Optimal Input Design for Reduction of Parameter Correlations

Ke Wang, Hong Yue, Hui Yu

Department of Electronic & Electrical Engineering
University of Strathclyde
Glasgow, UK
Email: k.wang@strath.ac.uk; hong.yue@strath.ac.uk; hui.yu.100@strath.ac.uk

Abstract— An new scalarisation criterion is proposed for optimal experiment design (OED) of input intensity so as to obtain the most informative experimental data for parameter estimation with reduced parameter correlations. This criterion is a linear combination of logarithm function of the A-optimality and the modified E (ME)-optimality. It can be used to improve the estimation quality from the A-optimal design, and to reduce parameter correlations from the ME-optimal design. The proposed algorithm has been examined through simulation study of an enzyme reaction system model. The results are compared with A-optimal design, ME-optimal design, and other designs with a focus on reducing parameter correlations such as the C- and the CE- designs.

Keywords - parameter estimation; optimal experimental design (OED); optimal input design; reduction of parameter correlation

I. INTRODUCTION

Optimal experimental design (OED) is a useful and efficient technique for systematically building mathematical models. OED has great advantages in situations where conducting experiments is resource and time consuming. The idea of OED was first introduced by Fisher in 1937 [1], then developed by many other researchers [2-4]. Most works on OED only consider the relationships between input and output, which is referred to as the "black-box" design. Another type of design uses underlying physics of the system or process to propose one or several models, then the OED is used to support model discrimination and parameter estimation. The latter is called "model-based OED."

As discussed above, one crucial part in building mathematical model is estimating the unknown parameters in the model. It is usually called parameter estimation (PE) problem or model fitting problem. One typical approach to solve PE problem is the Least square estimation (LSE) which is based on the assumption of measurement error being normally distributed, and if this assumption cannot be satisfied then other parameter estimation structure such as guaranteed parameter estimation will be adopted [5]. The quality of estimation in the context of LSE is often evaluated statistically using the Student t-value for each parameter [6]. However, for multi-parameter models, high correlations between parameter pairs will result in very low t-values which may lead to inappropriate analysis with the estimation significance [7]. Therefore, joint confidence regions are often employed in multi-parameter models to assess the statistically quality of parameter estimation.

Commonly, ellipsoids will be used to (over-)approximate joint confidence regions, while other type of over-approximation method such as orthotope, zonotope also can be used [8].

In this paper, OED are used mainly for the purpose of designing the most informative data which can improve the estimation precision of parameters. Information contained in data is measured using Fisher information matrix (FIM). Since the joint-confidence ellipsoid can be approximately calculated using the inverse of FIM, we can modify the joint-confidence ellipsoid by a optimisation problem with the objective function being the scalar function of FIM. Therefore, OED problem can be formed as an optimisation problem and its objective function is a scalarisation criterion of FIM. Standard alphabet criterion including A-optimality, D-optimality and E-optimality, but none of them can deal with the correlations between parameter pairs, and more often, increase the correlations [9].

Reference [10] has brought out a modified E (ME)-optimal criterion which uses the condition number (ration between the maximum eigenvalue and minimum eigenvalue) as the objective function. This criterion is used to reduce the correlations between parameter pairs by works such as [11]. However, as pointed in many works including [12], [13] and [14], this criterion has many drawbacks. First of all, since it only aims at reshaping the joint-confidence ellipsoids to make them as spherical as possible, it may result in very large volumes to those ellipsoids. Secondly, the conditional number is normally discontinuous which will cause convergence problems in gradient based optimisation methods. Thirdly, this criterion has the limitation that is only useful to two parameter models.

In order to tackle the first problem in ME-optimality, we introduce an improvement in this paper by combining the A-optimality with the ME-optimality together to ensure a satisfactory estimation precision while reducing the correlations between parameters. This criterion will be applied in an enzyme reaction system which is introduced in [15].

