공학박사 학위논문

Model-based Experimental Design for Computationally Efficient Parameter Estimation of Fed-batch Bioreactors

반회분식 생물공정의 파라미터 추정 계산 효율 향상을 위한 모델 기반 실험계획법

2019년 8월

서울대학교 대학원
화학생물공학부
김 정 훈

Abstract

Model-based Experimental Design for Computationally Efficient Parameter Estimation of Fed-batch Bioreactors

Jung Hun Kim
School of Chemical and Biological Engineering
The Graduate School
Seoul National University

Identification of batch dynamical systems is a tricky task because of its complexity and nonlinearity. If the macroscopic structure of a model is available, one can utilize Model-based Design of Experiments (MBDOE) method to facilitate the identification process, more precisely, the parameter estimation. However, a few crucial problems arise in utilizing MBDOE for estimating parameters of batch dynamical systems. First, the whole design depends on the initial estimate of the parameters. Second, the gigantic size of the problem prevents one from obtaining reliable solution in practical amount of time. Third, correlation between the parameters inhibits calculation process of MBDOE. In this thesis, we propose two new schemes of MBDOEs that solve issues of the existing MBDOE schemes. The first MBDOE modifies the existing on-line MBDOE into a form that can be efficiently used in large models, solving initial parameter depen-

dency issue, computation time and sensitivity matrix singularity issue. The second MBDOE improves the existing anti-correlation MBDOE into a form suitable for iterative experiments and causes no numerical instability. Finally, we apply the combined scheme of proposed methodologies to the microalgal bioreactor model to demonstrate its use, as well as study various issues that can occur when the algorithm is applied in actual cases.

Keywords: Batch process, System identification, Parameter estimation, Model-based design of experiment
Student Number: 2013-20962

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Identification of batch processes and experimental designs

Batch processes are widely utilized in chemical industry, especially in the production of specialized chemicals [1], biological product [2, 3, 4, 5, 6], pharmaceuticals [7, 8, 9, 10, 11] and polymers [12, 13, 14, 15, 16]. The above processes are increasing in their importance, due to the current trend towards small quantity production of various specialized products. Just like continuous processes, batch processes require real-time control [17, 18] and process optimization [19, 18, 20] in order to secure price competitiveness. Various types of techniques can be used for control and optimization of batch processes, which can be divided into methods based on a physical model and methods that do not require models. Model-free techniques include methods such as extremum-seeking control [21, 22] and iterative learning control [23, 24, 25], which have a few obvious drawbacks. First, several batch operations must be attempted until the optimum operating condition is found. Secondly, because each attempts of achieving optimality is case-specific, little knowledge can be obtained in the process of optimization. In other words, if the process

constraint changes due to new regulation,or if the objective function changes due to a change in the raw material price, the optimization calculation must be performed again from the zero basis. In industrial practice, the time given to achieve process optimization is limited and the production conditions change almost regularly. This lack of adaptability of the model-free methods makes it unsuitable for batch process optimization at the industrial site.

On the other hand, model-based optimizations can be carried out without burdensome experiments and provide a variety of knowledge about the process. In utilizing model-based methods, the biggest difficulty lies in obtaining a reliable model. The task of modeling can be divided into two parts, structural modeling and the parameter estimation. In structural modeling, what happens inside the reactor is described using first-principle equations and empirical equations. For example, polymerization process is expressed with a set of equations that describes the degree of polymerization with mass and energy balance equations [26, 27]. Likewise, most of the commercial chemical batch processes are fairly well studied for their underlying principles. In other words, structural part of the modeling consists mainly of literature survey, which can be done in relatively short amount of time and little cost. This leaves the parameter estimation process as the only remaining procedure for modeling. The parameter values are usually found by fitting experimental data to the model, where the resulting parameter estimates correspond to the parameters that best account for the data. Using a good quality data for fitting is important here, because accuracy of the parameter estimates depend on the quality of data. Especially when the amount of data is scarce, statistical features of the estimated parameters such as confidence intervals

are highly influenced by the experiment. In the case where one has to invest limited time and money to obtain experimental data, even more importance is given to the informational value of the data.

In order to maximize the informational content of the limited experimental data, one can take advantage of the methodology called model-based design of experiments (MBDOE). MBDOE, which is considered as a type of optimization problem, uses the model structure explicitly to calculate the optimal experimental condition that is expected to produce the most information-rich measurements, which in turn yield the most accurate parameter estimates. The method is already being widely used in the identification of batch chemical and biological processes [28, 29, 30, 31].

1.2 Issues of existing MBDOEs

The problem of finding the optimal experimental conditions for parameter estimation of batch dynamical systems can be defined in a straightforward way using MBDOE, as we will see in Chapter **2.1.1**. However, without any systematic modifications to 'naive' MBDOE, the result of MBDOE will be quite poor for following reasons.

1.2.1 Dependence on the initial parameter estimate

One of the most widely known issues with MBDOE is that the calculation requires, and therefore depends, on the guesses of model parameter values. In other words, there is a contradiction in which the objective of the calculation affects the calculation itself. This problem has been pointed out in literatures on MBDOE, and being considered as the innate limitation of the method [32, 33].

1.2.2 Numerical size of the problem

Another set of problems is caused by the size of the MBDOE problem. In the MBDOE, the amount of information is quantified using a sensitivity matrix which consists of a large number of sensitivity indices. Calculation of each sensitivity index requires numerical integration, which makes the calculation of the sensitivity matrix complicated. Solving MBDOE, which is an optimization problem, requires evaluating the sensitivity matrix over and over. The computational burden of MBDOE is therefore enormous. This leads to a number of other problems beyond simply making calculations take longer. For example, it is very difficult to choose the right solver because the

problem is not only very large, but also extremely nonlinear. When using a global optimization solver such as the genetic algorithm, it takes an unrealistically long time to obtain a solution which is not even reproducible. Using local optimization methods starting from the initial solution has a problem of being very sensitive to the choice of the initial solution. Regardless of the method, it is also a problem that it is impossible to interpret the solution due to the complexity of the problem.

1.2.3 Correlation between the parameters

Another set of problems is caused by correlation between parameters [34]. The estimate of one parameter depends on the parameter estimate of the other, which disables unique determination of parameter estimates. This issue can also be termed as the *practical identifiability* problem. [35] analyzed that, without a specially designed experiment, no unique set of parameter estimates can be determined for Michaelis-Menten kinetic model. Various methods for detecting identifiability have been suggested, including the ones suitable for relatively small-sized problems [36, 37, 38] and the ones suited for nonlinear dynamic models [39, 40].

Moreover, correlation between the parameters leads to numerical problems in solving MBDOE. A column of the sensitivity matrix, i.e., a sensitivity vector, corresponds to one parameters. As the correlation between the two parameters increases, two vectors becomes nearly parallel, which in turn makes the sensitivity matrix nearly singular. This (near–)singularity introduces large numerical errors in calculating performance measures in optimization stage.

1.3 Current approaches to the issues

Fortunately, the issues have been recognized by many researchers, and many suggestions have been made to relive them. In this chapter, we briefly review notable improvements of the MBDOE and their limitations.

1.3.1 Dependence on the initial parameter estimate

To eliminate the effects that comes from the uncertainty of the initial parameter estimates, one can consider directly taking account of the parameter uncertainty. One can try to maximize the minimum possible information content of the experiment, rather than maximizing the objective function based on all possibilities of parameter estimate [41]. If the probability distribution of the parameter estimates is known or can be assumed, a Bayesian experimental design can be calculated [42, 43]. One problem with the above-mentioned methods is that it is very difficult to quantify parameter uncertainty in actual cases. In addition, the actual calculation of above designs is usually very complex, contrary to the simplicity of the idea itself. Another practical approach to deal with the problem is to perform MBDOEs repeatedly. Once the MBDOE is performed with initial parameter estimates, the data is obtained and the parameters are re-estimated using the data. Then, a second MBDOE is calculated using the re-estimated parameters, resulting in further improved parameter estimates. This approach is called sequential design and is commonly used because of its procedural simplicity [44]. The method requires no additional computational complexity, but has a drawback that it takes a lot of time because the experiment has to be repeated and it does not uti-

lize the information obtained during the operation until batch termination time. If the iteration between the parameter update–MBDOE re-calculation is performed in real time instead of batch-to-batch, the real-time data can be utilized, which will further reduce the dependence on the initial parameter estimate. Although this method does not solve the dependency problem perfectly, it is the most advanced form of the existing sequential MBDOEs. There have been a few studies that focuses on the merits of this 'online' MBDOE. Some researchers approached the problem from the adaptive control point of view, where the 'adaptation' refers to the real-time re-estimation of the parameters. Stigter et al. [45, 46] formalized the problem and utilized it in finding parameters of a bioreactor model. Galvanin et al. ,[47, 31] Jayasankar et al. [48], and Zhu et al. [49] used similar frameworks for parameter estimation of relatively simple nonlinear dynamic models. Rathousky et al. [50], Patwardhan et al. [51], Larsson et al. [52, 53], Heirung et al. [54] and Telen [55] respectively suggested a special form of dual adaptive model predictive control, employing a form of DOE metric to ensure a persistent excitation condition.

1.3.2 Numerical size of the problem

The theoretical method for reducing the size of the MBDOE problem, or solving the problems caused by the size, has not been the subject of a devoted study. Instead, practical approaches are taken according to the case. For example, to resolve the dependency problem with regard to the initial guess of the MBDOE, one can establish several initial solutions using expert knowledge and then obtain the

optimal solution by comparing several local solutions. Another common practice is to reduce the number of parameters that are the object of the experimental design. The relative importance of the model parameters, as measured by the expected difference of the measurement caused by the the parameter value differences, varies by orders of magnitude. One can reduce the size of the problem by focusing on a subset of parameters that are far more important than the rest of the parameters. It is notable that on-line MBDOE mentioned in the previous section has problem size-reducing effect, as well as solving the initial parameter estimate dependency problem. This is because on-line MBDOE computes input action for a relatively short time span, rather than calculating the entire batch time. By far, studies that present on-line MBDOEs use a relatively simple model (linear or slightly nonlinear) in their demonstration. When applied to larger and complex models, the simple formulation of online MBDOE is likely to present various problems, as we will see in the Preliminary section.

1.3.3 Correlation between the parameters

There has been a few methods that have been proposed in order to eliminate the correlation between the parameters, in terms of MBDOE. In the simplest case, there is a method of expressing the correlation of all the model parameters as one value and calculating the MBDOE to minimize this value. A set of schemes using a combination of conventional MBDOE and MBDOE that reduces parameter correlation have presented. These methods have been proven successful, yet they are applicable only to a relatively simple model and have

a disadvantage that their calculation is very complicated.

An intuitive way of reducing the correlation should be to provide additional data for the parameter estimation. In [39] and [56], the authors simply provided additional data to the existing dataset, relieving the correlation between the parameters as the result. In another study [40], result of the firstly conducted experiment was analyzed, and then used to design subsequent expeirment aimed at reducing the parameter correlations.

Another set of approaches used the MBDOE framework, using some measure of parameter correlation as the objective function. The simplest design criteira used in this sense is E-optimality, which indicates the smalled eigenvalue of the Fisher's information matrix. A modified E-optimality is similarly defined as the ratio of the minimum eigenvalue to the maximum eigenvalue of the information matrix, and is one of the most widely used anti-correlation design criteria [57, 58]. In the study of Pritchard [59], elements of the correlation matrix was explicitly used as the design criterion and the authors achieved 5% decrement of the correlation indices. This was however at the cost of larger variances for individual parameters. A series of methods that balances between reducing the correlation and minimizing the individual variances was suggested by [60, 61, 62], where anti-correlation criterion is used as the objective function and the conventional criteria is used as constraints, or vice versa.

1.4 Scope of the study

The purpose of modeling is not to obtain a model itself, but to utilize the obtained model to various applications such as model-based optimization, model-based control, scheduling, and so on. If the model parameters are not accurate enough, the results of the optimization and control calculation will also be questionable. The ultimate goal of MBDOE is to find the most accurate possible parameter value so that one can maximize the reliability of the following calculations. Considering this, one can see that the accuracy of the required model depends on the ultimate purpose of the modeling. However, there are so many different areas that the model can be used, so there can be no general answer to that question. Therefore, we consider the variance and correlation index, which are general and simple statistics of parameter estimates, as the primary measure of MBDOE performance, with no consideration of future use of the model. In addition, we will use efficiency in the MBDOE calculation process as the secondary measure of the proposed scheme.

1.5 Outline of the thesis

In Chapter 1, we talked about the necessity of the study and briefly discussed current issues and previous researches. The remaining of thesis will be constructed as follows. In Chapter 2, we will discuss the theory of MBDOE and then discuss the issues described in chapter 1 in more detail. In Chapter 3, we propose a more advanced form of on-line MBDOE, which solves the first two problems of the three problems mentioned previously. Chapter 4 proposes a successive complementary anti-correlation MBDOE as a way to further improve the existing anti-correlation MBDOE. In Chapter 5, we apply the algorithm proposed in the previous two chapters to a microalgal fed-batch bioreactor model and deeply analyze the results from various perspectives.

Chapter 2

Preliminaries

2.1 Model-based design of experiments (MBDOE)

2.1.1 Basic formulation

First of all, it is assumed in our study that our system of interest can be described using a set of ordinary differential equations.

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \boldsymbol{\phi}; \boldsymbol{\theta}) \tag{2.1}$$

$$\mathbf{y} = \mathbf{h}(\mathbf{x}) + \boldsymbol{\epsilon} \tag{2.2}$$

$$\boldsymbol{\epsilon} \sim N(0, \Sigma) \tag{2.3}$$

The dynamics of the states $\mathbf{f}$ is described as a function of the states $\mathbf{x} \in R^{N_x}$, experimental design variables $\boldsymbol{\phi} \in R^{N_\phi}$ and model parameters $\boldsymbol{\theta} = [\theta_1, \theta_2, ..., \theta_P]^T$ as in (2.1). Output variables, or measurements $\mathbf{y}$ is related to the states $\mathbf{x}$ using the function $\mathbf{h}$ and the measurement error $\boldsymbol{\epsilon}$ as in (2.2). $\mathbf{h}$ is usually a selector function, that is, $\mathbf{h}(\mathbf{x}) = \mathbf{h}^T\mathbf{x}$ with $\mathbf{h} = [\delta_1, \delta_2, ..., \delta_{N_y}]$ where $\delta_i = 1$ when the state x_i is measurable and $\delta_i = 0$ when unmeasurable. Measurement noise $\boldsymbol{\epsilon}$ follows a normal distribution with a known diagonal covariance

matrix Σ as in (2.3).

$$\boldsymbol{\theta} \in [\boldsymbol{\theta}_{LB}, \boldsymbol{\theta}_{UB}] \tag{2.4}$$

$$\boldsymbol{\phi} \in [\boldsymbol{\phi}_{LB}, \boldsymbol{\phi}_{UB}] \tag{2.5}$$

Model parameters as well as the design variables are constrained with lower and upper bounds as in (2.4) and (2.5). Design vector $\boldsymbol{\phi}$ in (2.6) contains all elements that can change the measurement values, such as initial state variables $\mathbf{x}(0)$, time-independent control inputs $\mathbf{w}$ and time-dependent control inputs $\mathbf{u}$, and sample instants $\mathbf{t}_{sp}$. Although the time-dependent control input $\mathbf{u}(t)$ can be changed continuously, it needs to be expressed in finite dimension to enable calculation. This dimensional reduction is called control vector parameterization(CVP) and one typical way to perform it is to express control trajectory with a zero-th order hold(ZOH) with control-switching instants $\mathbf{t}_{sw}$.

$$\boldsymbol{\phi} = [\mathbf{x}(0)^T, \mathbf{t}_{sw}^T, \mathbf{u}^T, \mathbf{w}^T, \mathbf{t}_{sp}^T]^T \tag{2.6}$$

Next, we define sensitivity matrx S_t in (2.7) as a partial difference of the output vector $\mathbf{y}(t)$ with regard to the parameters. $\mathbf{y}(t)$ refers to the vector of measurements obtained at the time instant t, i.e., $\mathbf{y}(t) = [y_1(t), y_2(t), ..., y_{N_y}(t)]^T$. This makes the size of the matrix S_t to be $(N_y \times N_p)$, where N_y is the number of output variables and N_p is the number of parameters.

$$S_t(\boldsymbol{\phi}; \hat{\boldsymbol{\theta}}) = \frac{\partial \hat{\mathbf{y}}(t)}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\phi}, \hat{\boldsymbol{\theta}}} \tag{2.7}$$

The dependency shown in (2.7) indicates that in order to evaluate the value of the elements of S_t, one not only needs to provide the design vector $\boldsymbol{\phi}$ but also the current estimate of the parameters $\hat{\boldsymbol{\theta}}$. This point will be discussed in more detail in the upcoming section. If there is only one sampling instant, one can easily obtain the Fisher's information matrix(FIM) M as $M = {S_t}^T \Sigma^{-1} S_t$. When there are more than two intsants of measurement, FIM can be obtained as (2.8) by adding the information matrices of each time instant.

$$M(\boldsymbol{\phi}; \boldsymbol{\theta}) = \sum_{i=1}^{N_{sp}} \sum_{j=1}^{N_{sp}} \Sigma^{-1} S_i(\boldsymbol{\phi}; \boldsymbol{\theta})^T S_j(\boldsymbol{\phi}; \boldsymbol{\theta}) \tag{2.8}$$

$$\boldsymbol{\phi}^* = \arg\max_{\boldsymbol{\phi}} \mathcal{F}\left(M(\boldsymbol{\phi}; \boldsymbol{\theta})\right) \tag{2.9}$$

The values σ_{ij} correspond to the (i, j)–th component of the measurement covariance matrix Σ. The size of the resulting FIM M is $(N_p \times N_p)$, where $(N_y \times N_{sp} \times N_p)$ sensitivity induces have to be calculated therein. A scalar function $\mathcal{F}$ is taken with regard to the information matrix S, which corresponds to some measure of the parameter accuracy. The most common choice of $\mathcal{F}$ is the determinant(D-optimality), which represents the area of the parameter confidence region [32]. Another common choices include the smallest eigenvalue(E-optimality), and trace(A-optimality) [63, 64, 65]. Lastly, optimal experimental design $\boldsymbol{\phi}^*$ is found as an experimental design that minimizes $\mathcal{F}(M)$ in (2.9).

2.1.2 Issues seen in detail

- ***Dependence on the initial parameter estimate***

As can be see by the expression in (2.7), one requires parameter values to calculate the sensitivity matrix. Since the actual parameter values are not known, the best one can do is to use the initial parameter estimate values instead in the MBDOE calculation. This results in erroneous calculation of the information matrix M, which in turn makes the optimal experimental design ϕ^* inaccurate. If the parameter estimate is relatively accurate, the damage due to the parameter uncertainty is small, but it can be serious if the parameter estimate is far from actual value.

The following example demonstrates the parameter estimate dependency issue. MBDOE is performed to a single-input single-output system (2.10), using D-optimality criterion.

$$\begin{aligned} y &= \frac{u}{1-\theta u+1.5\theta u^3}+\epsilon \\ \epsilon &\sim N(0,\sigma^2) \end{aligned} \tag{2.10}$$

Figure 2.1 shows the change of the D-optimality criterion according to the change of the input.

