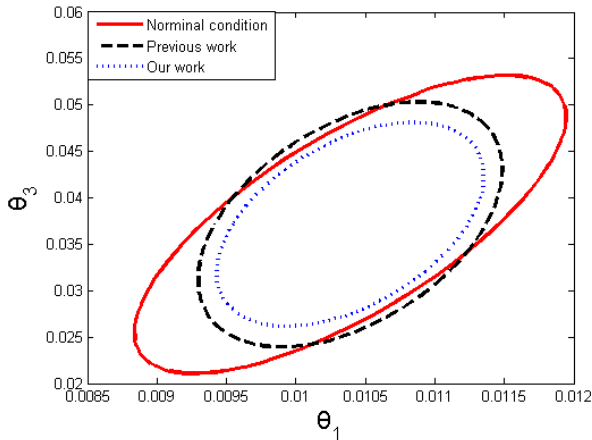


Fig. 6. Comparison of our work and other researcher's simulation result

It should be noted that correlation between parameter pairs is another adverse factor that would reduce parameter identifiability since the change in one parameter will result in proportional changes in another one or more parameters [17]. It is therefore useful to reduce correlations between parameters to be estimated if possible. We next examine the impact of OED to parameter correlations. The result without OED is shown in Table I where the three largest parameter pair correlation coefficients are highlighted with underlines. It has been mentioned in previous work [16] that the upper bound of parameter pair correlation coefficient is 0.75 for this biodiesel system. In the non-designed case, two of these coefficients exceed this set up bound and one of them nearly reaches 0.9

which is not acceptable. The parameter pair correlation matrix from the proposed observation design strategy is given in Table III. The three largest coefficients are significantly reduced and all the correlation coefficients are under the bound limit.

TABLE III
PARAMETER PAIR CORRELATIONS UNDER OBSERVATION DESIGN
CONDITION

parameters	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6
θ_1	1	<u>0.4333</u>	-0.5833	-0.3245	0.0299	0.0306
θ_2		1	-0.2172	-0.0875	0.0287	0.3171
θ_3			1	<u>0.6220</u>	0.1841	0.1346
θ_4				1	0.1084	0.0452
θ_5					1	<u>0.7171</u>
θ_6						1

V. CONCLUSION

In this work, a new observation design strategy which combines both sampling time design and measurement set selection is proposed. The two design problems are integrated into one single optimization problem which is convex and thus can be easily solved by standard optimization algorithms. The observation design method has been applied to a biodiesel production system for the reduction of parameter estimation uncertainties. Three different numerical methods are used to solve this OED problem and all of them can generally make consistent results in reducing parameter estimation uncertainties. It can also be observed from the simulation study that the correlations between parameter pairs can be reduced through collected data with OED. This will increase parameter identifiability and therefore further improve the overall parameter estimation quality.

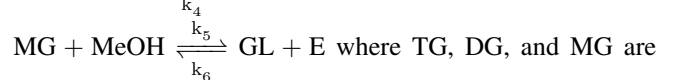
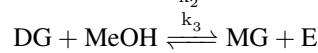
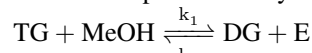
The effectiveness of the proposed OED strategy and its computational efficiency has been validated by the case study biodiesel system. Further work will be conducted to expand the integrated observation design to wider systems with higher complexities. The robustness of this OED method by considering parameter uncertainties and non-Gaussian measurement noise is another interesting topic which will be considered in our future work.

ACKNOWLEDGMENT

This work is partially supported by Chinese NSFC project no. 61473025.

APPENDIX

The following three reversible reactions are considered in the bio-diesel production system:



where TG, DG, and MG are the tri-, di-, and monoglycerides. MeOH, GL, and E indicate methanol, glycerol, and the mixture of methyl esters, which form biodiesel. All the reactions follow second order kinetics.

By defining x_i as variables representing species concentrations, a set of ordinary differential equations can be written to describe the concentration variations of the 6 components. Here the subscript $i = 1..6$ corresponds to triglycerides, methanol, diglycerides, monoglycerides, glycerol, and ester, respectively.

$$\begin{aligned} \frac{dx_1}{dt} &= -\theta_1 \frac{x_1 \cdot x_2}{vol} + \theta_2 \frac{x_3 \cdot x_6}{vol} \\ \frac{dx_2}{dt} &= u - \theta_1 \frac{x_1 \cdot x_2}{vol} + \theta_2 \frac{x_3 \cdot x_6}{vol} - \theta_3 \frac{x_2 \cdot x_3}{vol} \\ &\quad + \theta_4 \frac{x_4 \cdot x_6}{vol} - \theta_5 \frac{x_2 \cdot x_4}{vol} + \theta_6 \frac{x_5 \cdot x_6}{vol} \\ \frac{dx_3}{dt} &= \theta_1 \frac{x_1 \cdot x_2}{vol} - \theta_2 \frac{x_3 \cdot x_6}{vol} - \theta_3 \frac{x_2 \cdot x_3}{vol} + \theta_4 \frac{x_4 \cdot x_6}{vol} \\ \frac{dx_4}{dt} &= \theta_3 \frac{x_2 \cdot x_3}{vol} - \theta_4 \frac{x_4 \cdot x_6}{vol} - \theta_5 \frac{x_2 \cdot x_4}{vol} + \theta_6 \frac{x_5 \cdot x_6}{vol} \\ \frac{dx_5}{dt} &= \theta_5 \frac{x_2 \cdot x_4}{vol} - \theta_6 \frac{x_5 \cdot x_6}{vol} \\ \frac{dx_6}{dt} &= \theta_1 \frac{x_1 \cdot x_2}{vol} - \theta_2 \frac{x_3 \cdot x_6}{vol} + \theta_3 \frac{x_2 \cdot x_3}{vol} \\ &\quad - \theta_4 \frac{x_4 \cdot x_6}{vol} + \theta_5 \frac{x_2 \cdot x_4}{vol} - \theta_6 \frac{x_5 \cdot x_6}{vol} \\ vol &= \frac{x_1}{\rho_1} + \frac{x_2}{\rho_2} + \frac{x_3}{\rho_3} + \frac{x_4}{\rho_4} + \frac{x_5}{\rho_5} + \frac{x_6}{\rho_6} \end{aligned}$$