Other criteria aiming at reducing parameter correlations including the C-optimality criterion and the C-optimality criterion with the constraints on eigenvalues (which will be called the CE-optimality criterion in this paper) are proposed in [12]. C-optimality criterion uses the sum of square values of all correlation coefficients as the objective function, but it also has the problem that it does not take

information content into consideration which will cause large confidence interval. CE-optimality criterion design which uses the same objective function as the C-optimality criterion but added a constraint that the maximum eigenvalue of FIM should be larger than a certain threshold to ensure the information contained in the data. However, how to choose the threshold value becomes a big problem. Improperly choosing the threshold value will cause poorly design results which still has large confidence interval. Those methods will also be conducted in this paper as comparisons to our method.

The remaining of the paper is organised as follows. In Section II, the methodology for parameter estimation and OED, as well as our proposed criterion will be presented. A case study is introduced in Section III, for which the performance using the standard design criteria and the new criterion will be compared. Besides, comparison between the new criterion and other designs for correlation reduction will be discussed. Conclusions are given in Section IV.

II. METHODOLOGY

A. Parameter Esitmation and Optimal Experimental Design

In model-based OED, the model is often described by a series of nonlinear ordinary differential equations (ODEs):

$$\dot{X}(t) = f(X(t), \theta), \; X(t_0) = X_0 \tag{1}$$

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

where $X = [x_1, x_2, \cdots x_n]^T$ is a $n \times 1$ vector of state variables, $f(\cdot)$ is a set of state transition functions of the system dynamics which are assumed to be continuous and first-order differentiable, $\theta = [\theta_1, \theta_2, \cdots \theta_p]^T$ is a $p \times 1$ vector of parameters, $Y = [y_1, y_2, \cdots y_m]^T$ is a $m \times 1$ vector of measureable outputs, h is normally a selection function and ξ is a vector of measurement errors which is assumed to be *i.i.d.* zero mean Guassian noise.

Parameters in (1) and (2) can be estimated using least-square estimation:

$$\hat{\theta} = arg \min_{\theta \in \Theta} \sum_{l=1}^{N} \left(Y(t_l) - \hat{Y}(t_l) \right)^T \cdot Q^{-1} \cdot \left(Y(t_l) - \hat{Y}(t_l) \right) \tag{3}$$

where $Y(t_l)$ is the measurement data in time t_l, $\hat{Y}(t_l)$ is the model prediction, N is the total number of sampling time and Q is the measurement error covariance matrix.

As discussed before, the statistical quality of estimations is assessed by joint-confidence ellipsoid. This region can be calculated approximately using the first order Taylor series approximation to the expected function [6]:

$$\left(\theta - \hat{\theta} \right)^T \cdot F \cdot \left(\theta - \hat{\theta} \right) \leq p \cdot F_{dist}(p, N - p, \alpha) \tag{4}$$

$$F = \sum_{i=1}^{N} S^T(t_l) \cdot Q^{-1} \cdot S(t_l) \tag{5}$$

where F is the Fisher information matrix, F_{dist} is the Fisher distribution, α is the level of probability which is often chosen as 95% or 99% and S is the local sensitivity matrix with elements $s_{ij} = \frac{\partial x_i}{\partial \theta_j}$.

Before conducting the OED, we should define the design vector φ. Variables which can be manipulated by the experimenter and optimised during the OED procedure should be selected to the design vector according to the purpose of design. For example, sampling time, system input, system initial condition and measurement set selection are all can be selected as design vector.

Then, the OED can be formed as:

$$\varphi^* = arg \min_{\varphi \in \Phi} g(F) \tag{6}$$

where g is a scaler function that can reflect certain metrics of the FIM. Standard criteria including D-, A-, E-optimality are devised to maximise the determinant, trace and minimum eigenvalue of FIM respectively. Geometrically, they are minimising the volume, enclosing box and major axis of joint-confidence ellipsoids, respectively.

B. Parameter Correlations

The estimation accuracy is often assessed by standard statistical tests such as Student t-test. However, t-test is based on the hypotheses that samples are statistically independent and this hypothesis will be severely damaged by the high correlations among the samples. This means only in the situation that parameters are uncorrelated, parameters can be estimated using confidence interval and statistical test individually and reliably [12].