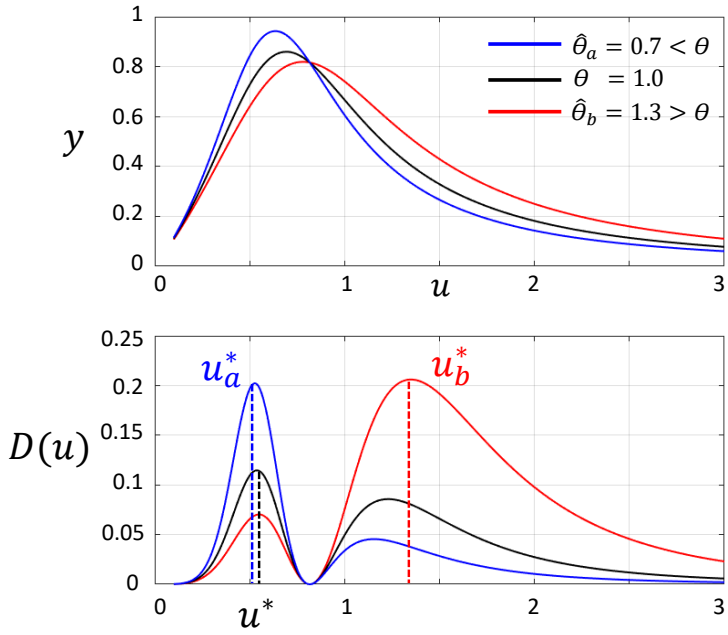


Figure 2.1: Trajectories of output y and D-optimality calculated by different parameter estimates

The optimality critetrion function $(\partial y/\partial\theta)^2$ depends on θ, different curves can be drawn according to the parameter estimate. The actual optimality trajectory, drawn by providing true parameter value $\theta = 1$ is shown in red. Trajectory drawn with the parameter estimate $\hat{\theta}_a = 0.7$ is in blue, and the trajectory drawn with the parameter estimate $\hat{\theta}_b = 1.3$ is shown in yellow. If we perform MBDOE using the parameter estimate $\hat{\theta}_b$, we obtain $u_b^* = 0.527$ as a result, which is a fairly good experiment compared to the actual optimal experiment $u^* = 0.538$. However, when MBDOE calculation is performed starting with $\hat{\theta}_a$, we obtain $u_a^* = 1.349$ that is far different from the actual optimum value. In this case, the amount of benefit one can get from MBDOE is limited.

- ***Numerical size of the problem***

If the input-output relationship of the model is given regardless of the time such as in (2.10), sensitivity values can be easily calculated by differentiating. When the system is described by differential equations (2.1), (2.2) and (2.3), sensitivity values can be calculated by integrating the equation (2.11) obtained by the chain rule.

$$\frac{d}{dt}\frac{\partial \mathbf{y}}{\partial \theta_j} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\frac{\partial \mathbf{y}}{\partial \theta_j} + \frac{\partial \mathbf{f}}{\partial \theta_j} \tag{2.11}$$

The values of the elements of the matrix $\partial \mathbf{f}/\partial \mathbf{x}$ and the vector $\partial \mathbf{f}/\partial \theta_j$, requires the state trajectory $\mathbf{x}(t)$, which means that one has to perform numerical integration of the equation (2.1) in prior. As a result, large number of numerical integrations must be performed to compute a single sensitivity matrix. Again, in order to perform the optimization calculation which sets norm of the FIM as the objective function,

sensitivity matrices have to be evaluated repeatedely. Therefore, the numerical burden of the MBDOE is usually very heavy.

The following example demonstrates the numerical complexity issue. Presented in (5.1) is a fed-batch bioreactor model proposed by Yoo [66].

$$
\begin{aligned}
\frac{dX}{dt} &= \mu X - XD \\
\frac{dS_N}{dt} &= -\rho X + S_N^i \frac{u_N}{V} - S_N D \\
\frac{dS_C}{dt} &= -\frac{1}{Y_{XS}}\mu X - \frac{1}{Y_{LS}}\pi X + S_C^i \frac{u_C}{V} - S_C D \\
\frac{dQ}{dt} &= \rho X - \mu Q - QD \\
\frac{dL}{dt} &= \pi X - vL - LD \\
\frac{dV}{dt} &= u_N + u_C - f_0
\end{aligned}
$$

where

$$
\begin{aligned}
\mu &= \mu_m \left(1 - \frac{q_0}{q}\right)\left(1 - \frac{l_0}{l}\right)\left(\frac{S_2}{K_{S_2} + S_2}\right)\left(\frac{I}{K_I + I}\right) \\
\rho &= \rho_m \left(\frac{S_1}{K_{S_1} + S_1}\right)\left(\frac{q_m - q}{q_m - q_0}\right) \\
\pi &= \pi_m \left(\frac{S_2}{K_\pi + S_2}\right)(1 - q)(1 - l) \\
v &= v_m \left(\frac{K_v}{S_2 + K_v}\right)\left(1 - \frac{l_0}{l}\right) \\
l &= \frac{L}{X + Q + L} \\
q &= \frac{Q}{X + Q + L} \\
D &= \frac{u_N + u_C}{V}
\end{aligned}
\tag{2.12}
$$

Suppose that we try to evaluate sensitivity matrix of the system, with $\mathbf{t}^{sp} = [12, 24, ..., 300]$. All the other simulation conditions such as initial parameter estimate and input trajectory are identical to the conditions given in the original study. First, we obtain $\mathbf{y}(t)$ by numerical integration of the dynamic equation (2.1) and (2.2). And then sensitivity matrices are calculated by integrating (2.11) along with the $\mathbf{y}(t)$ obtained previously. When `ode45` function is used for both integrations, it requires more than 77,000,000 time steps and 42 hours to inquire a single sensitivity matrix S. Stiffness of the sensitivity matrix is responsible for this computation time, so the integration methods suitable for stiff differential equations should be used. When `ode23` function is used instead, it takes 24 seconds for evaluation with 70,000 time steps. For the optimization calculation, this objective function needs to be iteratively calculated. For example, if we use a genetic algorithm with a population of 500 and suppose that it takes 300 generations to obtain an answer. The total time it requires is about 500 * 300 * 24 seconds = 41.7 days, which is makes it impossible to perform optimization for practical uses.

- ***Correlation between the parameters***

The correlation between the parameters can be quantified by the follwing computations. In (2.13), the approximate variance-covariance matrix V is defined as the inverse matrix of the information matrix Z. Elements of the correlation matrix C in (2.14) are found by the computation given in (2.15). Each elements $c_{i,j}$ indicates the degree of correlation between the parameters θ_i and θ_j. Dependencies with

regard to $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ for v_{ii} and c_{ij} are omitted for notational simplicity.

$$V(\boldsymbol{\phi};\boldsymbol{\theta}) = M(\boldsymbol{\phi};\boldsymbol{\theta})^{-1} = \begin{bmatrix} v_{11} & \cdots & v_{1P} \\ \vdots & \ddots & \vdots \\ v_{P1} & \cdots & v_{PP} \end{bmatrix} \tag{2.13}$$

$$C(\boldsymbol{\phi};\boldsymbol{\theta}) = \begin{bmatrix} c_{11} & \cdots & c_{1P} \\ \vdots & \ddots & \vdots \\ c_{P1} & \cdots & c_{PP} \end{bmatrix} \tag{2.14}$$

$$c_{ij} = v_{ij} \,/\sqrt{v_{ii}}\,/\sqrt{v_{jj}} \tag{2.15}$$

2.2 On-line MBDOE

On-line MBDOE maintains the theoretical framework of MBDOE described previously. The difference is that MBDOE problem of reduced size is solved in real-time, and the re-estimated parameter is used in MBDOE calculation of the next time step. When performing real-time control in this manner, it is rational to set the control-switching time $\mathbf{t}^{sw}$ and sampling time $\mathbf{t}^{sp}$ identical. Also, assume that switching and sampling occur at a constant time interval T during the operation, i.e., $\mathbf{t}^{sw} = \mathbf{t}^{sp} = [t_1^{sp}, t_1^{sp}, ..., t_{N_{sp}}^{sp}] = T \times [1, 2, ..., k_f]$. At a given time instant t_k^{sp}, on-line MBDOE solves problems (2.6) through (2.8) using the sensitivity matrix given by (2.16).

$$S_k = S_k(\boldsymbol{\phi}[k]; \hat{\boldsymbol{\theta}}[k]) = \frac{\partial \hat{\mathbf{y}}_i[k]}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\phi}[k], \hat{\boldsymbol{\theta}}[k]} \tag{2.16}$$

$$M_k = M_k(\boldsymbol{\phi}[k]; \boldsymbol{\theta}[k]) = \sum_{m=k+1}^{k+H_p} \sum_{n=k+1}^{k+H_p} \Sigma^{-1} {S_k}^T S_k \tag{2.17}$$

$$\boldsymbol{\phi}^*[k] = \underset{\boldsymbol{\phi}[k]}{\arg\max}\, \mathcal{F}\left(M_k\right) \tag{2.18}$$

In (2.16), $\hat{\boldsymbol{\theta}}[k]$ is the parameter estimate given in time instant k. The measurement vector $\mathbf{y}_i$ is substituted by $\mathbf{y}_i[k]$, which is vector of the measurement y_i of the time instants $(k+1), ..., (k+H_p)$ rather than the entire sampling time instants. The change of the design vector notation from $\boldsymbol{\phi}$ to $\boldsymbol{\phi}[k]$ indicates that the control inputs of the reduced time range is being considered. H_p is the prediction horizon used similarly in the model predictive control (MPC), which can be selected somewhat arbitrarily. Smaller values of H_p indicates that one makes prediction, and bases one's decision upon, relatively small time pe-

riod.

The size of the sensitivity matrix is reduced from $(N_{sp} \times N_p)$ to $(H_p \times N_p)$. This will undoubtably reduce the computational load of the problem, solving numerous problems arises from it. However, two aspects should be noticed:

First, the computation time must be very fast compared to the off-line case in order to proceed in real time without causing problems. For example, if the computation time is equal to the sampling time T, one can set up the theoretical framework of on-line MBDOE without much change. However, the instant at which the new parameter estimate is used in MBDOE calculation is delayed by T. This deteriorates the MBDOE performance accordingly. In order to avoid this, the calculation speed should be very short in comparison with the sampling interval T. In the previous studies on on-line MBDOEs, computation time has never been a issue because they were all applied to linear models or small nonlinear models. In dealing with more complex and highly nonlinear models, computation will certainly be a problem, and it is therefore necessary to further reduce the size of the problem to reduce computation time. The simplest way to reduce the problem size should be to make H_p smaller. However, there is a lower limit imposed on the value of H_p because of the rank of the information matrix. Suppose that one set H_p as 1 to minimize the computatioin. For typical batch models, the number of measured outputs is much less than the number of parameters ($N_y N_p$). In this case, rank of the information matrix M_k equals to N_y assuming no collinearity between sensitivity vectors. Because M_k is not a full rank (N_p) matrix, it is impossible to calculate commonly used optimality criterion. To avoid this, H_p should be equal to or bigger than $[N_p/N_y] + 1$,

which is the minimum H_p that makes M_k full rank. The larger the difference between N_y and N_p, the larger the minimum H_p and heavier calculation. Also, even if M_k meets the full rank condition, it is highly likely that M_k is in ill-conditioned state. To illustrate this, we show the sensitivity matrix in the form (2.19).

$$S_k = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_{Ny} \end{bmatrix}_k \tag{2.19}$$

$$\begin{aligned} M_k &= \sum_{m=k+1}^{k+H_p} \Sigma^{-1} {S_m}^T S_m \\ &= \Sigma^{-1} \Big\{ S_{k+1}^T S_{k+1} + \cdots + S_{k+H_p}^T S_{k+H_p} \Big\} \\ &= \Sigma^{-1} \Big\{ \left(\mathbf{r}_1^T \mathbf{r}_1 + \mathbf{r}_2^T \mathbf{r}_2 + \cdots \mathbf{r}_{N_y}^T \mathbf{r}_{N_y} \right)_{k+1} + \cdots \\ &+ \left(\mathbf{r}_1^T \mathbf{r}_1 + \mathbf{r}_2^T \mathbf{r}_2 + \cdots \mathbf{r}_{N_y}^T \mathbf{r}_{N_y} \right)_{k+H_p} \Big\} \end{aligned} \tag{2.20}$$

Here, $\mathbf{r}_i$ $(i = 1, 2, ..., N_y)$ are row sensitivity vectors representing the sensitivity of a single measurement y_i with regard to all model parameters $\boldsymbol{\theta}$. The information matrix M_k in (2.17) can be expressed as the sum of each information matrix as in (2.20). As exjplained previously, M_k is full rank assuming that there is no parallel columns and $H_p N_y > N_p$, However, dynamic inside the the narrow time horizon is not significantly different, i.e., row vectors $\mathbf{r}_i|_m (m \in [k+1, k+H_p])$ are nearly parallel. As a result, M_k becomes an ill-conditioned matrix which makes the numerical accruacy of the optimization(MBDOE)

calculation poor.

2.3 Anti-correlation MBDOE

Here, we briefly introduce methods for reducing the correlation of parameters using the MBDOE framework. The earliest study of this kind [59] was simply to minimize the elements of the correlation matrix (2.14). A series of advanced and more sophisticated forms of anti-correlation MBDOE were studied [61, 62]. One of methods that the authors have proposed named PAC method is shown below.

$$\begin{aligned} &\min_{\boldsymbol{\phi}\in\boldsymbol{\Phi}} c_{ij}^2 \text{ with i,j such that} \\ &\qquad c_{ij} = \max c|_{basepoint} \\ &\text{s.t. } c_{kl}^2(\hat{\boldsymbol{\theta}}, \boldsymbol{\phi})|_{k\neq l} < \epsilon_{kl}^c \\ &\qquad k, l \in \{1, 2, \cdots, N_p\} \end{aligned} \tag{2.21}$$

In this method, a representative operating policy is set as a reference, and then a primary analysis based on information matrix, variance matrix and correlation matrix is made. Then, the minimization problem is solved with regard to the correlation indices that exceeds the predetermined threshold ϵ_{kl}^c. The other methods proposed in the same paper are variants such as the variances of the individual parameters are used as constraints. The authors used this method to estimate the parameters of the bioreactor model and verified it by experiments.

When this method is applied to a larger sized bioreactor model, a few difficulties are expected. First, the proposed method can give out drastically different results depending on the basepoint one chooses. Moreover, result of the primary anlayses also depends on the value of the parameter estimates. For the above reasons, it is dangerous to use the above anti-correlation method from a point where uncertainty of

parameter estimates is large.

The second reason is that it is not suitable for designing experiments using batches. When the model is large and complex, there is a need to repeat several experiments in order to reduce correlation. However, the existing methods include steps that are complicated and requires expert knowledge. This makes them powerful when performing single experiment, but is not suitable for performing several experimetns in sequence because of the amount of time and effort required to design all experiments. There is a need for a methodology that is simple and consistent regardless of the number of batches, and that can visually observe improvements over each experiment.

Chapter 3

Parameter subset selective on-line MBDOE

3.1 Objective of the methodology

As described in Chapter 1, applying the existing MBDOE methodology directly to a large, nonlinear batch system results in numerous theoretical as well as practical problems. On-line MBDOE is a methodology worth developing further because two of the three issues described in former chapters can be addressed by it. However, as explained in Chapter 2, existing on-line MBDOE methodology is impractical to be applied to a large-sized, highly nonlinear batch systems. In this chapter, we present a way to make the existing on-line MBDOE method more efficient so that it can be used in identification of large and nonlinear models. This method differs from the existing on-line MBDOE in that it involves the process of obtaining a subset of parameters at each time step. The MBDOE problem solved in real time is formulated with regard to a subset of parameters. Using this method, the amount of computation required to calculated optimal input is greatly reduced while the optimality of MBDOE is largely retained. In addition, numerical issues according to the parameter correlation is also solved.

3.2 Theoretical formulation

The procedures described in subsections $\mathbf{3.2.1}$ through $\mathbf{3.2.3}$ are repeated for the time indices $k = 0, 1, ..., k_f$. In this study, we consider the time-varying input $\mathbf{u}[k]$ as the only member of experimental design $\boldsymbol{\phi}[k]$, i.e., $\boldsymbol{\phi}[k] = \mathbf{u}[k]$.

3.2.1 Parameter subset selection

At each time instant k, we select the parameters that dominates the dynamics of the system according to the algorithm presented below. Provided a certain experimental design $\mathbf{u}[k]$ with prediction horizon H_p, sensitivity matrix $S_k(\mathbf{u}[k])$ is evaluated and expressed as a collection of sensitivity vectors (3.4). One should notice that the algorithm $\mathbf{1}$ is repeated at every sampling instant, although the time-dependency is purposely ommitted in order to avoid notational complexity. The algorithm is largely adapted from the subset selection procedure suggested by Chu and Hahn (2012). The difference from the original algorithm is that in the original algorithm, a predetermined number of subset parameters are selected in step 3. At each iteration, one selects one parameter at a time that has the largest norm of the sensitivity vector (3.1), projected by the sensitivity vectors that correspond to all previously selected parameters (3.2).

- Step 1. ***Initiation.*** Number the column vectors of the sensitivity matrix S as $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_{N_p}$. Starting at the iteration index $k = 1$, let the projected sensitivity vectors as $\mathbf{s}_p^{(k)} = \mathbf{s}_p,\ p \in [1, N_p]$.

- Step 2. ***Pre-selection.*** Choose a parameter with the largest norm of *projected* sensitivity vector. Chosen parameter is indexed

with i_k.

$$i_k = \arg\max_j \mathbf{s}_j^{(k)^T} \mathbf{s}_j^{(k)} \tag{3.1}$$

The 2-norm value of the projected sensitivity matrix, i.e. $\mathbf{s}_{i_k}^{(k)^T} \mathbf{s}_{i_k}^{(k)}$ is named as *orthogonal magnitude*, and is recored as m_p if $i_k = p$.

- Step 3. ***Projection.*** Calculate the orthogonal projection matrix $P(\mathbf{s}_{i_k}^{(k)})^{\perp}$. I is the $(N_p \times N_p)$ sized identity matrix.

$$P(\mathbf{s}_{i_k}^{(k)})^{\perp} = I - \frac{\mathbf{s}_{i_k}^{(k)} \mathbf{s}_{i_k}^{(k)^T}}{\mathbf{s}_{i_k}^{(k)^T} \mathbf{s}_{i_k}^{(k)}} \tag{3.2}$$

Using $P(\mathbf{s}_{i_k}^{(k)})^{\perp}$, calculate the next step of projected sensitivity vectors as $\mathbf{s}_p^{(k+1)} = P(\mathbf{s}_{i_k}^{(k)})^{\perp} \mathbf{s}_p^{(k)}$. Set the iteration index $k = k+1$ and return to step 2. Repeat the iteration until the $k = N_p$, i.e., when all the parameters are align in the order $i_1, i_2, ..., i_{N_p}$.

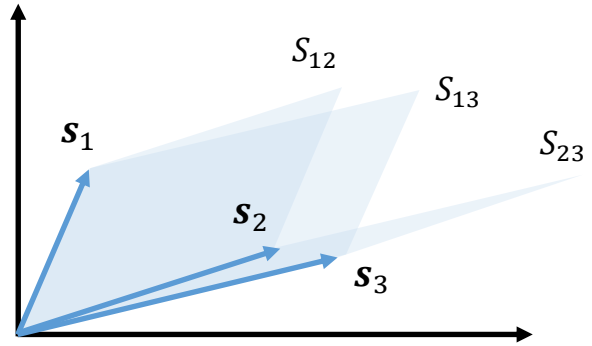
- Step 4. ***Selection.*** After the iteration of steps 2 and 3 is completed, a total of N_p *orthogonal magnitude* values $m_1, m_2, ..., m_{N_p}$ is obtained. Using these, the parameter subset $\tilde{\boldsymbol{\theta}}$ can be selected using different criteria. For example, one can choose the least number of parameters that account for the predetermined portion of the orthogonal magnitudes starting from θ_{i_1}. In this study, we simply choose a predetermined number (N_r) of most significant parameters, that is, N_r parameters with the largest m_p's as the parameter subset.