TABLE IV
NECESSARY CONDITION FOR THE MODEL

state variables (mol/L)	x_1 0.5656	x_2 1.0034	x_3 0	x_4 0	x_5 0	x_6 0
Parameters (1/(mol · min))	θ_1 0.01039	θ_2 0.03715	θ_3 0.03161	θ_4 0.398	θ_5 0.7413	θ_6 0.03548
densities (mol/L)	ρ_1 1.0233	ρ_2 24.5246	ρ_3 1.4732	ρ_4 2.6463	ρ_5 13.639	ρ_6 2.9250

TABLE V
FEEDING RATE OF TIME-VARYING INPUT SIGNALS

Time interval (min)	0-52	52-75	75-120
Feeding rate (mol/min)	0	0.0919	0.0061

REFERENCES

- [1] G. Franceschini and S. Macchietto, "Model-based design of experiments for parameter precision: State of the art," *Chem. Eng. Sci.*, vol. 63, no. 19, pp. 4846–4872, 2008.
- [2] A. C. Atkinson, "The usefulness of optimum experimental designs," *J. Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, pp. 59–76, 1996.
- [3] S. Asprey and S. Macchietto, "Designing robust optimal dynamic experiments," *J. Proc. Control*, vol. 12, no. 4, pp. 545–556, 2002.
- [4] J. R. Banga, K. J. Versyck, and J. F. Van Impe, "Computation of optimal identification experiments for nonlinear dynamic process models: a stochastic global optimization approach," *Ind. Eng. Chem. Res.*, vol. 41, no. 10, pp. 2425–2430, 2002.
- [5] D. Faller, U. Klingmüller, and J. Timmer, "Simulation methods for optimal experimental design in systems biology," *Simulation*, vol. 79, no. 12, pp. 717–725, 2003.
- [6] M. H. Asyali, "Design of optimal sampling times for pharmacokinetic trials via spline approximation," *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 18, no. 6, pp. 1019–1030, 2010.
- [7] E. Van Derlinden, L. Mertens, and J. F. Van Impe, "The impact of experiment design on the parameter estimation of cardinal parameter models in predictive microbiology," *Food Control*, vol. 29, no. 2, pp. 300–308, 2013.
- [8] H. Yue, M. Brown, F. He, J. Jia, and D. B. Kell, "Sensitivity analysis and robust experimental design of a signal transduction pathway system," *Int. J. Chem. Kinetics*, vol. 40, no. 11, pp. 730–741, 2008.
- [9] M. Brown, F. He, and L. F. Yeung, "Robust measurement selection for biochemical pathway experimental design," *Int. J. Bioinformatics Res. Appl.*, vol. 4, no. 4, pp. 400–416, 2008.
- [10] F. He, M. Brown, and H. Yue, "Maximin and bayesian robust experimental design for measurement set selection in modelling biochemical regulatory systems," *Int. J. Rob. Nonl. Control*, vol. 20, no. 9, pp. 1059–1078, 2010.
- [11] D. J. Pritchard and D. W. Bacon, "Prospects for reducing correlations among parameter estimates in kinetic models," *Chem. Eng. Sci.*, vol. 33, no. 11, pp. 1539–1543, 1978.
- [12] G. Franceschini and S. Macchietto, "Anti-correlation approach to model-based experiment design: application to a biodiesel production process," *Ind. Eng. Chem. Res.*, vol. 47, no. 7, pp. 2331–2348, 2008.
- [13] L. Ljung, *System identification*. Springer, 1998.
- [14] L. Hosten, "A sequential experimental design procedure for precise parameter estimation based upon the shape of the joint confidence region," *Chem. Eng. Sci.*, vol. 29, no. 11, pp. 2247 – 2252, 1974.
- [15] H. Nouredini and D. Zhu, "Kinetics of transesterification of soybean oil," *J. American Oil Chemists' Society*, vol. 74, no. 11, pp. 1457–1463, 1997.
- [16] G. Franceschini and S. Macchietto, "Novel anticorrelation criteria for model-based experiment design: Theory and formulations," *AIChE J.*, vol. 54, no. 4, pp. 1009–1024, 2008.
- [17] K. A. P. McLean and K. B. McAuley, "Mathematical modelling of chemical processes obtaining the best model predictions and parameter estimates using identifiability and estimability procedures," *The Canadian J. Chem. Eng.*, vol. 90, no. 2, pp. 351–366, 2012.