Moreover, the correlations between parameter pairs have many problems in parameter estimation and OED procedure. First of all, the search for optimal design point can be badly hampered since the correlation will cause an elongated valley surface in objective function [18]. Besides, the effect of one parameter will be counteract by another due to the high correlation between them.

Therefore, in order to make a confident conclusion on the estimation of parameters, we need to acquire data samples with most information and least correlations.

Correlations between parameters can be assessed by correlation matrix C, the coefficients of which can be calculated approximately by the coefficients of variance-covariance matrix V. According to [16] and [17], the variance-covariance matrix can be estimated using FIM:

$$V \approx F^{-1} \tag{7}$$

Then the coefficients of the correlation matrix C can be calculated by:

$$c_{ij} = \frac{v_{ij}}{\sqrt{v_{ii}} \cdot \sqrt{v_{jj}}} \tag{8}$$

C. Our Proposed Optimality Design

Since standard criteria do not consider the correlations between parameters, the ME-optimality is developed, which minimises the conditional number of the FIM.

$$\varphi^* = arg \min_{\varphi \in \Phi} \frac{maxeigen(F)}{mineigen(F)} \tag{9}$$

Geometrically, this criterion can be interpreted as making the joint-confidence ellipsoid as spherical as possible.

In order to overcome the limit that the ME-optimality may cause very large volume of joint-confidence ellipsoid, we proposed a criterion that combines the A-optimality with the ME-optimality linearly. This means we not only want to reduce the parameter pair correlations but also want to restrict the ellipsoid in a small enclosing box. This criterion can be described as:

$$\boldsymbol{\varphi}^* = arg \min_{\boldsymbol{\varphi} \in \Phi} \left(\beta \cdot \log(trace(\boldsymbol{F}^{-1})) + (1 - \beta) \cdot \log\left(\frac{maxeigen(\boldsymbol{F})}{mineigen(\boldsymbol{F})}\right) \right) \quad (10)$$

This proposed criterion will be examined on an enzyme reaction system. Also, the C-optimality criterion and the CE-optimality criterion [12] will be applied, in order to compare the results with our proposed method on the reduction of parameter estimation quality and correlations. These two methods will be introduced briefly here.

1. C-optimality design:
$$\boldsymbol{\varphi}^* = arg \min_{\boldsymbol{\varphi} \in \Phi} \sum_{i=1}^{p-1} \sum_{j=2}^{p} c_{ij}^2 (\hat{\boldsymbol{\theta}}, \boldsymbol{\varphi}) \quad (11)$$

2. CE-optimality design:
$$\boldsymbol{\varphi}^* = arg \min_{\boldsymbol{\varphi} \in \Phi} \sum_{i=1}^{p-1} \sum_{j=2}^{p} c_{ij}^2 (\hat{\boldsymbol{\theta}}, \boldsymbol{\varphi})$$
$$s.t. \ \lambda_i - \lambda_i^0 > \varepsilon_i^\lambda \quad (12)$$

where c_{ij} is the ij-th element of the correlation matrix $\mathbf{C}$, λ_i is an eigenvalue of the FIM $\boldsymbol{F}$ and ε_i^λ is the threshold for λ_i.

III. CASE STUDY

In this section, our proposed criterion will be tested using an enzyme reaction system which is proposed by [15]. First of all, the proposed criterion will be used in OED procedure and the results will be compared with the A-optimality and the ME-optimality designs. Then the C-optimality criterion and the CE-optimality criterion will be applied in the OED procedure to choose the best initial input values. The ability to reduce correlations between parameter pairs will be compared with the proposed criterion.

The model description and the nominal value of parameters are given in the appendix. According to [15], the initial concentrations of S, N, E, which will be written as S_0, N_0, E_0, are chosen as the controllable factor and are selected as the design vector $\boldsymbol{\varphi}$. Parameters k_2, k_{-3} and k_{-5} are determined as the three most important parameters through the sensitivity analysis. These three parameters will be selected for the OED.

A. Optimal Input Design Using the proposed design

Since our proposed criterion is the combination of the A-optimality and the ME-optimality, OED will be conducted using the A-, the ME- and the proposed criterion, and their parameter estimation quality will be compared. The standard deviation is chosen to be 0.1 in this case study when simulating the measurement data.