Chu and Hahn [67] have pointed out the possibility that the successive selection of parameters can result in suboptimal subset selec-

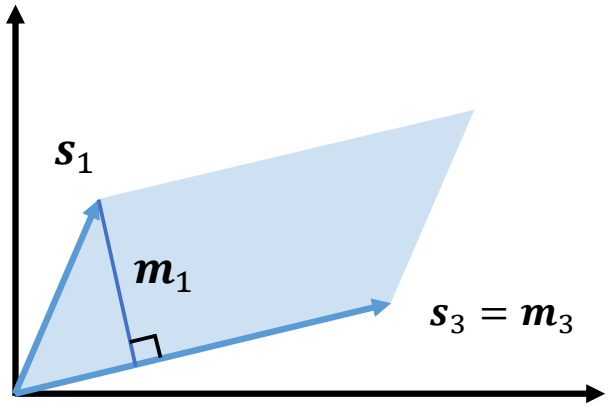
tion compared to the case where all subset parameters are selected simultaneously. However, the chance of suboptimality is considerably lower compared to the case where all instants are considered at once. This is because only a few time instants are considered, where only a few parameters take effect. In fact, subset parameters selected using either method showed no difference in both case studies presented in the current and the later chapter (data not shown). We selected the process given by steps **1** through **4** because the advantage of a significantly faster computation easily overrides its potential weaknesses.

Note that because the evaluation of the sensitivity matrix depends on the selection of experimental design $\boldsymbol{\phi}[k]$, the selection of the parameter subset also depends on $\boldsymbol{\phi}[k]$. To eliminate the effect of the selection of $\boldsymbol{\phi}[k]$, we obtain multiple samples of design vectors that satisfy (2.5). Samples can be drawn from $[\boldsymbol{\phi}_{LB}, \boldsymbol{\phi}_{UB}]$ by dividing them into equally spaced grids or using Latin hypercube sampling if the number of grids becomes excessively large. Orthogonal magnitude values obtained for each grid point are collected, and the total accumulated values for orthogonal magnitude can be used instead for subset selection.

The principle behind the parameter subset selection process can be explained as follows. The D-optimality used as an objective function of MBDOE corresponds to the hyper-volume of the n-dimensional body made by N_p column vectors of sensitivity matrix. For simplicity, consider the case where $N_y = 2$ and $N_p = 3$ where the hypervolume is area as in Figure 3.1(a). In this case, we can define three different areas by combination of three different sensitivity vectors. Each area is equal to the square root of the D-optimality defined by the reduced sensitivity matrix, which consists of two out of three sensitivity vectors. In other words, the D-optimality value changes according to the choice of the parameters. One can use this as the criterion for subset selection; a subset that maximizes the D-optimality(area) is the best choice of the selection. If the number of parameters is small, one can compare all ${}_{N_p}C_{N_r}$ areas and choose the optimal subset. However, as the number of parameters increases, the number of combinations becomes too large, so a simplified suboptimal approach is used where the maximum hypervolume is searched one dimension at a time. Because the hypervolume of the $n + 1$ dimension is hypervolume of the n dimension times the orthogonal length, the notion of orthogonalization comes naturally. However, the purpose of this orthogonalization is not to go all the way through N_p parameters and find orthogonal sets of vectors but to compute the hypervolume of vectors. The way that the orthogonal vectors is obtained in each iteration is the same as that of Gram-Schmidt process. This method is known to cause large numerical errors in the calculation process, compared to other orthogonalization methods such as Householder transformation. When the parameter subset is obtained in this way, the calculation result may become erroneous especially when the number of parameters N_p and



(a)



(b)

Figure 3.1: (a) Geometrics of he sensitivity vectors and the D-optimality of the sensitivity matrices comprised of subset sensitivity vectors. (b) Geometry of the parameter subset selection process and the reduced D-optimality critrion.

subset parameters N_r are both large and the difference between the orthogonal magnitude values is large. In this case, one should keep in mind that the obtained subset can be erroneous.

3.2.2 Optimal input calculation

The reduced form of the MBDOE problem is formulated using the parameter subset $\tilde{\boldsymbol{\theta}}[k]$ found from the previous analysis. The design vector has the form of (3.3), and the expression for the reduced sensitivity matrix is shown in (3.4). To keep the computational load tractable for online application, the zeroth order hold given by (3.5) is introduced for input variables.

$$\mathbf{U}[k] = [\mathbf{u}[k]^T, \mathbf{u}[k+1]^T, \cdots, \mathbf{u}[k+H_p-1]^T]^T \tag{3.3}$$

$$S_k(\mathbf{U}[k]) = \begin{bmatrix} \frac{\partial \hat{\mathbf{y}}[k+1]}{\partial \tilde{\boldsymbol{\theta}}}\Big|_{\mathbf{u}[k]} \\ \frac{\partial \hat{\mathbf{y}}[k+2]}{\partial \tilde{\boldsymbol{\theta}}}\Big|_{\mathbf{u}[k],\mathbf{u}[k+1]} \\ \vdots \\ \frac{\partial \hat{\mathbf{y}}[k+H_p]}{\partial \tilde{\boldsymbol{\theta}}}\Big|_{\mathbf{u}[k],\mathbf{u}[k+1],\ldots,\mathbf{u}[k+H_p-1]} \end{bmatrix} \tag{3.4}$$

$$\mathbf{u}[k] = \mathbf{u}[k+1] = \cdots = \mathbf{u}[k+H_p-1] \tag{3.5}$$

Solving optimization problems (2.18),(2.17),(2.16) with conditions (3.3) and (3.5) yields optimal input $\mathbf{U}^*[k]$. Now, a parameter subset is found again using the sensitivity matrix calculated by substituting $\mathbf{U}^*[k]$ in (3.4).

3.2.3 Implementation and parameter re-estimation

The first element $\mathbf{u}^*[k]$ of the optimal input calculated in the previous step $\mathbf{U}^*[k]$ is implemented to the plant, and the measurement is performed according to (2.2) and (2.3). The parameters are re-estimated as the values that best describe the measurement values up to that moment.

$$\hat{\boldsymbol{\theta}}[k] = \arg\max_{\boldsymbol{\theta}} J[k] \tag{3.6}$$

$$J[k] = \sum_{\kappa=1}^{k} \|\mathbf{y}[\kappa] - \hat{\mathbf{y}}[\kappa]\|^2 \tag{3.7}$$

$$\hat{\mathbf{y}}[\tau] = \mathbf{y}[\tau-1] + \int_{(\tau-1)T}^{T} \mathbf{f}\left(\mathbf{x}, \mathbf{u}^*[\tau-1]; \hat{\boldsymbol{\theta}}[\tau-1]\right) dt \tag{3.8}$$

The parameter estimate obtained at the final time instant $\hat{\boldsymbol{\theta}}[k_f]$ is the final parameter estimate. Figure 3.2 briefly summarizes the scheme of the proposed methodology.

3.3 Demonstration

In order to demonstrate the use of the method and prove, the scheme is applied to a fed-batch bioreactor model. This two-state, four-parameter model was has been utilized in a few MBDOE studies.

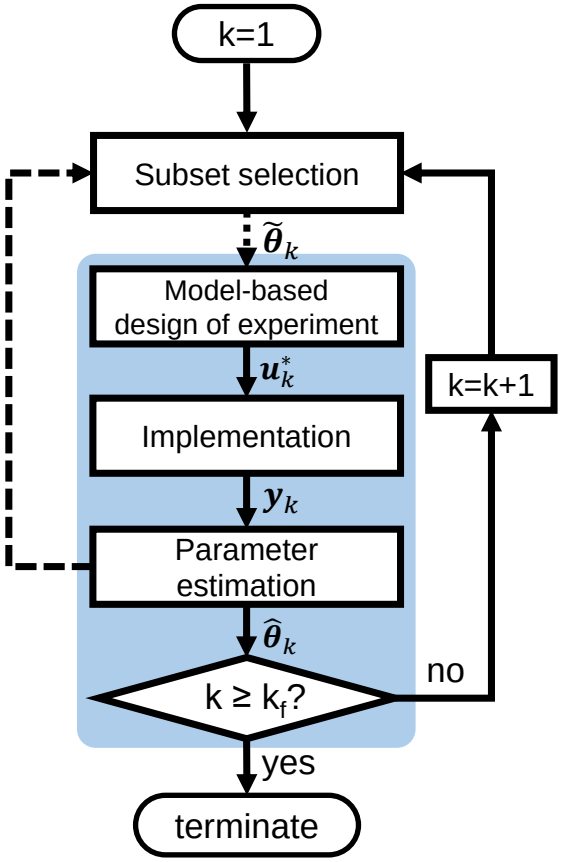


Figure 3.2: Scheme of the subset-selective on-line MBDOE methodology

3.3.1 Model description and problem settings

$$\begin{aligned} \dot{X} &= (\mu - D - \theta_4)\, X \\ \dot{S} &= \frac{1}{\theta_3}\mu X - D(S_{in} - S) \\ \mu &= \frac{\theta_1 S}{\theta_2 + S} \end{aligned} \tag{3.9}$$

X denotes the concentration of the biomass in g/L, and S is the concentration of the substrate in the media, also given in g/L. $D[h^{-1}]$ is the dilution factor, which is a time-varying controllable input of the system. The second time-variant control element is $S_{in}[g/L]$, which is the concentration of the substrate feed. It is assumed that the during the batch duration $T = 40h$, control-switching, sampling, and parameter re-estimation is performed every 4 hours. As a result, both time-variant controls changes 10 times and 20 measurements in total are obtained for parameter estimation. Between each control-switching time instants, both control variables D and S_{in} are assumed to hold as the same value. The admissible ranges for D, S_{in}, and $X(0)$ are [0.05,0.2], [5,35] and [1,10], respectively. The variance matrix Σ is set as $diag([0.1, 0.03])$ and the initial values of the state variables were set as $X[0] = 5.5$ and $S[0] = 0$. The actual values for the parameters $\theta_1, \theta_2, \theta_3, \theta_4$ are $0.31, 0.18, 0.55, 0.05$. For the initial estimate of the parameters, i.e. $\hat{\boldsymbol{\theta}}[0]$, $[0.62, 0.09, 0.8, 0.1]$ is chosen, representing double or half of the actual values. The number of subset parameters N_r selected at each time point is set to 3. In this demonstration, the on-line MBDOE which does not choose the parameter subset is

also calculated, and the two results are compared with each other. To solve this problem, SQP method was used and all calculations were performed in MATLAB 2017b.

3.3.2 Result

The trajectory of the optimal input value $\mathbf{u}^*[k]$ obtained from reduced MBDOE is shown in the following figure. In order to distinguish from the optimum input value obtained from the reduced MBDOE ($\mathbf{u}^*[k]$), the optimum input value obtained from the full-sized MBDOE is indicated as $\mathbf{u}^*_{full}[k]$.

In Figure 3.3, we see that the optimal input trajectory in case of subset parameter selection and no selection results in similar trajectory. It can be interpreted that this is caused by the fact that the relative importance of the parameters at a certain time instant differs by orders of magnitude. The significance of the four parameters during the batch, calculated with the method presented in Chapter $\mathbf{3.2.1}$ is shown in Figure 3.4.

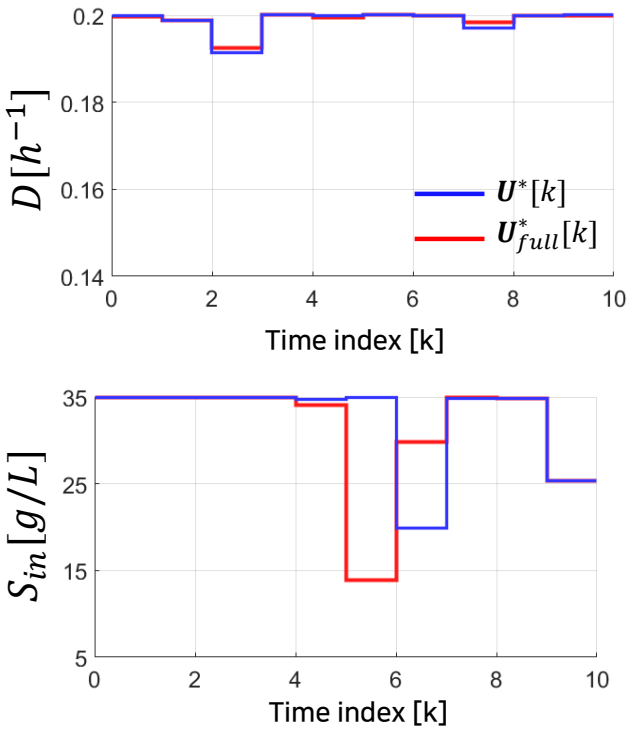


Figure 3.3: Optimal input trajectories for reduced-sized on-line MBDOE and full-sized on-line MBDOE

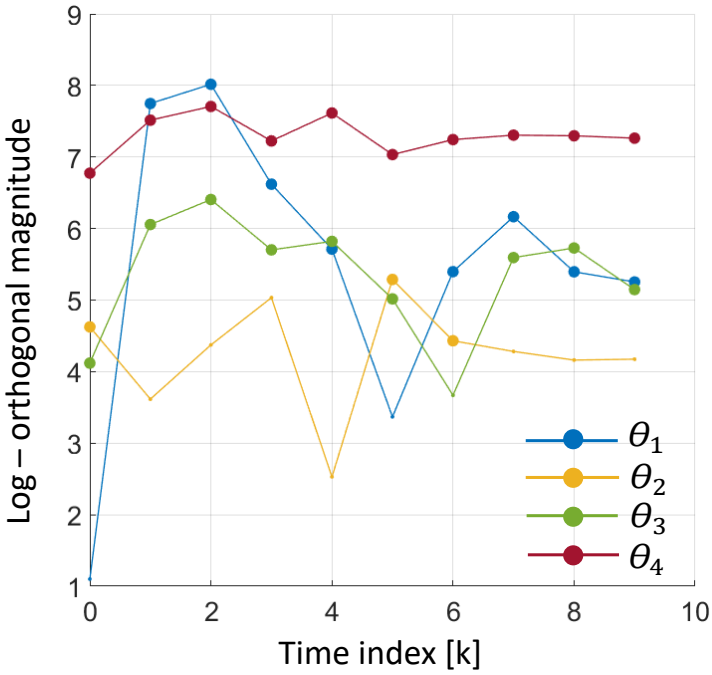


Figure 3.4: Magnitude of orthogonal elements

The parameters selected as subset parameters are indicated with large markers. Here we can see indirectly that the influence of non-critical parameters on the information matrix is negligible, and inclusion of these parameters in the design process has no significant effect on the result. The state variables for both cases follows similar trajectory (Figure 3.5), and also the parameter estimate values obtained online (Figure 3.6). In order to compare the performance between the full-sized D-optimal design and the reduced-sized D-optimal design, D-efficiency is defined in the following manner.

$$D_{eff}[k] = \frac{log[det(S_k(\mathbf{U}^*[k];\boldsymbol{\theta})^T \Sigma^{-1} S_k(\mathbf{U}^*[k];\boldsymbol{\theta}))]}{log[det(S_k(\mathbf{U}^*_{full}[k];\boldsymbol{\theta})^T \Sigma^{-1} S_k(\mathbf{U}^*_{full}[k];\boldsymbol{\theta}))]} \quad (3.10)$$

Simply put, this value represents the ratio of the logarithm of the D-optimality of the two optimal inputs obtained by the reduced MBDOE and the full-sized MBDOE, respectively. Because the full-sized optimal design $\mathbf{U}^*_{full}[k]$ is the optimized value with regard to the denominator, the nominator value cannot be larger than the dominator if the optimization is successfully solved. In other words, $D_{eff}[k]$ has a theoretical maximum value of 1, and smaller value than this indicates larger loss of optimality from reducing the problem size by choosing the subset of parameters. The result of this calculation is shown in Figure 3.7. The Figure shows that the efficiency is relatively low at $k = 6$, which is the same instant where the difference between reduced optimal design $\mathbf{u}^*[k]$ and full design $\mathbf{u}^*_{full}[k]$ is evident in Figure 3.3. However, at all other points, the D-efficiency is close to 1, and the lowered D-efficiency does not significantly affect the accuracy of the parameter estimate. In summary, the process of selecting a parameter subset does little to compromise the performance of the

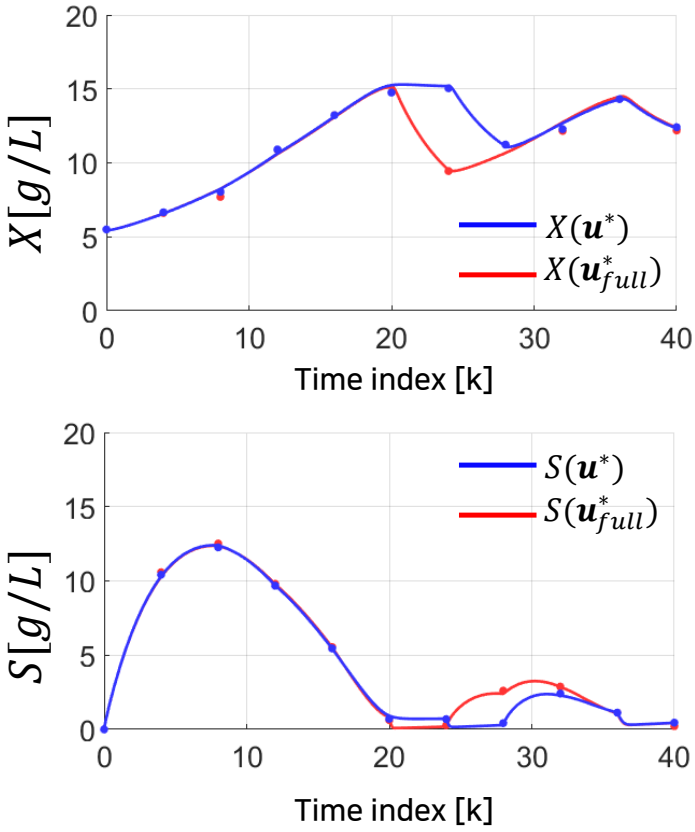


Figure 3.5: Trajectories of state variables when the optimal input sequence is applied

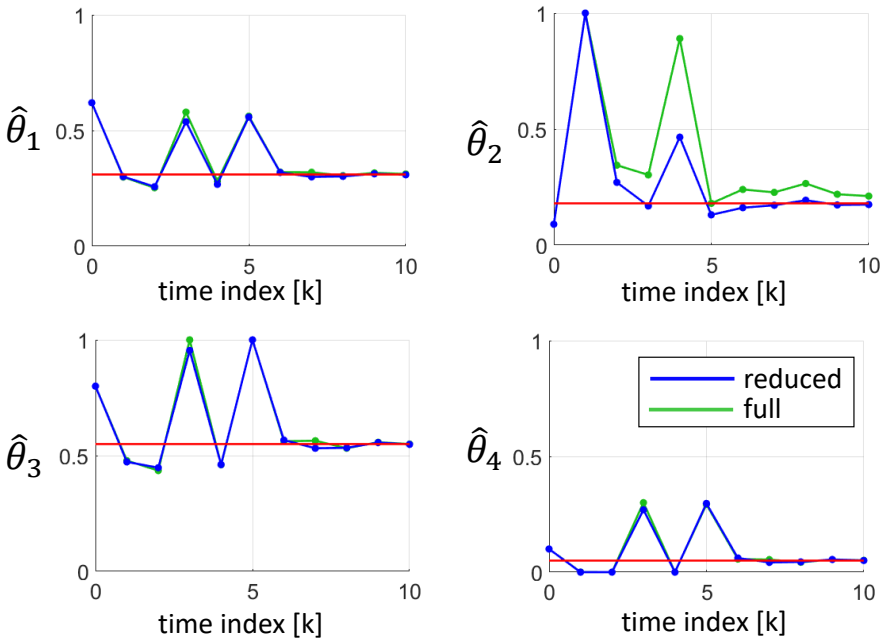


Figure 3.6: Progression of the parameter estimate values

algorithm. Finally, the optimality of the total optimal trajectory calculated online from $k = 1$ to $k = 10$ was compared with the optimality of the input trajectory calculated by the traditional off-line method. The first bar of Figure 3.7 represents the D-optimality of the optimal input calculated assuming that the actual parameter values are known. Obviously, this calculation is impossible in real situation because the true parameter values are unknown. This value stands for the theoretical maximum of D-optimality that can be obtained by any MBDOE calculation. The second bar represents the *actual* D-optimality of the off-line MBDOE calculation done with the inaccuracy parameter estimate. Here, the word *actual* implies that the value computed by MBDOE with an unknown parameter is re-evaluated using the true parameter values. The third and fourth bars represent the *actual* optimality of the two on-line MBDOEs shown in Figure 3.3. We see that the loss of optimality when the entire MBDOE is performed with the inaccurate parameter estimate is considerably large. This loss can be minimized by using on-line MBDOE schemes, and it can be expected that on-line MBDOE can be effectively used in the early stages of parameter estimation of batch systems.

3.3.3 Comparison for different number of subset parameters

In the previous simulation, it has been shown that even if all the parameters are not considered in formulating MBDOE, it has no detrimental effect on MBDOE performance or parameter estimation. However, a single simulation is influenced by the effects of initial parameter estimates as well as random measurement noise. Therefore, we should perform repeated simulations with random initial parameter values, in order to fairly evaluate the effect of N_r. The table 3.1 summarizes the mean values obtained from 100 repeated simulations, for different number of subset parameters $N_r = 1, 2, 3$, and 4(full design). The accuracy of the final parameter estimates were not significantly different from the full design for $N_r = 3$ and $N_r = 2$ cases. However, the uncertainty of the parameters increased sharply when $N_r = 1$. Variances of each parameters were slightly magnified in $N_r = 3$ case, compared to the full design case. Even when the subset parameter was reduced to 2, the variance of the parameters remained largely unchanged, except for the second parameter. When only one parameter was chosen for MBDOE, variances of the parameters 1,2, and 3 were significantly enlarged. However, variance of the fourth parameter was the smallest in this case. This is because when $N_r = 1$, the information is collected exclusively for the parameter 4, which has the biggest sensitivity with regard to the output variables.

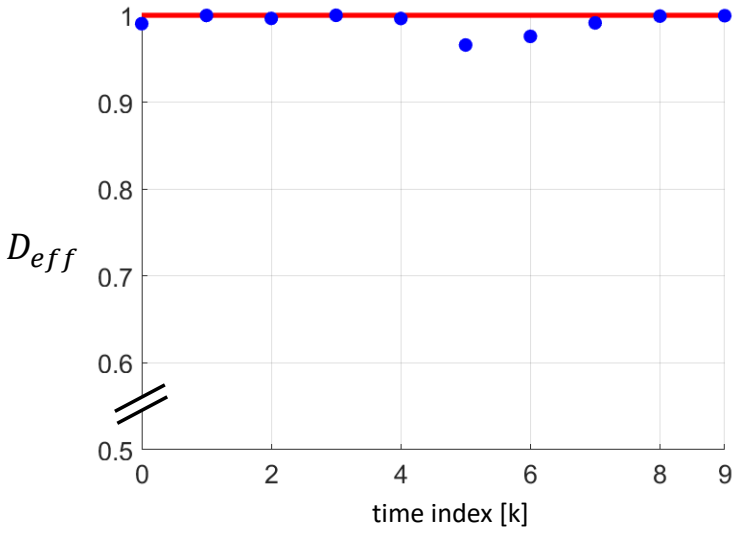


Figure 3.7: Progression of D-efficiency values defined in (3.10)

Number of subset parameters	N_r=4	N_r=3	N_r=2	N_r=1
Sum of squared error of parameters	0.0099	0.0100	0.0118	0.1310
Variance of parameter 1	$0.1060 * 10^{-4}$	$0.1088 * 10^{-4}$	$0.1043 * 10^{-4}$	1.9850
Variance of parameter 2	$0.2984 * 10^{-4}$	$0.2858 * 10^{-4}$	$0.2970 * 10^{-3}$	0.7875
Variance of parameter 3	$0.4544 * 10^{-4}$	$0.4560 * 10^{-4}$	$0.6965 * 10^{-4}$	$1.7318 * 10^{-4}$
Variance of parameter 4	$0.0943 * 10^{-4}$	$0.1007 * 10^{-4}$	$0.1037 * 10^{-4}$	$0.9376 * 10^{-5}$

Table 3.1: Parameter estimation performances with regard to the variation of number of subset parameters

3.3.4 Effect of model conditions and hyper-parameters on the performance of the scheme

- Number of parameters N_p

The various problems of full-scale MBDOEs that do not utilize subset parameters become worse as the number of parameters increases. The larger the number of parameters, the larger the FIM, resulting in a larger nonlinearity of the MBDOE, a longer computation time, and a greater singularity of the sensitivity matrix. In this case, the relative advantage of the reduced-sized online MBDOE becomes greater. No matter how many parameters are used, the number of parameters N_p itself does not affect the performance of the scheme, since the computational and numerical characteristic of reduced MBDOE depend exclusively on N_r rather than N_p.

- Number of subset parameters N_r

As we saw in Chapter 3, as the number of subset parameters N_r increases, the performance of the estimated parameters tends to improve, at the cost of lower computational efficiency. Because both characteristics are crucial in practical implementation of on-line MBDOE, how to choose N_r becomes a very important question. The size of the sensitivity matrix is $H_p N_y \times N_r$, and in order for the FIM to be non-singular, the condition $H_p N_y \geq N_r$ must be met. $N_y \geq N_r$ is obtained for the extreme case $H_p = 1$. The same conclusion can be drawn from an empirical observation of the bioreactor dynamic model. There is usually only one parameter that dominates the behavior of each output value. Therefore, there are N_y parameters that dominants the total N_y output variables, and neglecting the remaining

$N_y - N_p$ parameters in the MBDOE calculation process has little effect on the result. Using the above rule of thumb, setting $N_r = N_y$ is the simplest choice. In fact, in the study performed in current chpater, there was only small difference in parameter estimation performance of in the case of $N_r = 2$ compared to the full design case. However, $N_r = 1$ case showed a significant performance degradation. One way to more rigorously determine N_r is to compare the orthogonal magnitude values of the subset parameters during the process of subset selection. Increasing subset parameters can stop when abrupt decrease of the magnitude value is observed. This method can be used to further generalize the proposed algorith, where the hyper-parameter N_r is also simultaneously calculated as a part of the on-line MBDOE.

- Number of control variables N_u

If the number of inputs is only 1 or 2, the initial search space of the MBDOE is very small, enabling accurate MBDOE calculation. When the number of input variables increases, the initial search space of the MBDOE increases exponentially, and the computation time also exponentially increases accordingly. When the initial search space becomes too large, it becomes impossible to implement on-line MBDOE efficiently. If the number of inputs is too large, one can consider replacing the initial grid for the MBDOE into the a more efficient method such as latin hypercube.

Chapter 4

Successive complementary anti-correlation MBDOE

4.1 Objective of the method

In chapter **1.3.3**, two methods for relieving parameter correlation have been briefly introduced. One method was to simply accumulate sufficient data by repeating several experiments(i.e., batches), and the other was to perform a carefully designed experiment using anti-correlation MBDOE. When only one of the two aforementioned methods is used in obtaining parameter estimates, there is a high chance that the result will not be satisfactory. When additional data is collected using conventional sequential experimental designs, the resulting measurement has little or no effect on reducing the correlations between parameters. As a result, the issue of parameter correlation remains largely unsolved, while the precision of uncorrelated parameters is improved. By contrast, parameter estimates obtained from a single anticorrelating experiment are likely to be relatively uncorrelated, yet the precision of each parameter tends to be inaccurate. Therefore, in order to obtain parameter estimates that are both precise and uncorrelated, it is necessary to incorporate the benefits from both approaches.

A simple way of achieving an anticorrelation feature in the context of multiple experiments may be to design multiple (parallel) experiments as in [68], using anticorrelation design criteria. Despite being viable, this approach is naive and is limited for two reasons. First, the number of design variables for a multiple-experiment MBDOE is much larger than that of a single-experiment MBDOE. The expected consequence is that the calculational burden of MBDOE becomes excessively large, making one unable to obtain a reliable solution in a manageable amount of time. Second, all experimental designs for multiple (parallel) batches depends on initial parameter estimates, which makes the resulting experimental designs unreliable. Although this is a problem for any MBDOE, the effect of the initial estimates is reduced when a sequential strategy is used in which the recursive parameter re-estimation is performed between each experiment. This strategy cannot be utilized for parallel experiments, so it is much more sensitive to the initial parameter estimate.

In this study, we incorporate the anticorrelation approach into the sequential experimental design framework. In the same manner as the conventional sequential experiment, one iterates between the experimental design and the parameter re-estimation. What is different from the existing method is that the design objective function for each batch is defined according to the result from the previous batch and analyses based on it. In other words, a type of information that is lacking from the previous batch is realized and is sought during later experiments. This method utilizes both the anticorrelation and sequential design methods, helping one to obtain the most precise parameter estimate, in terms of variance as well as correlation. Moreover, one can decide when to terminate the sequential design

by comparing the result of each successive experiment. This prevents one from performing unnecessary experiments, which is an additional advantage of the proposed method.

4.2 Theoretical formulation

The scheme consists of three steps. The first experiment is designed in the same way as a conventional MBDOE ($\mathbf{4.2.1}$). By analyzing the result of the first experiment, the parameter estimates are updated ($\mathbf{4.2.2}$), and a new objective function is defined as well ($\mathbf{4.2.3}$). The second set of optimal experimental designs is found with regard to the new objective function, and the procedure is repeated until termination ($\mathbf{4.2.4}$). The entire process is summarized in Figure 4.1.

4.2.1 Initial experimental design

By solving the MBDOE problem (2.7),(2.8) and (2.9) with the initial parameter estimate $\hat{\boldsymbol{\theta}}[0]$, one obtains the first optimal experiment $\boldsymbol{\phi}_1^*$. As for the scalar function $\mathcal{F}$, a determinant (i.e., D-optimality) is recommended because in this way, one is expected to obtain the most 'balanced' experimental design in terms of parameter variance and correlation (wp90Dopt). The resulting design vector $\boldsymbol{\phi}_1^*$ is implemented to obtain the measurement vector $\mathbf{Y}_1$.

4.2.2 Complementary design formulation

In terms of MBDOE, the informational value of the experiment $\boldsymbol{\phi}_1^*$ is summarized in the matrix $M(\boldsymbol{\phi}_1^*; \boldsymbol{\theta})$. The dependency of M on $\boldsymbol{\theta}$ indicates that the exact informational value of the experiment $\boldsymbol{\phi}$ is obtained only when the true parameter value $\boldsymbol{\theta}_{true}$ is supplied. Because this is not the case, one can only resort to the parameter estimate instead. A better parameter estimate is expected to obtain a

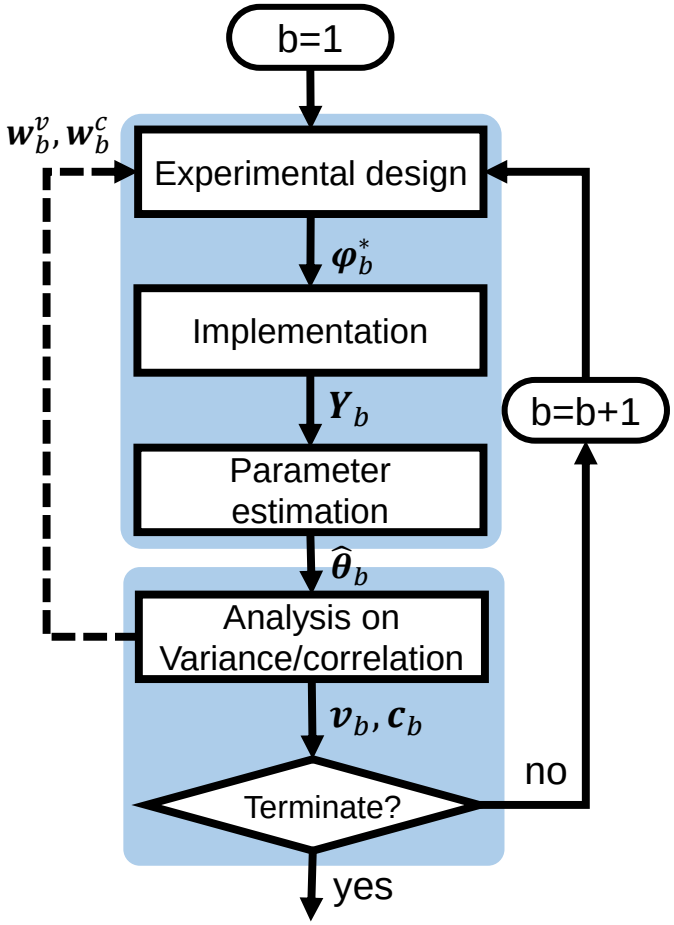


Figure 4.1: Outline of the successive complementary anti-correlation MB-DOE

better approximation of the true informational value of a certain experimental design. In this respect, using a newly obtained parameter estimate $\hat{\boldsymbol{\theta}}_1$ will improve the estimate of the informational value of the experiment $\boldsymbol{\phi}_1^*$, and we use $M(\boldsymbol{\phi}_1^*; \hat{\boldsymbol{\theta}}_1)$ instead of $M(\boldsymbol{\phi}_1^*; \hat{\boldsymbol{\theta}}_0)$ in subsequent analyses of the result. Using (2.13) and (2.14), the approximate values of the parameter variance and correlation indices are calculated: $V_1 \equiv M(\boldsymbol{\phi}_1^*; \hat{\boldsymbol{\theta}}_1)^{-1}$ and C_1 as (2.14) and (2.15) using V_1.

As stated earlier, diagonal elements of V_1 (denoted as $\mathbf{v}_1 = [v_{1,11}, v_{1,22}, ..., v_{1,PP}]^T$) and non-diagonal elements of C_1 (denoted as $\mathbf{c}_1 = [c_{1,12}, c_{1,13}, ..., c_{1,P-1,P}]^T$) encapsulate the statistics of the parameter estimate $\hat{\boldsymbol{\theta}}_1$. Because small values of $v_{1,ii}$ $(i \in [1, P])$ and $c_{1,ij}$ $(i, j \in [1, P])$ indicate precise parameter estimation, we can re-express our goal for the experiments into minimizing each element of $\mathbf{v}_b$ and $\mathbf{c}_b$ for some b under given thresholds $\bar{\mathbf{v}} = [\bar{v}_{11}, \bar{v}_{22}, ..., \bar{v}_{PP}]$ and $\bar{\mathbf{c}} = [\bar{c}_{12}, \bar{c}_{13}, ..., \bar{c}_{P-1,P}]$. One intuitive way of setting the variance threshold values is to set $\bar{v}_{ii} = \alpha\, \theta_i^{nom}$, where θ_i^{nom} indicates the nominal value of the parameter θ_i that represents the parameter magnitude, and α indicates the relative precision that one wants for that parameter. Correlation threshold values $\bar{c}_{ij}$ can be set as constants ranging from 0 and 1. Of course, 'better' threshold values can be chosen by careful examination of the system. For example, one can set $\bar{v}_{ii}$ as the value in which the effect of varying the estimate of parameter θ_i inside the range $\hat{\theta}_i \pm \bar{v}_{ii}$ is negligible for output prediction. However, burdensome calculations such as a global sensitivity analysis need to be performed for this purpose.

Comparing the vectors $\mathbf{v}_1$ and $\mathbf{c}_1$ to $\bar{\mathbf{v}}$ and $\bar{\mathbf{c}}$ reveals the current status of the parameter precision: which parameters' precisions

are satisfied and which are not, and the correlations between which parameters are left to be minimized. We quantify this analysis using the weight coefficients $\mathbf{w}_b^v = [w_{b,11}^v, w_{b,22}^v, ..., w_{b,pp}^v]^T$ and $\mathbf{w}_b^c = [w_{b,12}^c, w_{b,1P}^c, ..., w_{b,P-1,P}^c]^T$ (the batch index b=1 in this case), described in (4.1) and (4.2).

$$w_{b,ii}^v = max\left(1 - \frac{\bar{v}_{ii}}{v_{b,ii}},\ 0\right), \quad for\ \ i \in [1, P] \tag{4.1}$$

$$w_{b,ij}^c = max\left(1 - \left(\frac{\bar{c}_{ij}}{c_{b,ij}}\right)^2,\ 0\right), \quad for\ \ i, j \in [1, P] \tag{4.2}$$

When the parameter variance $v_{b,ii}$ is larger than the threshold $\bar{v}_{ii}$, a positive-valued weight is given. The magnitude of the weight is proportional to the ratio between the two values: a heavier weight is imposed when the parameter variance is too large compared to the desired variance. By contrast, when the desired variance is satisfied, no weight is given to that parameter. The same reasoning is applicable to the weight values $c_{b,ij}$. The difference is that we put a square to the ratio $\bar{c}_{ij}/c_{b,ij}$, considering that the correlation indices can have negative values.

Weight coefficients from the previous experiment are used to define a new objective function for the next set of experiments, as in (4.3).

$$\boldsymbol{\phi}_{b+1}^* = \arg\min_{\boldsymbol{\phi}} \left[(\mathbf{w}_b^v)^T \mathbf{v}_{b+1}(\boldsymbol{\phi}) + \gamma\,(\mathbf{w}_b^c)^T \mathbf{c}_{b+1}(\boldsymbol{\phi})\right] \tag{4.3}$$

Here, the expected variance vectors of the new experiment $\mathbf{v}_{b+1}(\boldsymbol{\phi})$ and $\mathbf{c}_{b+1}(\boldsymbol{\phi})$ are calculated from the information matrix (4.4). The equation (4.4) is based on the additive property of information, that

is, an information matrix of multiple experiments is the sum of the information matrices of the individual experiments. Here, k former experiments are already fixed as $\boldsymbol{\phi}_1^*, \boldsymbol{\phi}_2^*, ..., \boldsymbol{\phi}_b^*$, and the new experiment $\boldsymbol{\phi}_{b+1}^*$ is the only one to be determined. Therefore, the optimization (4.3) yields the experimental design that is dedicated to finding only what is lacking from the previous experiments in terms of information.

$$M(\boldsymbol{\phi}_1^*, \boldsymbol{\phi}_2^*, ..., \boldsymbol{\phi}_b^*, \boldsymbol{\phi}_{b+1}; \hat{\boldsymbol{\theta}}(b)) = \sum_{i=1}^{b} M(\boldsymbol{\phi}_i^*; \hat{\boldsymbol{\theta}}_b) + M(\boldsymbol{\phi}_{b+1}; \hat{\boldsymbol{\theta}}_b) \tag{4.4}$$

Another advantage on can obtain from this formulation is that the numerical instability in MBDOE calculations is greatly reduced. Since anti-correlation MBDOE naturally handles MBDOE with bad condition number in the calculation process, the error in the calculation process is considerably large, and due to the stiffness in calculating the objective function, the calculation time also becomes longer. The 'balanced' information matrix from the previous step acts as a buffer for preventing these numerical stiffness problems. The coefficient γ in (4.3) is the relative weight between the two different objectives, and can be chosen as an arbitrary positive number. Moreover, additional constraints of (4.5) are imposed on the optimization problem (4.3).

$$\begin{aligned} &c_{b+1,ij} \leq \bar{c}_{ij} \\ &\text{for } (i,j) \text{ such that } c_{b,ij} \leq \bar{c}_{ij}. \end{aligned} \tag{4.5}$$

For the correlation indices that were decreased under the threshold value in the kth experiment, these constraints ensure the correlation

indices to be lower than the thresholds in later experiments as well. Otherwise, correlation indices that were sufficiently small can be increased at the cost of minimizing some other correlation values.