In the proposed criterion we simply select 0.5 to be the value of β. The optimal design vector that obtained using the active-set method from those three criteria are shown in Table 1.

Table. 1 Optimal input intensities with the A-, the ME- and the proposed ME design

Design criterion	E_0	S_0	N_0
A-optimality	3.5023e-6	0.3982	1
ME-optimality	1.5e-6	0.0018	0.0762
Proposed optimality	1.5e-6	0.0250	0.1069

Unit: mol/L

Using those designed values, we can generate experimental data and then estimate parameters k_2, k_{-3}, and k_{-5}. The estimation quality is presented using joint-confidence ellipse which is shown in Fig.1-Fig.3.
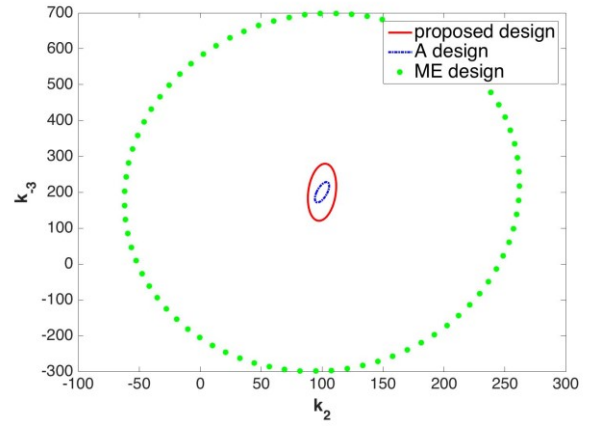

Figure 1. Joint-confidence region for k_2 and k_{-3} using the A−, the ME− and the proposed design criteria
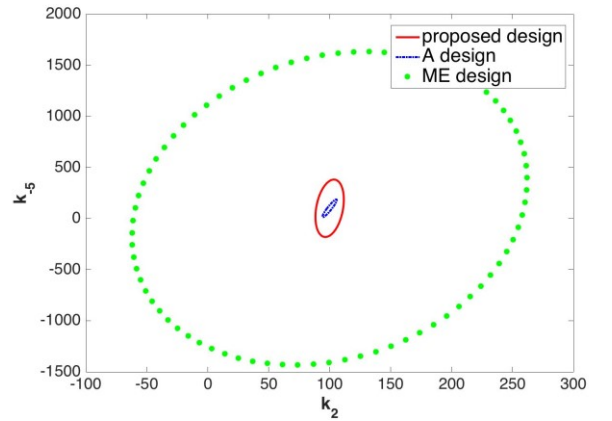

Figure 2. Joint-confidence region for k_2 and k_{-5} using the A-, the ME- and the proposed design criteria
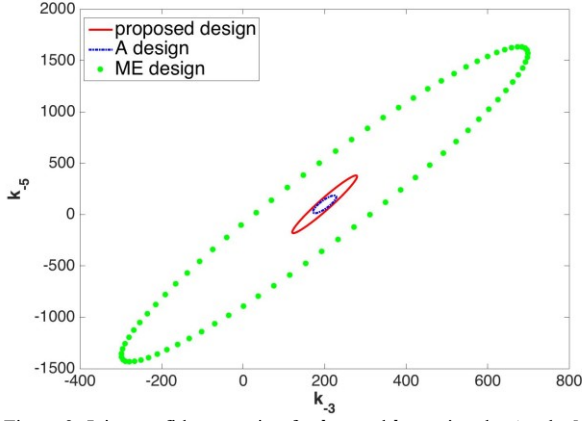
Figure 3. Joint-confidence region for k_{-3} and k_{-5} using the A-, the ME- and the proposed design criteria

It shows that the confidence interval using ME-optimality is too large to make valuable conclusion on the estimation of parameters. Using the ME-optimality design, the reduction of parameter pair correlation sacrifices the parameter estimation precision.