4.2.3 Iteration and termination

The procedures described in **4.2.2** and **4.2.3** are iteratively performed until one chooses to terminate the iteration. One chooses to terminate the iteration when one of the three conditions are satisfied. First, one terminates when all the variances and correlation indices are sufficiently small so that no additional experiment is required. The ultimate goal of experimental design is satisfied and one can easily choose to terminate. However, this ideal case is rarely seen in practice. One commonly encounters the situation where a slight increase in parameter precision is expected for an additional experiment, mostly owing to the model structure . When the 'limit' is detected by analyzing the progress of confidence regions or intervals, one can choose to terminate the procedure. The last situation arises when neither the desired precision nor the 'limit' is reached, but the budget (or time) for operating an additional experiment runs out so that one has no other option.

4.3 Case study

In order to demonstrate how to utilize the proposed method, the scheme is applied to a fed-batch bioreactor model. The same model that utilized in Section **3.3** was again used here, however with a different definition of control vector and problem settings.

4.3.1 Model description

$$\begin{aligned} \dot{X} &= (\mu - D - \theta_4)\, X \\ \dot{S} &= \frac{1}{\theta_3}\mu X - D(S_{in} - S) \\ \mu &= \frac{\theta_1 S}{\theta_2 + S} \end{aligned} \tag{4.6}$$

D is expressed as a piecewise-constant input with three switching instants. $S_{in}[g/L]$ is the concentration of the substrate feed and is a time-invariant control of the system. The initial value of the biomass concentration $X(0)$ is another time-invariant control, where the initial substrate concentration $S(0)$ is assumed to be always 0. Measurements of state variables (2.2) are made three times during the operation. Measurement noise is generated according to (2.3) with $\Sigma = diag([0.1, 0.1])$. To summarize, the design control vector $\boldsymbol{\phi}$ consists of 12 elements as

$$\boldsymbol{\phi} = [X(0), t_1^{sw}, t_2^{sw}, t_3^{sw}, D_1, D_2, D_3, D_4, S_{in}, t_1^{sp}, t_2^{sp}, t_3^{sp}]\,. \tag{4.7}$$

The admissible ranges for D, S_{in}, and $X(0)$ are [0.05,0.2], [5,35] and [1,10], respectively. The batch duration T is fixed at 40 h. Moreover, the minimum difference between the control switching instants t_i^{sw} and the sampling instants t_i^{sp} is assumed to be 1.

4.3.2 Solution method

At $b = 1$, a D-optimal experiment $\boldsymbol{\phi}_1^*$ was calculated using a genetic algorithm (GA) with 200 populations. GA was chosen because it is capable of exploring the entire design space. Therefore, its solution is more likely to be near a global minimum. For the subsequent experiment designs $b = 2, 3, ...$, sequential quadratic programming (SQP) was used instead of GA, because it can handle the additional constraint (4.5) much more efficiently than GA. Different solutions from multiple starting points were obtained, because the solution from SQP is sensitive with regard to the initial guess. The initials were chosen according to the following rule.

- X(0) can take one of the values 1, 5.5 or 10.
- $\mathbf{t}^{sw}$ can be either [0.1,1,1,2.1], [8.825,17.55,26.275] or [33,34,35].
- Controllable inputs D_i can be 0.05, 0.125 or 0.2.
- S_{in} is chosen from 5, 20 and 35.
- Sampling instants $\mathbf{t}^{sp}$ is fixed as [10, 20, 30].

In total, $3^4 = 81$ initial points were generated, and the solutions from each initial point were compared. The one with the least objective value was chosen as the final solution. For the desired parameter variance and desired correlation, $\bar{\mathbf{v}} = [0.01, 0.01, 0.01, 0.01]^T$ and $\bar{\mathbf{c}} = [0.5, 0.5, 0.5, 0.5, 0.5, 0.5]^T$ were used. γ=0.5 was used in (4.3),

indicating that we put more emphasis on minimizing the variance of individual parameters. Parameter estimation was also performed by SQP, thus providing the parameter estimate of the previous instant as its initial estimate. For an initial estimate of the parameters, $\hat{\boldsymbol{\theta}}_0 = [0.62, 0.09, 1.00, 0.025]$ was used. Statistics of the parameter estimates calculated after each iteration were analyzed using the latest evaluation of the information matrix $M(\boldsymbol{\phi}_1^*, \boldsymbol{\phi}_2^*, ..., \boldsymbol{\phi}_b^*; \hat{\boldsymbol{\theta}}_b)$. Namely, point estimates and their marginal confidence intervals, and joint confidence regions of parameter pairs were found. A t-score test and chi-square test were performed as well. All calculations were performed in MATLAB R2017b.

4.3.3 Result

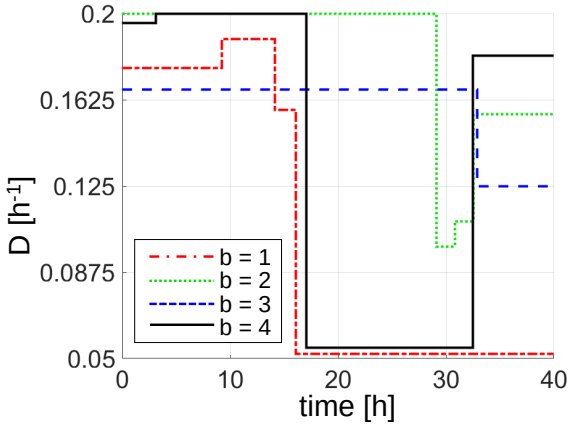
Iteration #1

The first experimental design $\boldsymbol{\phi}_1^*$ was found and implemented (Figures 4.2 and 4.3, indicated by the dotted-dashed red line). Based on the measurement $\mathbf{Y}_1$, a new parameter estimate $\hat{\boldsymbol{\theta}}_1$ was calculated, and the relevant inferences were calculated as well (Table 4.1, first row). Based on the χ^2-statistics, we could see that the measurement data was being successfully described by the model and the parameters. However, the precision of the two parameters θ_2 and θ_4 was questionable, as reflected in the t-score and their marginal confidence intervals. This can also be seen in the variance matrix in Table 4.2, where the $(2, 2)$ element of the matrix is obviously high compared to the other instances. This observation is reflected in the second component of the weight vector $\mathbf{w}_1^v$. In other words, the variance of the second parameter will be part of the objective of the second experiments,

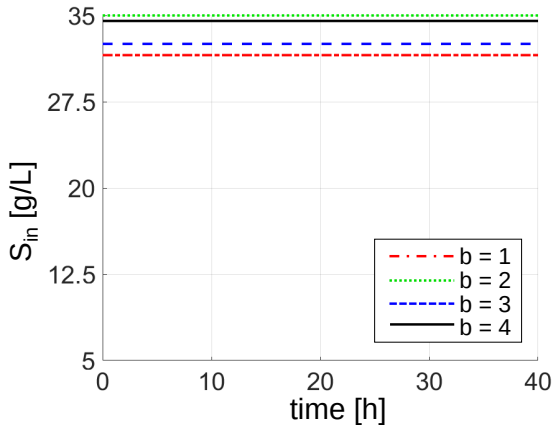
while the variances of the other parameters will not. The correlations of the parameters are presented in the first row of Table 4.3. Compared to the threshold correlation value $\bar{c}_{ij} = 0.5$, only two out of six correlation instances satisfy the condition $c_{ij} \leq \bar{c}_{ij}$. The remaining four correlation indices constitute part of the objective function for designing ϕ_2^*. Now, a total of six different elements of V_2 and C_2 are the objective of minimization for the next experiment. Moreover, two additional constraints were imposed on $c_{2,23}$ and $c_{2,24}$ so that their values do not exceed the threshold $\bar{c}_{23} = \bar{c}_{24} = 0.5$. Considering that $w_{2,22}^v$ is the largest weight coefficient and $\gamma = 0.5$, we expect the second experimental design to primarily minimize the variance of the parameter θ_2.

Table 4.1: Progression of parameter estimates and their inferences. Underlined items denote failed t-tests.

Iteration number (b)	Parameter estimate		t-score	Reference t-value	χ^2-score	Reference χ^2-value
1	θ_1	0.3137 $\pm$ 0.0169	3.731	2.920	0.807	5.991
	θ_2	0.1965 $\pm$ 0.2154	$\underline{0.183}$			
	θ_3	0.5605 $\pm$ 0.0415	2.714			
	θ_4	0.0513 $\pm$ 0.0083	$\underline{1.248}$			
2	θ_1	0.3081 $\pm$ 0.0142	18.726	1.860	5.970	15.507
	θ_2	0.1728 $\pm$ 0.0960	$\underline{1.556}$			
	θ_3	0.5468 $\pm$ 0.0394	11.998			
	θ_4	0.0489 $\pm$ 0.0105	4.029			
3	θ_1	0.3086 $\pm$ 0.0060	37.422	1.761	7.533	23.685
	θ_2	0.1737 $\pm$ 0.0641	1.988			
	θ_3	0.5479 $\pm$ 0.0146	27.495			
	θ_4	0.0491 $\pm$ 0.0026	13.796			
4	θ_1	0.3091 $\pm$ 0.0045	52.321	1.725	11.624	31.410
	θ_2	0.1847 $\pm$ 0.0591	2.382			
	θ_3	0.5470 $\pm$ 0.0128	32.543			
	θ_4	0.0489 $\pm$ 0.0029	12.807			

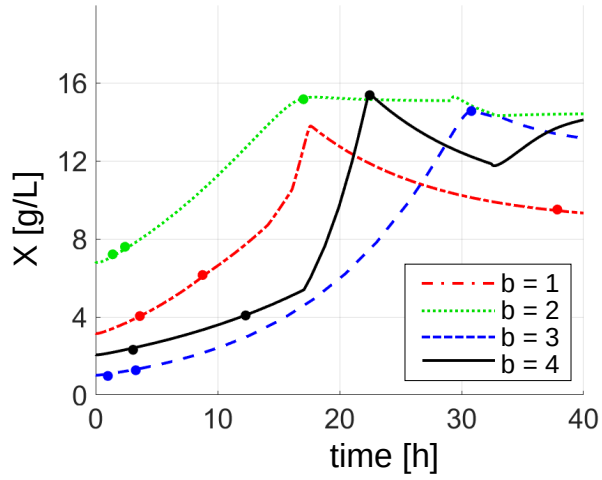


(a)

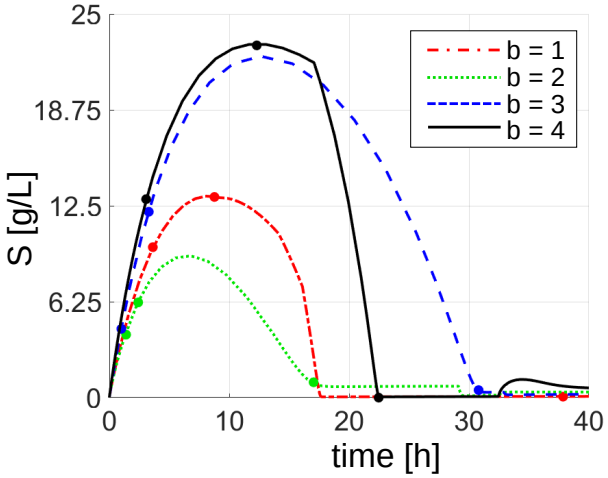


(b)

Figure 4.2: Trajectories of optimal experimental designs of (a) dilution factor and (b) substrate inlet concentration. Limits of y-axis correspond to allowed range of each input.



(a)



(b)

Figure 4.3: Trajectories of state variables where optimal experimental design is applied. Large dots correspond to sampling instants.

Table 4.2: Progression of variance matrix and variance weight coefficients

Iteration number (b)	Variance matrix	$\mathbf{w}_b^v$
1	$10^{-3} \times \begin{bmatrix} 0.381 & 3.268 & 0.731 & 0.137 \\ 3.268 & 62.109 & 1.107 & 0.069 \\ 0.731 & 1.107 & 2.304 & 0.431 \\ 0.137 & 0.069 & 0.431 & 0.091 \end{bmatrix}$	$[0,\ 0.839,\ 0,\ 0]^T$
2	$10^{-4} \times \begin{bmatrix} 0.509 & -0.206 & 1.157 & 0.300 \\ -0.206 & 23.203 & -4.564 & -1.446 \\ 1.576 & -4.546 & 3.905 & 0.923 \\ 0.300 & -1.448 & 0.923 & 0.277 \end{bmatrix}$	$[0,\ 0,\ 0,\ 0]^T$
3	$10^{-4} \times \begin{bmatrix} 0.148 & 0.714 & 0.192 & 0.007 \\ 0.714 & 16.607 & -0.790 & -0.477 \\ 0.192 & -0.790 & 0.863 & 0.072 \\ 0.007 & -0.477 & 0.072 & 0.028 \end{bmatrix}$	$[0,\ 0,\ 0,\ 0]^T$
4	$10^{-4} \times \begin{bmatrix} 0.080 & 0.277 & 0.151 & 0.013 \\ 0.277 & 13.815 & -0.533 & -0.480 \\ 0.151 & -0.533 & 0.679 & 0.070 \\ 0.013 & -0.480 & 0.070 & 0.033 \end{bmatrix}$	$[0,\ 0,\ 0,\ 0]^T$

Iteration #2

The solution to the second experimental design ϕ_2^* is shown in Figures 4.2 and 4.3. As expected, the variance of θ_2 was reduced (Table 4.2) and the t-score significantly increased. Moreover, the correlation index between the parameters θ_1 and θ_2 was minimized as well. Both aspects are illustrated in Figure 4.4 where the length of the approximate confidence region was considerably reduced along the y-axis and the skewness of the ellipsoid was reduced at the same time. However, values of the other correlation indices ($c_{2,13}, c_{2,14}, c_{2,34}$) remained largely the same. We can interpret this result that for each of the unchanged correlation indices $c_{2,13}, c_{2,14}, c_{2,34}$, one of the following two events occured. First, a correlation index can be decreased using some experimental designs; however, decreasing the other objectives $v_{2,22}$ and $c_{2,12}$ is a better way to minimize the overall objective function as defined by $\mathbf{w}_1^v$ and $\mathbf{w}_1^c$. Another possibility is that minimizing the correlation index is prohibited by the model structure and/or the experimental conditions. For example, we assumed that inlet substrate concentration S_{in} is static throughout the experiment, or that we can switch the value D only three times during the experiment. These constraints make the design space of ϕ much smaller, preventing one from reaching a state from which we obtain a measurement that possibly decreases a certain correlation index. For this moment, one cannot decide which of the two has occured for which parameter. One can discern between the two cases for each parameter only after an analysis of the next experiment is made. Regarding the design of the third experiment, no weight was imposed on the parameter variance ($\mathbf{w}_2^v = [0, 0, 0, 0]^T$) which makes the objective to minimize only the correlation indices $c_{3,13}, c_{3,14}, c_{3,24}$ and $c_{3,34}$. In other

words, our goal of achieving the desired level of variance is achieved, and the only remaining task is to reduce the correlation between the parameters. Comparing a new weight $\mathbf{w}_2^c$ to the former weight $\mathbf{w}_1^c$, the same number of positive weights were imposed for the experimental design. Although the weight coefficient $w_{2,12}^c$ was eliminated, a new coefficient $w_{2,24}^c$ was introduced. This may seem unexpected because we used the constraint $c_{2,24} \leq \bar{c}_{24} = 0.5$ in solving $\boldsymbol{\phi}_2^*$. This could happen because the values of $\mathbf{w}_b^v$ and $\mathbf{w}_b^c$ were evaluated based on the parameter estimate $\hat{\boldsymbol{\theta}}_b$. In other words, before obtaining the measurement $\mathbf{Y}_2$, one decides whether the experimental design $\boldsymbol{\phi}_2^*$ causes the constraint violation, based on the state prediction made from $\hat{\boldsymbol{\theta}}_1$. Because the parameter estimate $\hat{\boldsymbol{\theta}}_1$ is inaccurate, one erroneously decides that $\boldsymbol{\phi}_2^*$ does not cause a violation. After $\mathbf{Y}_2$ is measured and the new (and more accurate) estimate $\hat{\boldsymbol{\theta}}_2$ is made, the effects of the former experiments $\boldsymbol{\phi}_1^*$ and $\boldsymbol{\phi}_2^*$ are reevaluated according to $\hat{\boldsymbol{\theta}}_2$ and reach different conclusions for the correlation index $c_{2,24}$. These situations are undesirable to our process because in this case, the constraint (4.5) only decreases the search space of the experiment. Because θ_2 was the most inaccurate parameter among $\hat{\boldsymbol{\theta}}_2$, we speculate that the erroneous evaluation of $w_{1,24}^c$ was largely a result of the inaccurate estimate $\hat{\theta}_{2,2}$. We will briefly discuss how to address this issue in the concluding section.

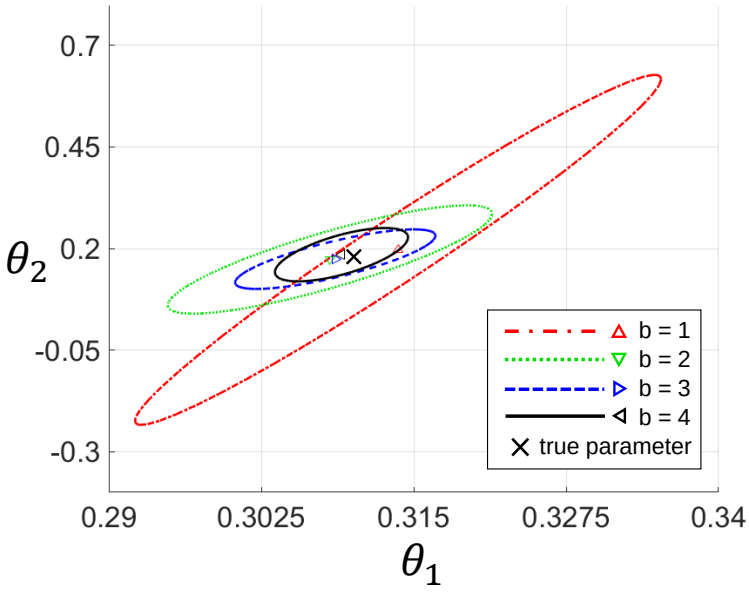


Figure 4.4: Joint confidence region between the parameters θ_1 and θ_2.

Table 4.3: Progression of correlation matrix and correlation weight coefficients

Iteration number (b)	Correlation matrix	$\mathbf{w}_b^c$
1	$\begin{bmatrix} 1 & 0.671 & 0.779 & 0.736 \\ 0.671 & 1 & 0.093 & 0.029 \\ 0.779 & 0.093 & 1 & 0.939 \\ 0.736 & 0.029 & 0.939 & 1 \end{bmatrix}$	$[0.445,\ 0.588,\ 0.538,\ 0,\ 0,\ 0.717]^T$
2	$\begin{bmatrix} 1 & -0.060 & 0.821 & 0.798 \\ -0.060 & 1 & -0.478 & -0.572 \\ 0.821 & -0.478 & 1 & 0.888 \\ 0.798 & -0.572 & 0.888 & 1 \end{bmatrix}$	$[0,\ 0.629,\ 0.608,\ 0,\ 0.235,\ 0.683]^T$
3	$\begin{bmatrix} 1 & 0.456 & 0.537 & 0.102 \\ 0.456 & 1 & -0.209 & -0.705 \\ 0.537 & -0.209 & 1 & 0.470 \\ 0.102 & -0.705 & 0.470 & 1 \end{bmatrix}$	$[0,\ 0.132,\ 0,\ 0,\ 0.497,\ 0]^T$
4	$\begin{bmatrix} 1 & 0.263 & 0.661 & 0.257 \\ 0.263 & 1 & -0.178 & -0.706 \\ 0.661 & -0.178 & 1 & 0.476 \\ 0.257 & -0.706 & 0.476 & 1 \end{bmatrix}$	$[0,\ 0.428,\ 0,\ 0,\ 0.499,\ 0]^T$

Iteration #3

A third optimal experiment $\boldsymbol{\phi}_3^*$ was implemented and $\hat{\boldsymbol{\theta}}(3)$ was obtained. The results of an analysis are summarized on the third rows of Tables 4.1–4.3. No significant decrease in the variances of individual parameters was achieved. However, a significant reduction in the correlation indices $c_{3,14}$ and $c_{3,34}$ was observed, as well as a moderate decrease in the correlation index $c_{3,13}$. Figure 4.5 depicts the confidence ellipsoid of parameters θ_1 and θ_4, where the skewness of the ellipsoid was relieved in the third iteration. However, this was at the cost of the index $c_{3,24}$ whose absolute value was increased from 0.572 to 0.705. Only two indices $c_{3,13}$ and $c_{3,24}$ violated the threshold $\bar{c} = 0.5$, from which we built our fourth experiment.

Iteration #4

In analyzing the statistics of the parameter estimate $\hat{\boldsymbol{\theta}}_4$ obtained from $\boldsymbol{\phi}_4^*$, we see little impact on relieving the parameter correlations, which was our goal for the experiment. Correlation indices $c_{k,13}$ and $c_{k,24}$ actually increased from 0.537 and 0.470 to 0.661 and 0.476, respectively. This is also indicated in the joint confidence regions in Figure 4.6, where little difference is observed in their shapes or sizes. Comparing the results of experiments $\boldsymbol{\phi}_3^*$ and $\boldsymbol{\phi}_4^*$, we conclude that we reached a state where little is expected from additional experiment(s). Therefore, we terminate the iteration at $k = 4$, gathering $\hat{\boldsymbol{\theta}}_4$ as the final parameter estimate. Statistics of the final estimate $\hat{\boldsymbol{\theta}}_4$ are given in the last rows of Tables 4.1, 4.2, and 4.3. All t-values of the parameters were below the reference t-value, and the lack-of-fit test showed that the measured data was successfully described by the model. All variances of the parameters were minimized under a desired variance of 0.01, and so were the four out of six correlation indices. Correlation indices that remained unsatisfied were the indices between $\theta_1 - \theta_3$ and $\theta_2 - \theta_4$. This indicates that the correlation exists between these two pairs owing to the model structure, which is difficult to decouple by means of experimental design.

4.4 Remarks on the choice of hyper parameters

- Weighting factors $w^v_{b,ii}$ and $w^c_{b,ij}$

In (4.1) and (4.2), There are two reasons for giving a zero weight weight of 0 for parameters and parameter pairs that do not exceed their reference values. First, one can actually calculate a level of variance and correlation below which is not required according to the end use of the model. It is more efficient to minimize the variance and correlation indices below the reference point rather than minimizing variance and correlation of all parameters. Moreover, minimizing the number of terms contained in the objective function makes it easier to interpret the calculation results.

- Relative weight between the variance and correlation γ

The implication of the value γ and its effect on MBDOE is clear. It is responsible for determining which of the two conflicting objective functions to put more emphasis on. For limiting cases where the value is 0, the MBDOE concentrates only on reducing the variance of the individual parameters without considering the correlation. On the other hand, the larger the gamma value, the smaller the priority on the variance and the greater the priority on the correlation indices. One important fact is that if the variance of a parameter is large, then the correlation indices calculated for this parameter estimates also lose their reliability. Ultimately, both variance and correlation should be minimized. However, in reality, it is a more efficient approach to minimize variance first and then minimize the related correlation values. This was naturally achieved in the previous simulation even if γ was fixed from beginning to the end. There is no guarantee that this

tendency will repeat all the time, so it is necessary to induce this tendency by adjusting γ. One can start with a very small γ, say 0.01, in order to concentrate on reducing the variances that are not sufficiently minimized by prior D-optimal batch. Then one can increase γ as the batch repeats, concentrating more on minimizing the correlation between the parameters.

4.5 Conclusion

An experimental design method for the parameter estimation of batch systems that combines anticorrelation criteria and sequential design is presented. The parameters of batch systems such as fed-batch bioreactors are generally highly correlated, and one should oftentimes apply both the sequential design and anticorrelation criteria in order to obtain a reliable set of parameter estimates. This study presented a method to utilize both approaches in an integrated way, as demonstrated in the case study. The case study showed that the resulting parameter estimates satisfied both the variance and the correlation. Moreover, by analyzing the progression of the weight coefficients, we could determine a point at which to terminate the iteration.

One undesirable occurrence in the case study was that of the occurrence of the weight coefficient $w^c_{2,24}$ as compared to the former coefficient $w^c_{1,24} = 0$. This incident can be detrimental to the performance of the algorithm (although it was not in our case), and therefore should be avoided if possible. One simple way of avoiding this problem is to use the weight coefficient $w^c_{b,ij} = 0$ whenever either of the variances $v_{b,ii}$ or $v_{b,jj}$ is not satisfied. This is based on the observation that the erroneous evaluation of $w^c_{1,24}$ was related to the invalid estimate of θ_2. In other words, we evaluate (and attempt to minimize) solely the correlation indices between the parameters that are individually well estimated.

Another question that can be asked is whether our scheme is applicable to a much larger system. When naively implemented on a large system, the objective function involves too many elements. In particular, the number of correlation indices in the objective function

can be up to $N_p(N_p - 1)/N_p$, which makes the objective function too complex and makes the resulting experimental design uninterpretable. Surprisingly, the same amendment proposed for solving the erroneous weight problem can be used here. By putting no (or little) weight on the parameters that are imprecise, one can focus each experimental design on minimizing a small number of indices. In addition, parameter subset selection methods such as those in [67, 69] can be used in order to exclude parameters that are essentially irrelevant to the model behavior.

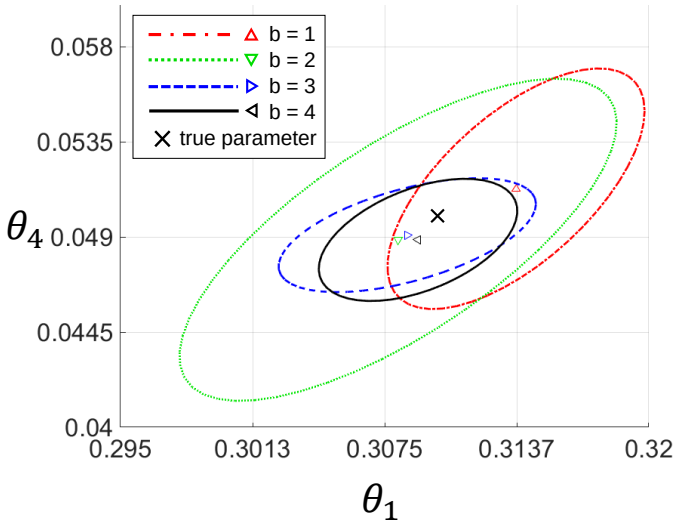
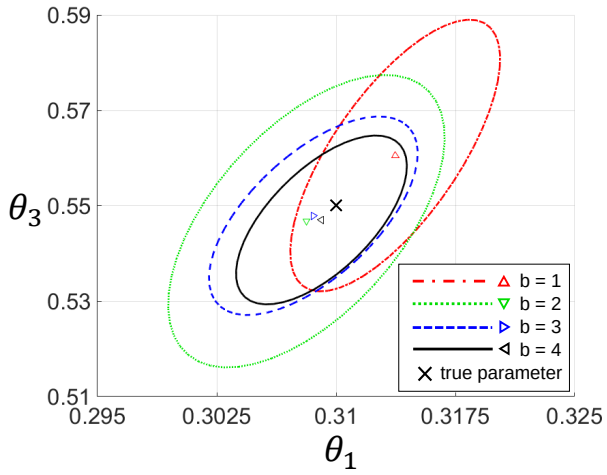
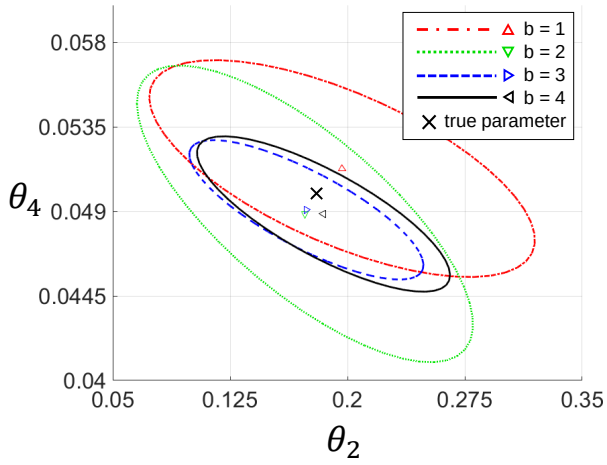


Figure 4.5: Joint confidence region between the parameters θ_1 and θ_4.



(a)



(b)

Figure 4.6: Joint confidence region between parameters (a) θ_1 and θ_3, (b) θ_2 and θ_4

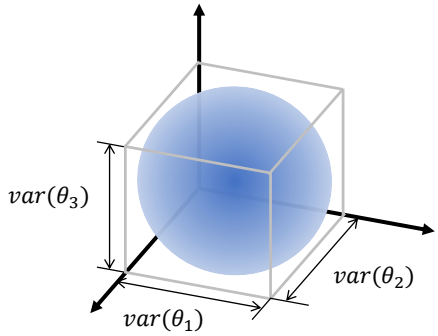
Chapter 5

Application to a microalgal fed-batch bioreactor

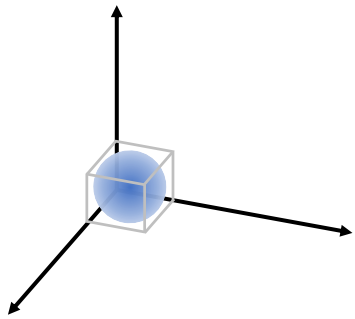
5.1 Necessity of the combined scheme

The statistic of the estimated N_p parameters is summarized in the Fisher's information matrix, which can be visualized by the hyper-ellipsoid of the N_p dimension in Figure $5.1(a)$. The D-optimality used as the objective function in Chapter **3** corresponds to the volume of this hyper-ellipsoid. In other words, MBDOE which uses D-optimality as the objective function, calculates an experiment that minimizes the volume of this ellipsoid. The resulting ellipsoid is expected to look like the ellipsoid shown in Figure $5.1(b)$. However, in most practical situations the resulting ellipsoid resembles to form shown in Figure $5.1(c)$. This is because, due to the model's structural characteristics, it is usually more advantageous to reduce the volume of ellipsoid by reducing the variance of specific parameters than to reduce it by reducing the variance of all parameters simultaneously. In both cases, the volume of the ellipsoid is the same, but the latter is worse in terms of parameter accuracy. The variance of the second parameter is unacceptably large, and also the correlation between the parameters is large. In order to fix this situation, we need the type of algorithm proposed in Chapter **4**. The variance and correlation index

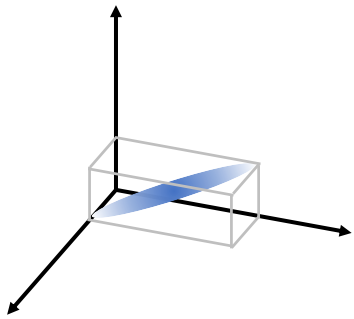
of each parameter, used as the objective function in scheme in Chapter 4, indicate the axial length and skewness of the ellipsoid, respectively. To summarize, each of the two proposed methods is imperfect, so it is necessary to use them in complementary sense.



(a)



(b)



(c)

Figure 5.1: (a) Hyperellipsoid representing the confidence region of parameters. (b) Ideal transformation of the confidence region. (c) Non-ideal transformation of the confidence region.

5.2 Overall scheme of the study

In the two previous chapters, we proposed two algorithms and demonstrated them using a relatively simple model. Both methods worked well, but in reality, most of the models used in practice are larger and more complex, so there is a necessity to validate the methods with more realistic models. Therefore, we have applied the methods to a 6 - state, 14 - parameter fed-batch bioreactor model to verify its usefulness. Given that different types of MBDOEs have different types of advantages and disadvantages, it is very important to decide which type of MBDOEs to use in which order. The first thing to consider in choosing the type of MBDOE is the dependence on the parameter initial estimates. We compared the performance of the off-line MBDOE and the on-line MBODE (both of full-sized design and reduced design) in Chapter **3** and confirmed the effect of on-line MBDOE that minimizes the effect of initial parameter inaccuracy. Therefore, we can conclude that it is best to operate the first batch using the on-line MBDOE. After the first batch has been completed, the accuracy of the estimated parameters is calculated. If the statistics of the parameter estimate is unsatisfactory, either on-line MBDOE or conventional off-line MBDOE can be performed again, using D-optimality criteria as an objective. Repeating this process comes to a point where the acceptable accuracy of the parameters is obtained. From this moment, the inaccuracy of a few remaining parameters and the correlation between the parameters are the most important tasks to be solved. We convert the scheme to successive complementary anti-correlation MBDOE at this moment and find the point at which we end the parameter estimation process by looking at the changes in

$\mathbf{w}^v$ and $\mathbf{w}^c$. The overall scheme is summarized in the Figure 5.2 on the next page.

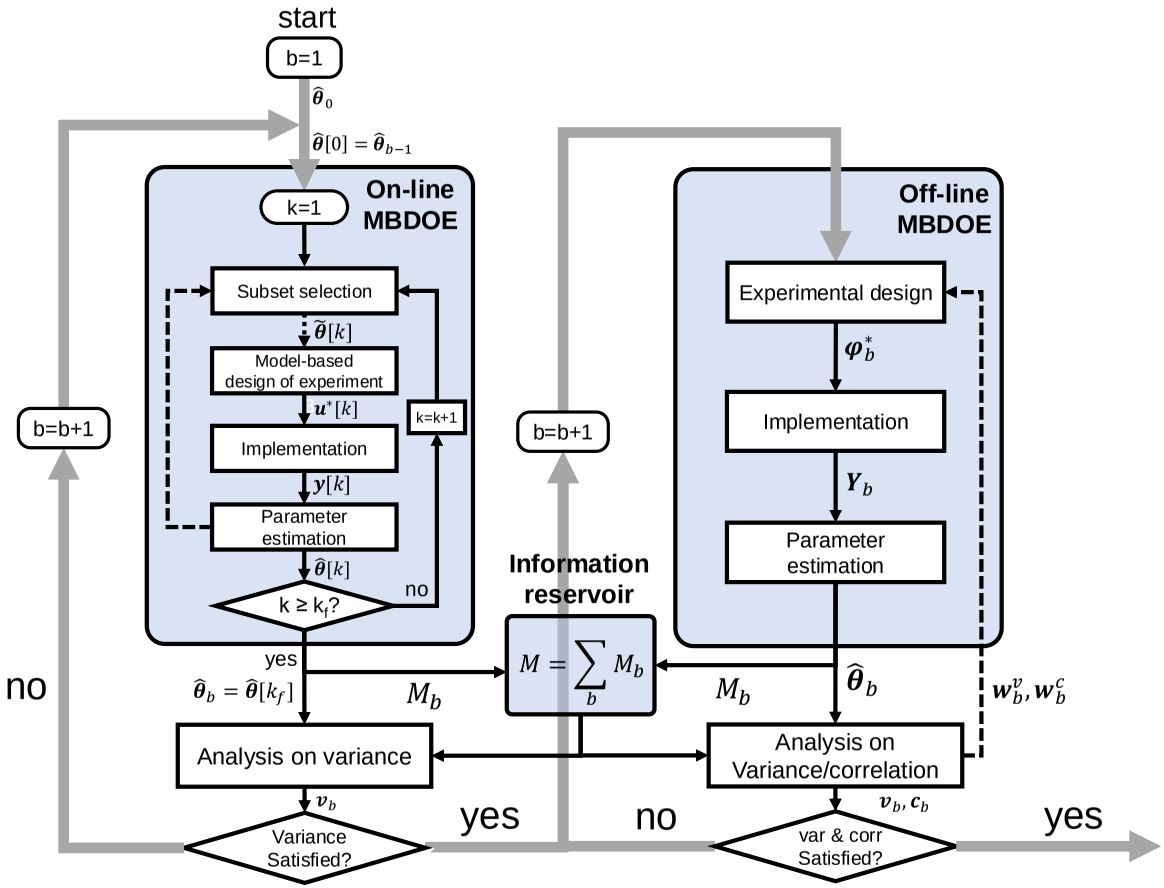


Figure 5.2: Overall scheme for integrating on-line and off-line MBDOEs

5.3 Model description

In this chapter, we simulate the fed-batch microalgal bioreactor model suggested by Yoo [66] referred in Chapter **2** earlier. There are two main reasons for choosing this model. First, this model has enough generality because it has the shape, size, and complexity that is typical of macroscopic bioprocess models commonly used in industry. Secondly, the model was established by the authors and collaborators from the development stage. Therefore, various simulation conditions such as initial state values and the magnitude of the measurement error can be set realistically. it is also advantageous to analyze the result and its implications from the practical point of view.

The model is shown in (5.1).

$$
\begin{aligned}
\frac{dX}{dt} &= \mu X - XD \\
\frac{dS_N}{dt} &= -\rho X + S_N^i \frac{u_N}{V} - S_N D \\
\frac{dS_C}{dt} &= -\frac{1}{Y_{XS}}\mu X - \frac{1}{Y_{LS}}\pi X + S_C^i \frac{u_C}{V} - S_C D \\
\frac{dQ}{dt} &= \rho X - \mu Q - QD \\
\frac{dL}{dt} &= \pi X - vL - LD \\
\frac{dV}{dt} &= u_N + u_C - f_0
\end{aligned}
$$

where

$$
\begin{aligned}
\mu &= \mu_m \left(1 - \frac{q_0}{q}\right)\left(1 - \frac{l_0}{l}\right)\left(\frac{S_2}{K_{S_2} + S_2}\right)\left(\frac{I}{K_I + I}\right) \\
\rho &= \rho_m \left(\frac{S_1}{K_{S_1} + S_1}\right)\left(\frac{q_m - q}{q_m - q_0}\right) \\
\pi &= \pi_m \left(\frac{S_2}{K_\pi + S_2}\right)(1-q)(1-l) \\
v &= v_m \left(\frac{K_v}{S_2 + K_v}\right)\left(1 - \frac{l_0}{l}\right) \\
l &= \frac{L}{X + Q + L} \\
q &= \frac{Q}{X + Q + L} \\
D &= \frac{u_N + u_C}{V}
\end{aligned}
\tag{5.1}
$$

The physical meanings of the parameters are given in Table 1. Their true values and lower and upper bounds for the estimation are also given. Parameter ranges and constraints are specified by inspecting each parameter. The initial parameter estimate $\hat{\boldsymbol{\theta}}[0]$ is randomly se-

lected such that it respects the given ranges and constraints. The batch termination time t_f is 300 hours and the time interval between the control/sampling instants T is 12 hours, making $k_f = 300/12 = 25$ and N_{sp}=24. Initial state vector $\mathbf{x}[0]$ is $[0.1, 0, 0, 0.5, 0.01, 2]$. The admissible ranges for each input variable are $u_N(in\ ml/h) \in [0, 10]$, $u_C(in\ ml/h) \in [0, 10]$ and $u_I(in\ \mu mol/m^2s) \in [0, 300]$. For the measurement noise Σ, values $[0.03, 0.001, 0.001, 0.005, 0.001, 0]$ is used. This reflects the actual error variances measurements of each states.