Then, we check the parameter correlations for those three designs:
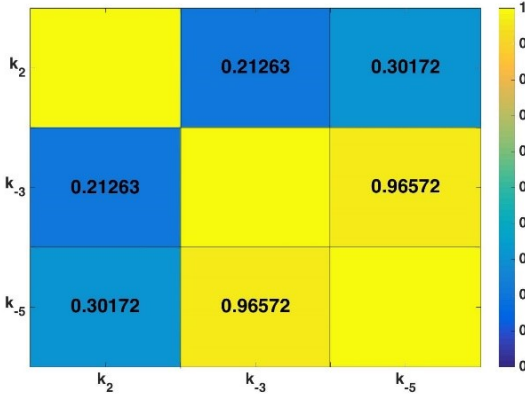

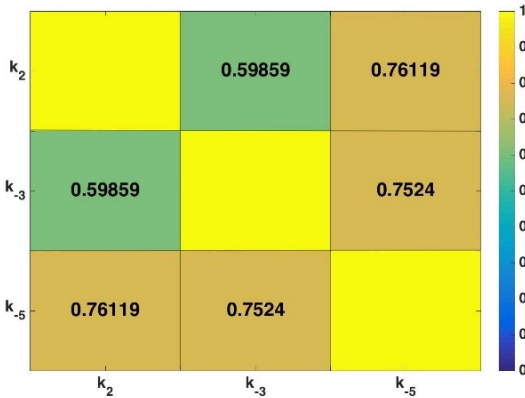Figure 4. Correlations between selected parameter pairs with the proposed design

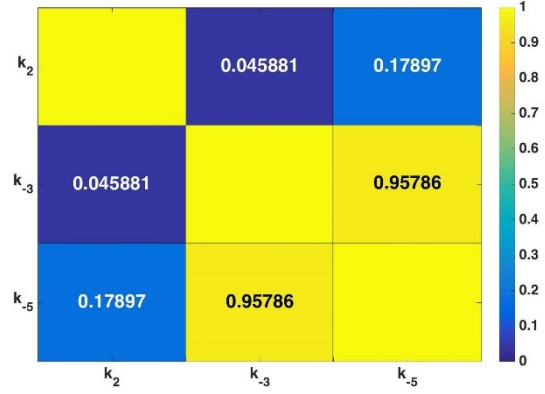Figure 5. Correlations between selected parameter pairs with the A design

Figure 6. Correlations between selected parameter pairs with the ME design

It can be seen that, compare to ME-optimal, the proposed criterion has largely reduced the confidence interval for all three parameters and preserved the low-correlations between parameter pairs $[k_2, k_{-3}]$ and $[k_2, k_{-5}]$. The proposed criterion is actually a compromise between improving parameter estimation precise which is obtained by A-optimality and reducing correlations which is obtained by ME-optimality.

B. Comparesion with Other Correlatin Reduction Designs

This section we will conduct other designs which aim at reducing correlations between parameters including the C-optimality criterion design and the CE-optimality criterion design. The designed results are shown in Table 2.

Table. 2 Optimal input intensities with C- and CE- designs

	E_0	S_0	N_0
C-optimality criterion	1.5e-6	0.01	1
CE-optimality criterion	3.1724e-6	0.0535	1

Unit: mol/L

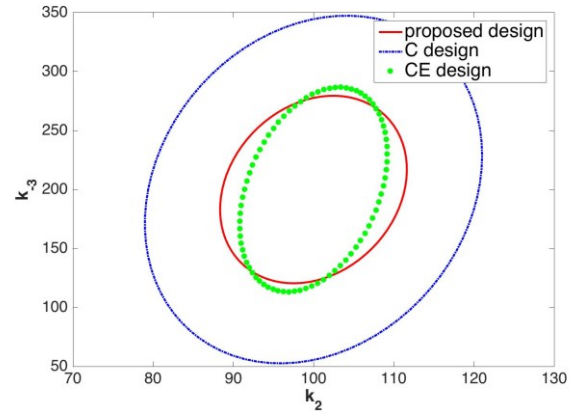Joint-confidence region calculated using designed values are shown in Fig. 7-Fig.9.