#	symbol	unit	meaning	range & constraints	true value	initial estimate
1	μ_m	$1/h$	maximum growth rate	0.01–0.044	0.0218	0.0281
2	q_0	g/g	minimum nitrogen quota for supporting growth	0–0.1, $q_0 < q_m$	0.008	0.0530
3	l_0	g/g	minimum lipid quota for supporting growth	0–0.1	0.001	0.0861
4	K_{S_2}	g/L	half saturation constant of carbon source for growth	0.0001–0.008	0.0008	0.0039
5	K_I	$\frac{\mu mol}{m^2 s}$	half saturation constant of light for growth	1–100	10.001	39.9562
6	ρ_m	$1/h$	maximum uptake rate	0.0355–0.142	0.071	0.1070
7	K_{S_1}	g/L	half saturation constant of nitrogen source for uptake	0–0.003	0.0003	0.0022
8	q_m	g/g	maximum quota of nitrogen above which uptake rate stops	0–1, $q_m > q_0$	0.5285	0.2829
9	π_m	$1/h$	maximum lipid production rate	0.107–0.428	0.214	0.2186
10	K_π	g/L	half saturation constant for oil production	5.413–541.3	54.13	85.7946
11	v_m	$1/h$	maximum lipid consumption rate	0.008–0.318	0.0159	0.0219
12	K_v	g/L	proportional constant of carbon source for lipid consumption	0.495-49.47	4.947	13.3333
13	Y_{xs}	g/g	yield coefficient of substrate to biomass	0.05–2	1.195	0.1367
14	Y_{ls}	g/g	yield coefficient of substrate to lipid	0.05–2	0.202	1.5221

Table 5.1: Model parameter description, values and ranges

5.4 Parameter subset selective on-line MBDOE

5.4.1 Simulation settings

As an indicator of design optimality $\mathcal{F}$, the log-determinant of FIM, i.e., D-optimality is used. In addition to its various advantageous properties [32], this is because the parameter subset selection method we use corresponds to finding the subset with the largest D-optimality criteria [67]. Prediction horizon H_p is set as 3. The sample experimental designs at each time instant $\boldsymbol{\phi}[k]$ are defined as grid points generated from the input range. The admissible ranges of input variables are divided into 3 equally spaced values, resulting in 9 samples. By comparing the orthogonal contribution to the sensitivity matrix S_k of each parameters, $N_r = 5$ parameters are selected. Here, N_r is equal to $N_y - 1$. This is because state variable V carries little informational value in parameter estimation because V can be directly calculated from input values u_N and u_C regardless of the parameter values. The columns of the sensitivity matrix S_k are normalized by multiplying them by nominal scale of each parameter, i.e., $\theta_{i,UB} - \theta_{i,LB}$ in order to avoid the dependency of scale of each parameters. The reduced-sized optimization (MBDOE) problem is formulated with subset $\tilde{\boldsymbol{\phi}}[k]$, and solved via interior-point algorithm. The method requires an initial point to begin with, and different initial points can lead to different local minima. We divide the solution space (experimental design space) into $3^{N_u} = 27$ grids, each of them performing as an initial point. All solutions obtained from the respective initial points are compared, and the solution with the least objective value is labeled as $\mathbf{U}^*[k]$. Additionally, solution to parameter estimation problem (3.6) is obtained by interior-point algorithm. The latest

estimate of parameter $\hat{\boldsymbol{\theta}}[k-1]$ is used as the starting point in the calculation of the next-step parameter estimate $\hat{\boldsymbol{\theta}}[k]$. After performing a batch experiment, one obtains Fisher's information matrix M_b from which you can calculate the variance of each parameter. This is compared with the reference variance value $\bar{v}_{ii}$ for each parameter, and if the parameter satisfying this criterion is 10 or more out of 14, one terminates the on-line MBDOE and proceeds to perform successive complementary MBDOE from the next batch. The reference variance values are set using lower and upper bounds of each parameter as in (5.2).

$$\begin{aligned} &\bar{v}_{ii} = 0.01(\theta_{i,UB} - \theta_{i,LB}) \\ &\text{for } i \in [1, N_p] \end{aligned} \tag{5.2}$$

All calculations were performed on MATLAB 2017b.

5.4.2 Result

Batch #1 : First on-line MBDOE

At each instant k, the significance of each parameters is represented by the orthogonal magnitudes previously defined in (3.1). Figure 5.3 shows the variation of these values over time. Y-axis in this graph shows the logarithm of the orthogonal magnitude value and each line corresponds to a parameter. Higher value indicates greater the importance of the corresponding parameter. The $N_r = 5$ most important parameters that are used for MBDOE of k-th time step is represented by larger markers. We can divide the total of 14 model parameters into three categories according to their relative importance. The most important parameters (parameters #1, 3, 6 and 11) were selected as a member of subset in large number of time instants. Other parameters (parameters #2, 4, 7, 8, 9, 13 and 14) were a member of subset at smaller number of instants. Three parameters (parameters #5, 10 and 12) had little significance at all time and were never selected. The most number of selection was 14 for parameter #11, out of k_f=25 possible instances. It can be confirmed by this that parameters show importance in limited time ranges, so it is efficient to concentrate the design objective only to a selected subset of parameters at each instant.

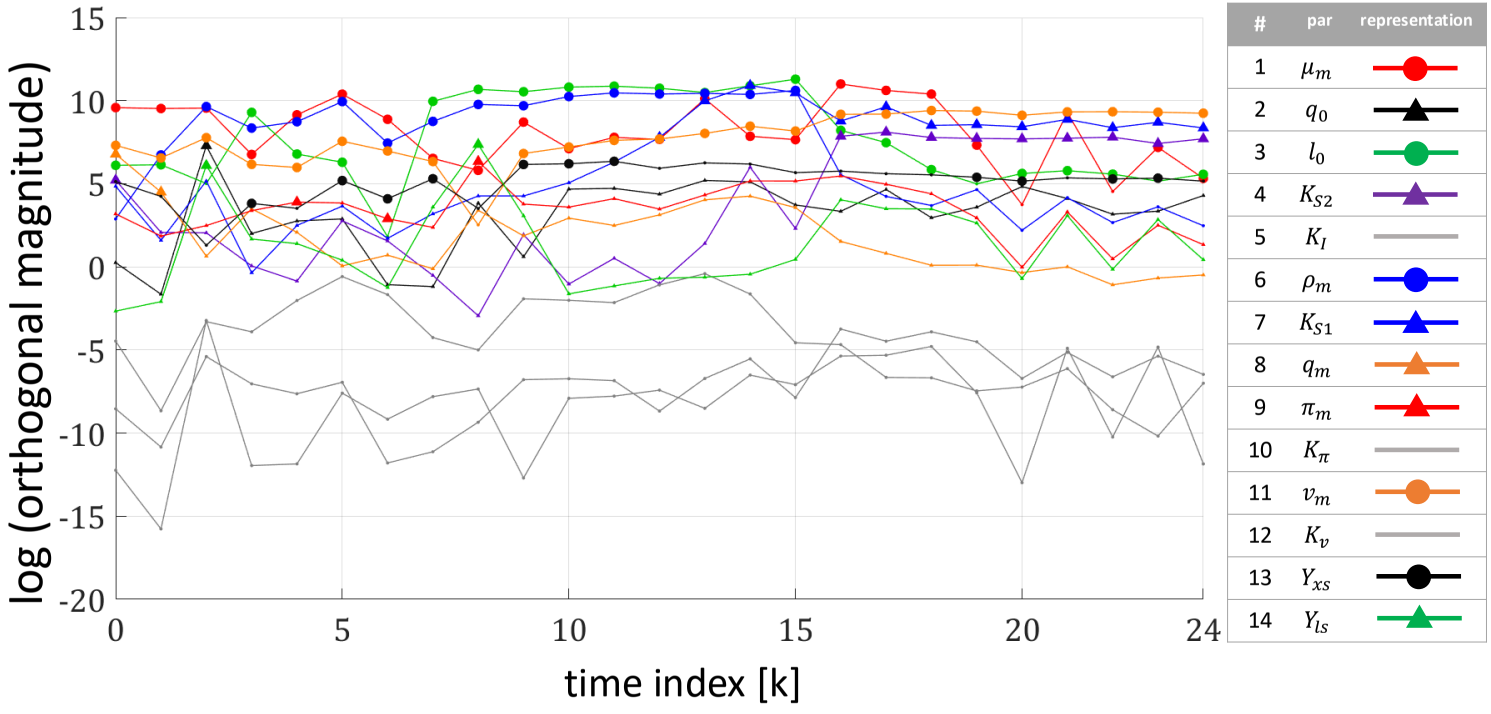


Figure 5.3: Progression of orthogonal magnitudes of parameters

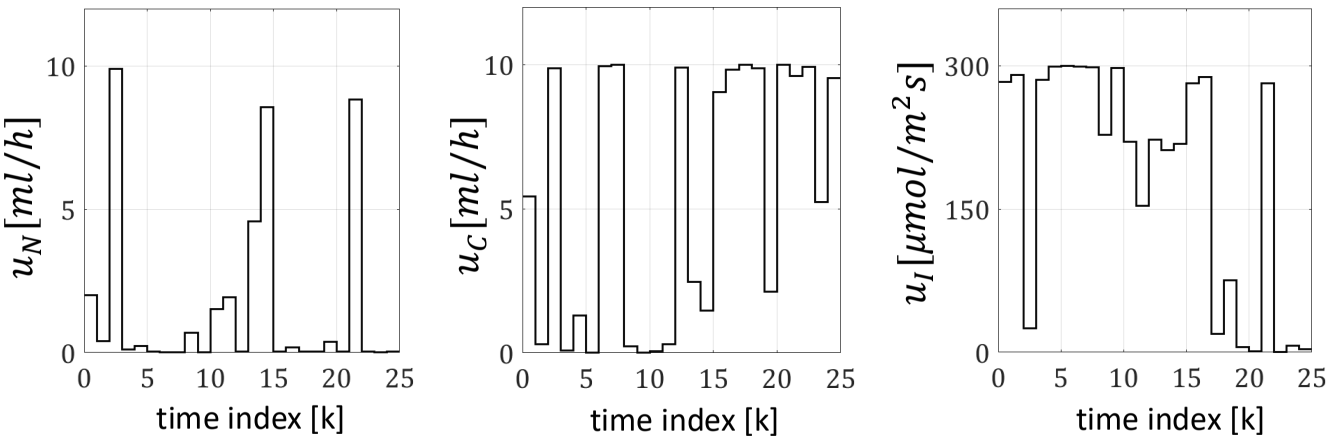


Figure 5.4: Optimal input trajectory obtained from batch #1

Figure 5.4 shows optimal input trajectories for each input, which is the product of reduced-sized MBDOE. Applying this input results the state trajectory shown in Figure 5.5. It also shows the measurements obtained from the state at each time instant. The progression of the parameter estimates calculated on-line are shown in Figure 5.6. We observe sharp changes of the parameter estimate values before $k = 10$. This is due to the fact that a small number of data, or small amount of error, significantly changes the residual function in the early phase. The fluctuating tendency of the parameter estimates seems to have affected the data trend of Figure 5.3 at early instants as well. Parameter estimate values converges after $k = 18$. When we calculate the residuals of the data using the final parameter estimates and actual parameters, the values are not significantly different — $2.91 * 10^5$ for the estimated parameter and $2.71 * 10^5$ for the real parameter. Comparing the final parameter estimate $\hat{\boldsymbol{\theta}}[k_f]$ to the true parameter value indicated as the red line, we see that some bias remains. This seems to be due to the fact that we use the previously estimate value $\hat{\boldsymbol{\theta}}[k-1]$ as the starting point for estimating the parameter at the current step, $\hat{\boldsymbol{\theta}}[k]$. In other words, the parameter estimate values are path-dependent due to the way we obtain it. This is a very important fact about the on-line MBDOE — On-line MBDOE is directly affected by the performance of the parameter estimation. This can be detrimental to the successful implementation of the scheme because the severe bias of the parameters can deteriorate the performance of the MBDOE as well. As a way to solve the bias problem of real-time parameter estimation, one can consider using multiple initial point for parameter re-estimation, just as we did for real-time MBDOE. The problem with this method is that the number of pa-

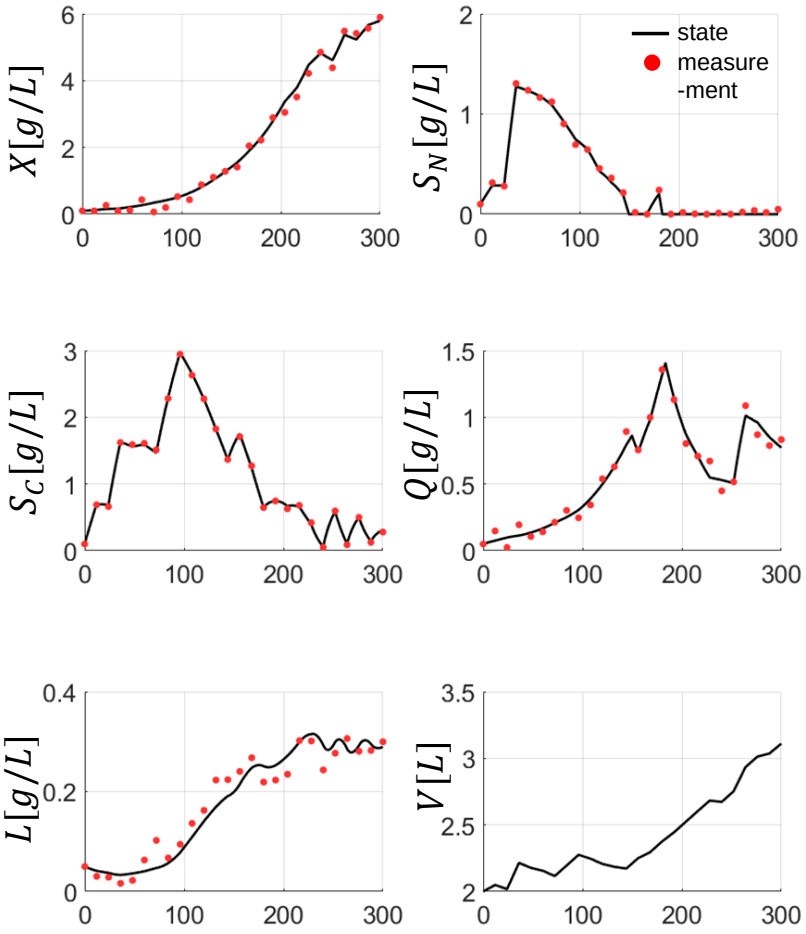


Figure 5.5: State trajectories obtained by applying optimal input from batch #1 and the measurements

rameters is very large, making the initial search space also large. In this case, real-time parameter estimation is unable to be performed smoothly on-line. One way to compensate for this is to perform real-time parameter estimation starting from the parameter estimate of the previous time step, and at the same time perform off-line parameter estimation using a much wider search space. The sudden change of parameter estimates in early batch stage is another practical problem shown in the simulation. One possible solution to this is to use as a buffer that collects as much data as possible before the start of the experiment. This buffer prevents the residual function from changing abruptly with an addition of small number of data. Alternatively, one can hold the parameter re-estimation process until a sufficient amount of data is collected in the early operation stage. In the preceding chapter, we mentioned that one of the reasons that we choose parameter subsets is that FIM may be too ill-conditioned when all parameters are considered. In this case, calculation of FIM and its norm causes large numerical errors, deteriorating the result of optimization computations. To study the effect of subset selection in solving this problem, the condition number of FIM defined with the subset parameters and the condition number of FIM defined with the original parameters is compared. Figure 5.7 shows that the condition number of the reduced FIM is much smaller at all time instant, relieving the numerical instability problem caused by ill-conditionedness of FIM.

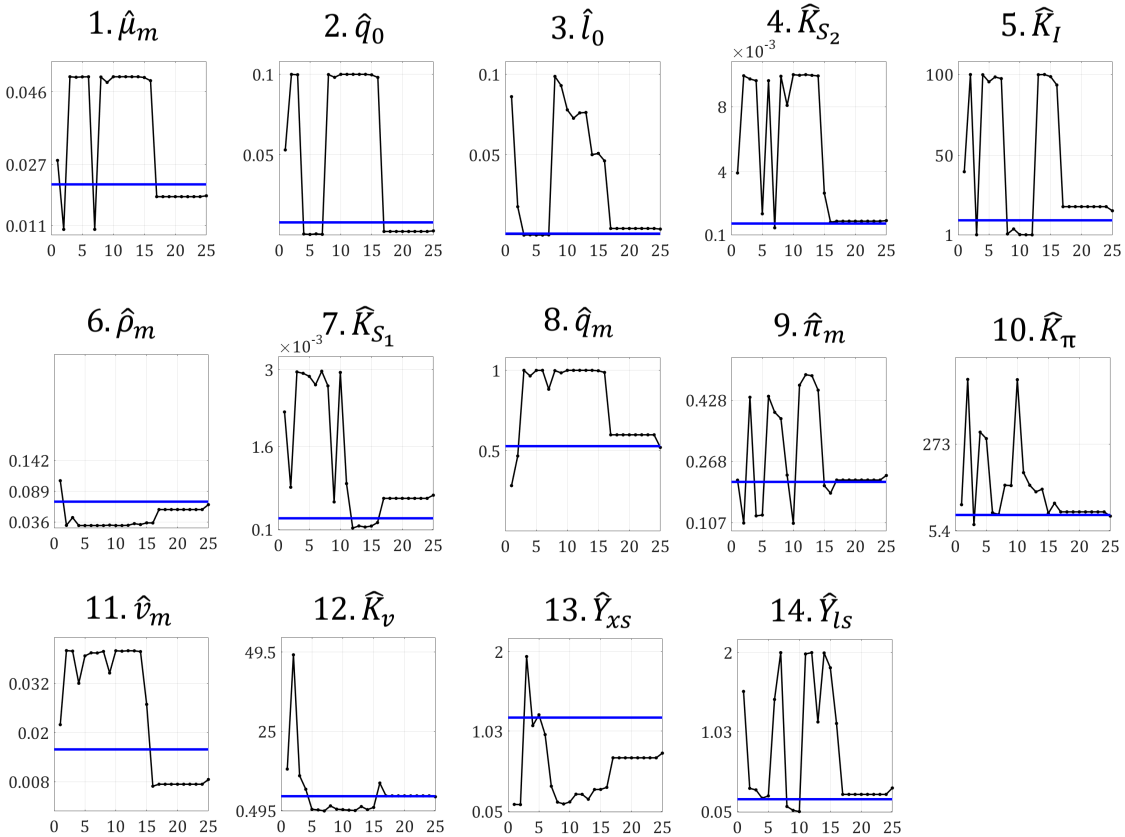


Figure 5.6: Progression of parameter estimate values in batch #1

Fast computation time is one of the constraints that must be satisfied in performing the on-line MBDOE. If the computation time in each step takes too long, the real-time information is reflected to the experiment in a delayed sense, deteriorating the efficiency of MBDOE scheme. In the suggested algorithm, there is an additional step compared to the existing on-line MBDOE, the process of subset selection. In order for this algorithm to work smoothly on-line, a total of three calculations(subset selection, MBDOE and parameter re-estimation) must be completed sufficiently fast. As shown in Figure 5.8, the time required for determining subset is very small compared to the time required for the other two tasks. In addition, the size of the sensitivity matrix S_k used for the MBDOE calculation becomes very small compared to the full-parameter case. This and the lowered condition number makes the MBDOE calculation step much faster for the reduced-sized MBDOE. The longest time took for all 3 calculations was 938 seconds, which is about 1.3% of the sampling interval $T = 12h$. In a nutshell, the computation time is not a problem for the proposed algorithm. However, in this study, we did not consider the time required for obtaining the measurement values. In an actual implementation, if no in-line sensor is utilized, the most time-consuming step for the on-line scheme should be the time for obtaining the measurement. If the time required for measurement is considerable, the formulation of the objective function should be modified to a form that considers the time delay.