Figure 7. Joint-confidence region for k_2 and k_{-3} with the proposed design, the C-optimality design and the CE-optimality design
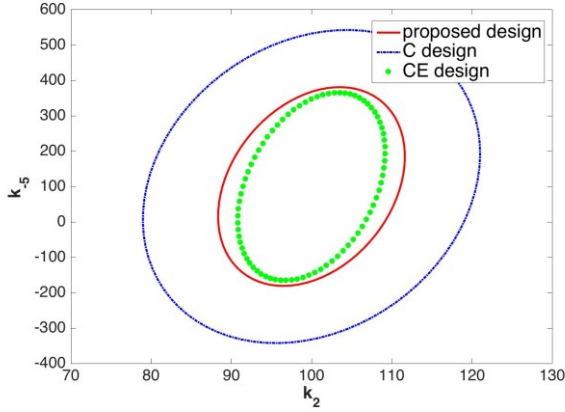
Figure 8. Joint-confidence region for k_2 and k_{-5} with the proposed design, the C-optimality design and the CE-optimality design
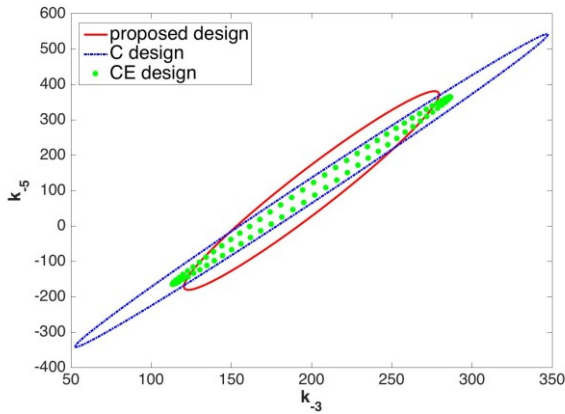

Figure 9. Joint-confidence region for k_{-3} and k_{-5} with the proposed design, the C-optimality design and the CE-optimality design

From the figure above we can see that the C-optimality criterion design has the best effect on reducing the correlation between parameters, but it has a larger joint-confidence region than the proposed design. However, if a constraint is added to the C-optimality criterion design to ensure the information content, which is the idea of CE-optimality criterion design, the effect of reducing parameter correlation is not obvious. The correlations between parameters are shown in Fig. 10 and Fig. 11.
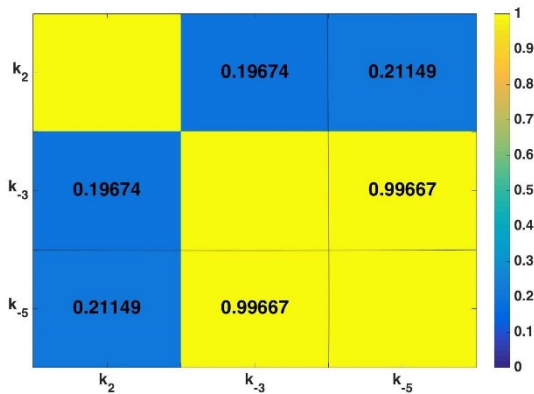

Figure 10. Correlations between selected parameter pairs with the C-optimality criterion design
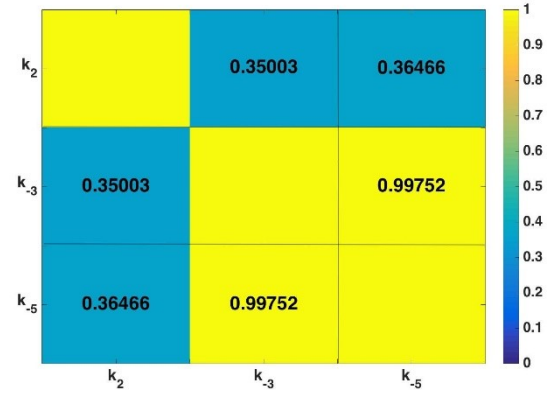

Figure 11. Correlations between selected parameter pairs with the CE-optimality criterion

From the results above, we can see that when the threshold for the CE-optimality design is set to be the same as our proposed criterion, the correlation coefficients is larger than the proposed criterion. This means our criterion has a better performance for the balance between the information content and parameter correlations.

IV. CONCLUSION AND FUTRUE WORK

In this paper, a new criterion has been proposed which linearly combines the A-optimality criterion and the ME-optimality together to acquire the experimental data with satisfactory estimation information and parameter pair correlation reduction. This new criterion has two advantages: 1. It greatly improves the performance of the ME-optimality which often causes convergence problem in gradient-based numerical optimisation approach. 2. It largely reduces the correlations between parameters compared to standard alphabetical criteria such as the A-optimal criterion. It also has better performance than other designs which are mainly designed for correlation reduction. However, reducing confidence interval and reducing correlations are always contradictory and trade-off is always needed during the design.

Since the proposed criterion in this paper is only applied to a simulation model, it should be tested on models with actual measurement data. Besides, since this proposed criterion combines the improvement of parameter estimation quality and the reduction of correlations between parameter pairs, it is actually a multi-objective design. Multi-objective design will be a promising and interesting field in the future work.

REFERENCES

[1] R. A. Fisher, *The Design of Eexperiments*. Oliver And Boyd; Edinburgh; London, 1937

[2] G. E. Box and N. R. Draper, *Empirical Model-building and Response Surfaces*. John Wiley & Sons, 1987.

[3] I. A. Swinnen, K. Bernaerts, K. Gysemans, and J. Van Impe, "Quantifying microbial lag phenomena due to a sudden rise in temperature: a systematic macroscopic study," *Int. J. Food Microbiology,* vol. 100, no. 1-3, pp. 85-96, 2005.

[4] A. Atkinson, A. Donev, and R. Tobias, *Optimum Experimental Designs, with SAS*. Oxford University Press, 2007.

[5] A. R. G. Mukkula and R. Paulen, "Model-based design of optimal experiments for nonlinear systems in the context of guaranteed parameter estimation," *Comput. Chem. Eng.,* vol. 99, pp. 198-213, 2017.

[6] S. P. Asprey and Y. Naka, "Mathematical problems in fitting kinetic models—some new perspectives," *J. Chem. Eng. Japan,* vol. 32, no. 3, pp. 328-337, 1999.

[7] 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.

[8] A. R. G. Mukkula and R. Paulen, "Model-based optimal experiment design for nonlinear parameter estimation using exact confidence regions," *IFAC-PapersOnLine,* vol. 50, no. 1, pp. 13760-13765, 2017.

[9] V. Maheshwari, M. Kandpal, and L. Samavedham, "Multi-criteria optimization based experimental design for parameter estimation of a double feedback gene switching model," *IFAC Proceedings Volumes,* vol. 46, no. 31, pp. 333-337, 2013.

[10] K. Versyck, J. Claes, and J. Van Impe, "Optimal experimental design for practical identification of unstructured growth models," *Mathematics Comput. Simu.,* vol. 46, no. 5-6, pp. 621-629, 1998.

[11] K. Bernaerts, K. J. Versyck, and J. F. Van Impe, "On the design of optimal dynamic experiments for parameter estimation of a Ratkowsky-type growth kinetics at suboptimal temperatures," *Int. J. Food Microbiology,* vol. 54, no. 1-2, pp. 27-38, 2000.

[12] G. Franceschini and S. Macchietto, "Novel anticorrelation criteria for model‐based experiment design: Theory and formulations," *AIChE Journal,* vol. 54, no. 4, pp. 1009-1024, 2008.

[13] F. Sidoli, A. Mantalaris, and S. Asprey, "Modelling of mammalian cells and cell culture processes," *Cytotechnology,* vol. 44, no. 1-2, pp. 27-46, 2004.

[14] D. De Pauw, "Optimal experimental design for calibration of bioprocess models: a validated software toolbox," Ghent University, 2005.

[15] H. Yue, P. Halling, and H. Yu, "Model development and optimal experimental design of a kinetically controlled synthesis system," *IFAC Proceedings Volumes,* vol. 46, no. 31, pp. 327-332, 2013.

[16] L. C. Zullo, "Computer aided design of experiments: an engineering approach," Department of Chemical Engineering and Chemical Technology, Imperial College, 1991.

[17] F. Galvanin, S. Macchietto, and F. Bezzo, "Model-based design of parallel experiments," *Industrial & engineering chemistry research,* vol. 46, no. 3, pp. 871-882, 2007.

[18] A. K. Agarwal and M. L. Brisk, "Sequential experimental design for precise parameter estimation. 1. Use of reparameterization," *Ind. Eng. Chem. Proc. Design and Development,* vol. 24, no. 1, pp. 203-207, 1985.

APPENDIX

The following ODEs are derived for the enzyme reaction system following the mass-balance principle.

$$\frac{dE}{dt} = -k_1 \cdot E \cdot S + k_{-1} \cdot ES + k_4 \cdot EQ$$
$$-k_{-4} \cdot E \cdot Q + k_6 \cdot ER \tag{A1}$$

$$\frac{dES}{dt} = k_1 \cdot E \cdot S - k_{-1} \cdot ES - k_2 \cdot ES$$
$$+k_{-2} \cdot E^* \cdot P \tag{A2}$$

$$\frac{dE^*}{dt} = k_2 \cdot ES - k_{-2} \cdot E^* \cdot P - k_3 \cdot E^* \cdot N + k_{-3}$$
$$\cdot EQ - k_5 \cdot W \cdot E^* + k_{-5} \cdot ER \tag{A3}$$

$$\frac{dEQ}{dt} = k_3 \cdot E^* \cdot N - k_{-3} \cdot EQ - k_4 \cdot EQ$$
$$+k_{-4} \cdot E \cdot Q \tag{A4}$$

$$\frac{dER}{dt} = k_5 \cdot W \cdot E^* - k_{-5} \cdot ER - k_6 \cdot ER \tag{A5}$$

$$\frac{dS}{dt} = -k_1 \cdot E \cdot S + k_{-1} \cdot ES \tag{A6}$$

$$\frac{dP}{dt} = k_2 \cdot Es - k_{-2} \cdot E^* \cdot P \tag{A7}$$

$$\frac{dN}{dt} = -k_3 \cdot E^* \cdot N + k_{-3} \cdot EQ \tag{A8}$$

$$\frac{dQ}{dt} = k_4 \cdot EQ + k_{-4} \cdot E \cdot Q \tag{A9}$$

$$\frac{dR}{dt} = k_6 \cdot ER \tag{A10}$$

where $k_1, k_{-1}, k_2, \cdots, k_6$ are reaction rates, S, Q, P, N and R are donor substrate, desired product, leaving group product, nucleophile and hydrolysis respectively and are measurable. E, ES, E^*, EQ and ER are different form of enzymes and are difficult to measure since concentrations of them are very low. W is water and is constant.

The initial concentration of S, N, E, which will be written as S_0, N_0, E_0 are chosen as the controllable factor and are selected to the design vector $\boldsymbol{\varphi}$. The design space is $S_0 \in [0,1]$, $N_0 \in [0,1]$, $E_0 \in [1.5e-6, 1.5e-4]$ and the nominal values for design vector are $S_0 = 0.8, N_0 = 0.9, E_0 = 1.5e-5$ (unit: mol/L). Concentration of other states are zeros. The nominal value of parameters $k_1, k_{-1}, k_2 \cdots k_6$ are shown in Table A1.

Table A1. Nominal parameter values

Kinetic parameters	Nominal values	Units
k_1	1e5	$mol^{-1} \cdot L \cdot s^{-1}$
k_{-1}	1e3	s^{-1}
k_2	100	s^{-1}
k_{-2}	1e4	$mol^{-1} \cdot L \cdot s^{-1}$
k_3	5e4	$mol^{-1} \cdot L \cdot s^{-1}$
k_{-3}	200	s^{-1}
k_4	1e3	s^{-1}
k_{-4}	2e4	$mol^{-1} \cdot L \cdot s^{-1}$
k_5	5e3	s^{-1}
k_{-5}	100	s^{-1}
k_6	500	s^{-1}