The variance of the parameter estimates obtained from the first batch, is shown in the first row of Figure 5.9. Among 14 parameters, all parameters except for 2 parameters had a sufficiently small value compared to the reference value $\bar{v}_{ii}$. According to the prede-

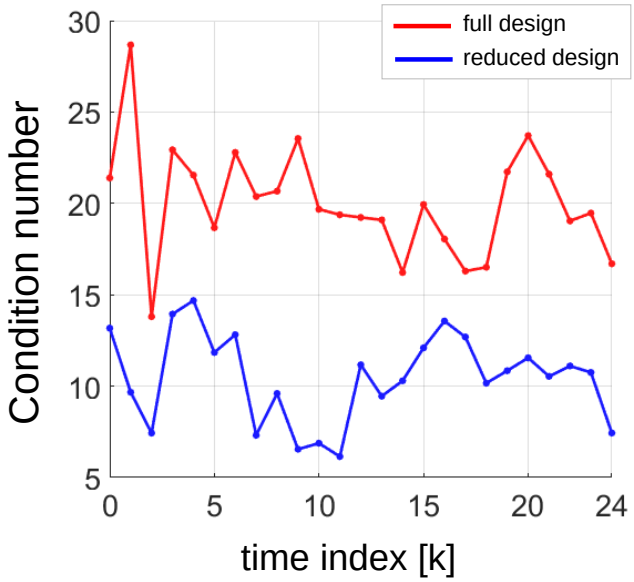


Figure 5.7: Comparison of condition numbers of Fisher's information matrix for reduced- and full design case

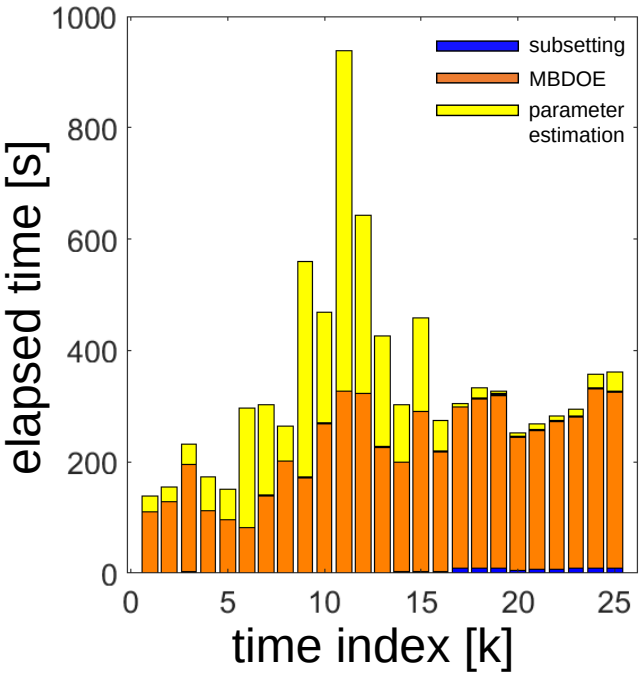


Figure 5.8: Elapsed time for calculation of each step in operating batch #1

termined criterion, it is determined that the remaining process of parameter estimation is performed with the successive complementary anti-correlation MBDOE.

5.5 Successive complementary anti-correlation MBDOE

5.5.1 Simulation settings

The initial state values are fixed and assumed to be known as $[0.1, 0, 0, 0.5, 0.01, 2]$, as in the previous batch. Each input is a time-dependent variable, and is control vector parameterized by 7 design parameters. The input is piecewise constant, characterized by $N_{sw} = 3$ control-switching instants. A total of $N_u(N_{sw} + N_{sw} + 1) = 21$ variables are used for characterizing all three inputs. Measurements for all state variables are made at three sampling instants, making the length of the design vector $\boldsymbol{\phi}_b$ to be 24. In addition, there are additional constraints between the sampling instants and sampling instants that there should be at least $1h$ difference between each instant. The MBDOE calculation was calculated using the interior point method with multiple starting points. We have created starting points according to the following rules.

- Switching instants for all 3 inputs can be either [1, 2, 3], [75, 150, 225] or [297, 298, 299].
- Controllable input u_N for all time instants can be either 0.001, 5 or 10.
- Controllable input u_C for all time instants can be either 0.001, 5 or 10.
- Controllable input u_I for all time instants can be either 0.001, 150 or 300.
- Sampling instants $\mathbf{t}^{sp}$ are fixed as [100, 200, 300].

In total, $3^4 = 81$ initial points were generated, and the solutions from each initial point were compared. The one with the least ob-

jective value was chosen as the final solution. For the desired parameter correlation values, $\bar{c}_{ij} = 0.7$ was chosen for all ${}_{14}C_2$ elements. γ=1 was used in for determining objective function (4.3), indicating that we put approximately the same emphasis on reducing the weights and on reducing the correlations. Parameter estimation was also performed by the interior-point method, providing the parameter estimate of the previous instant as its initial estimate. Statistics of the parameter estimates calculated after each iteration were analyzed using the latest evaluation of the information matrix, namely $M_b = M(\boldsymbol{\phi}_1^*, \boldsymbol{\phi}_2^*, ..., \boldsymbol{\phi}_b^*; \hat{\boldsymbol{\theta}}_b)$. All calculations were performed in MATLAB R2019a.

5.5.2 Result

Figure 5.9 shows the diagonal values $v_{b,ii}$ of variance matrices according to the batch index b. The red horizontal line represents the level at which the variance $v_{b,ii}$ of the parameter equals to the reference value $\bar{v}_{ii}$. If the bar goes higher than this line, the weight $w_{b,ii}^v > 0$ is given to the variance of the parameter θ_i. Figure 5.10 shows the comparison of the size of the correlation index to the reference value for 14 different parameter pairs selected from a total of ${}_{14}C_2 = 91$ parameter pairs. As in Figure 5.9, the weight $w_{b,ij}^c > 0$ is imposed on the parameter pair (i, j) when the blue bar cross over the red baseline in Figure 5.10. The parameter pairs shown in Figure 5.10 is the pairs that have exceeded the reference correlation at least once from $b = 1$ through $b = 4$. Correlation values of the remaining 77 pairs have never exceeded the reference value once. Figure 5.11 shows a graphical representation of the correlation index matrix.

Here, the correlation indices selected as the objective function are indicated by red squares, and the ones selected in a previous batch are indicated by orange squares.

The optimal input trajectory calculated by each successive complementary MBDOE is shown in Figure 5.12. Figure 5.12 shows the trajectories of state variables obtained by applying the optimal input trajectory as well as the instances of sampling. The sum of the weights calculated using FIM_b after each batch is shown in Figure 5.13.

Batch #2 : First successive complementary MBDOE

The optimal input trajectory and the final parameter estimate from the first batch $\hat{\boldsymbol{\theta}}[k_f] = \hat{\boldsymbol{\theta}}_1$ are used to Fisher's information matrix M_1. Variance and correlation indices are obtained from M_1, and are used to define weight values $w^v_{1,ii}$ and $w^c_{,ij}$ shown in the first row of Figures 5.9 and 5.10, respectively. As mentioned earlier, variances were found to be lower than their reference values except for the two parameters (parameters #10 and #12). Variances of those two parameters and the correlation indices for the seven parameter pairs were included in the objective function for MBDOE of the second batch experiment. The result of the first successive complementary MBDOE $\boldsymbol{\phi}^*_2$ was executed and parameter estimate $\hat{\boldsymbol{\theta}}_2$ was obtained using the measurement $\mathbf{Y}_2$. Little change of value was observed in the change of the parameter estimates $\hat{\boldsymbol{\theta}}_2$ compared to $\hat{\boldsymbol{\theta}}_1$. The re-evaluated value of the cumulative information matrix $M_1(\boldsymbol{\phi}^*_1; \hat{\boldsymbol{\theta}}_2) + M_2(\boldsymbol{\phi}^*_2; \hat{\boldsymbol{\theta}}_2)$ was used to calculate $v_{2,ii}$ and $c_{2,ij}$. As shown in the second rows in Figures 5.9 and 5.10 and the second matrix in Figure 5.11, variances of

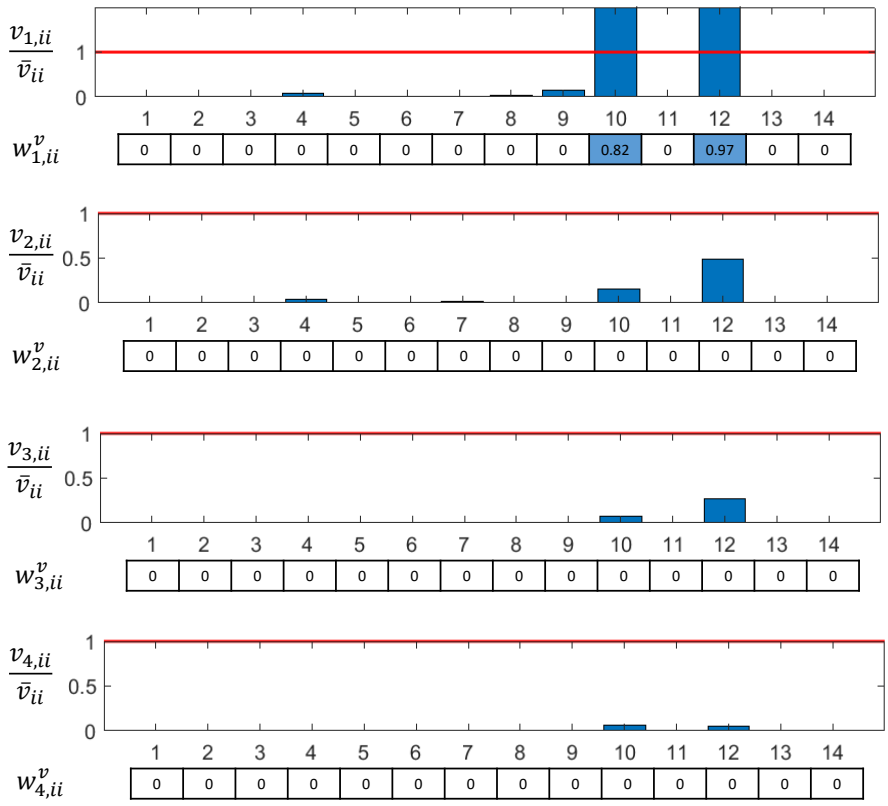


Figure 5.9: Change of variances and variance-weights over batches #1 through #4

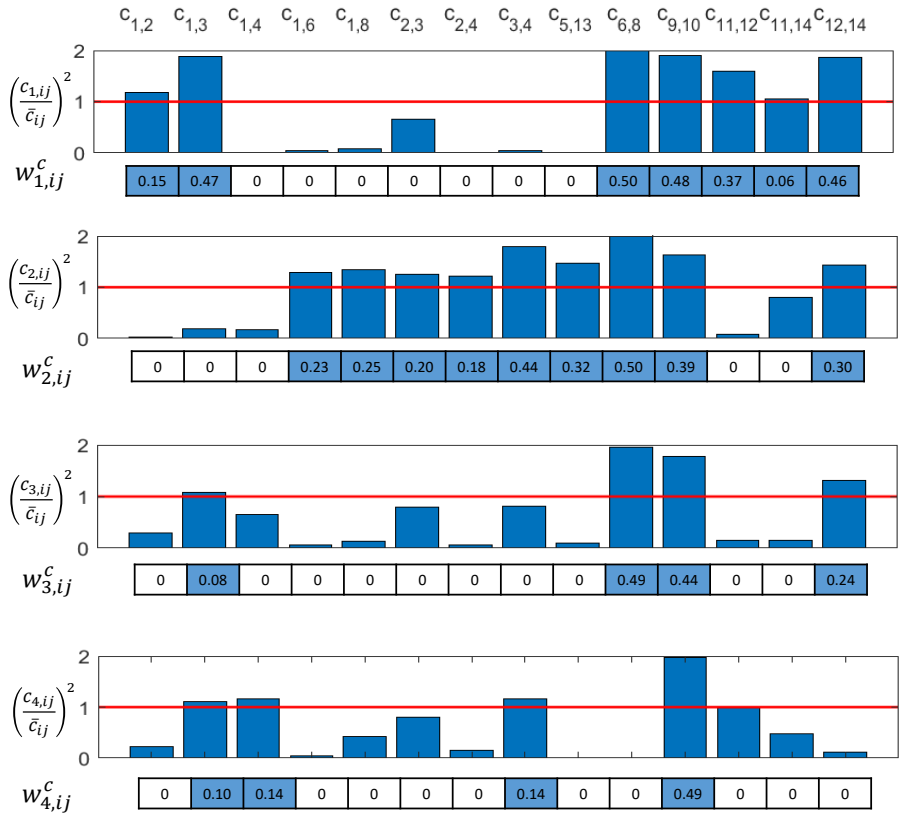


Figure 5.10: Change of correlation indices and correlation-weights over batches #1 through #4 for selected parameter pairs

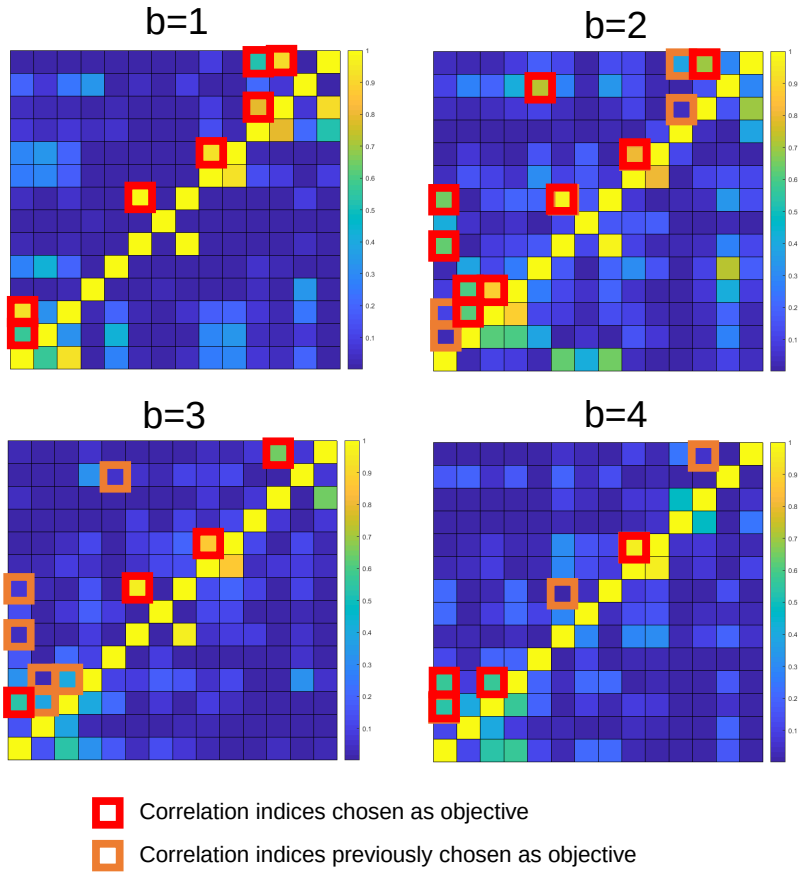
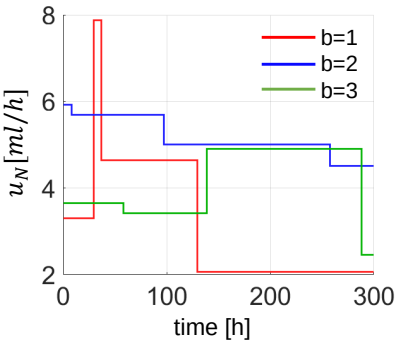


Figure 5.11: Change of correlation indices over batches #1 through #4

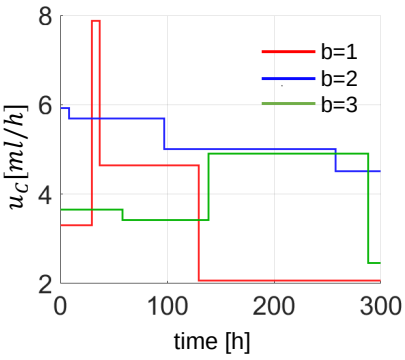
the two parameters were definitely below the reference value $\bar{v}_i i$. We succeeded in reducing the correlation values of 4 out of the 7 objective parameter pairs. However, an unexpected increase of correlation values for 6 parameter pairs was observed, which was below the reference value in the initial design. These increased correlation values were reflected in defining w_{ij}^c in designing $\boldsymbol{\phi}_3^*$.

Batch #3 and #4 : Second and third successive complementary MBDOEs

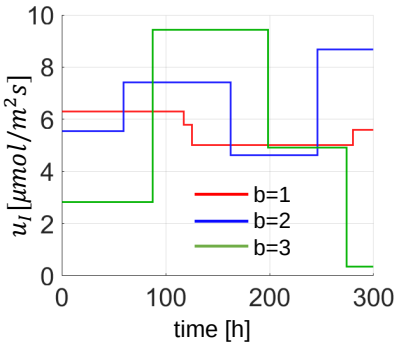
Using the resultant optimal design $\boldsymbol{\phi}_3^*$, $\mathbf{Y}_3$, $\hat{\boldsymbol{\theta}}_3$, M_3, $v_{3,ii}$, $c_{3,ij}$ were calculated in sequence. This time, the problematic correlation indices have shown significant reduction, leaving 4 correlation indices still above the reference value. The objective function for the fourth batch was designed aiming for reduction of those 4 variables. The sequence of MBDOE, implementation, and the analyses was performed as the same way as before. Comparing the correlation indices $c_{4,ij}$ to $c_{3,ij}$, two of the four correlations were relived, however at the cost of two other enlarged correlation indices. In the case of correlation index $c_{b,9,10}$, its value remained fairly large since the first simulation until $b = 4$, and it is speculated that it is the model structure itself that causes this the most. The total sum of the weight is still decreasing, but we terminated the iteration at $b = 4$ because the termination criteria by the sum of total weight (< 1) is satisfied. As a result, we obtained 14 estimated parameters with satisfactory values of variances and acceptable values of correlation indices. As a result of applying the two algorithms presented in the previous chapters to a larger-sized model, it is shown that both schemes work effectively together.



(a)



(b)



(c)

Figure 5.12: Optimal input trajectories for batches #2 thorugh #4

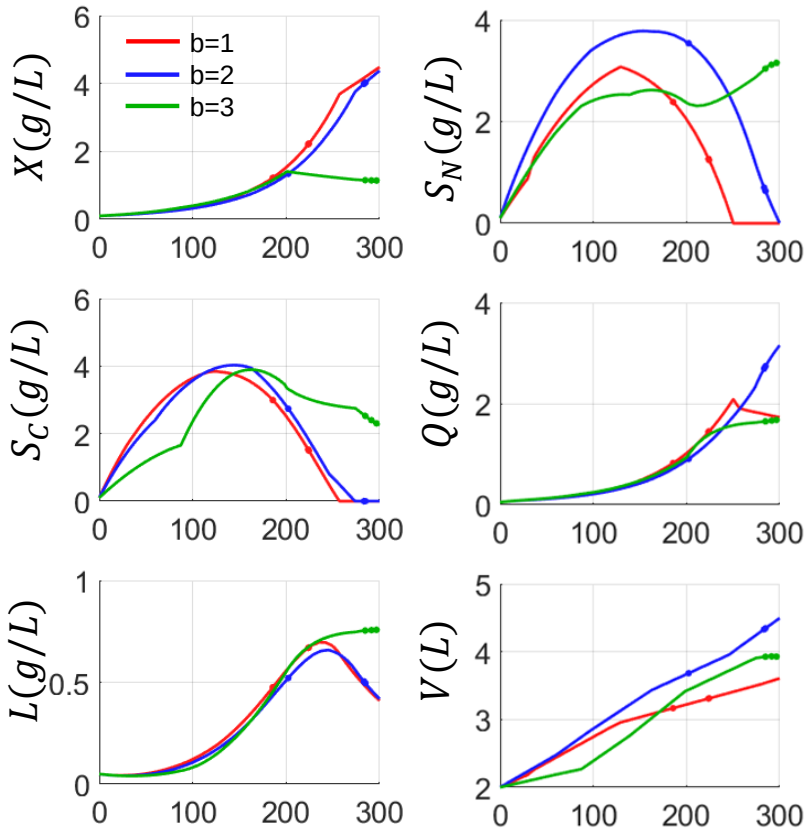


Figure 5.13: State trajectories for batches #2 through #4

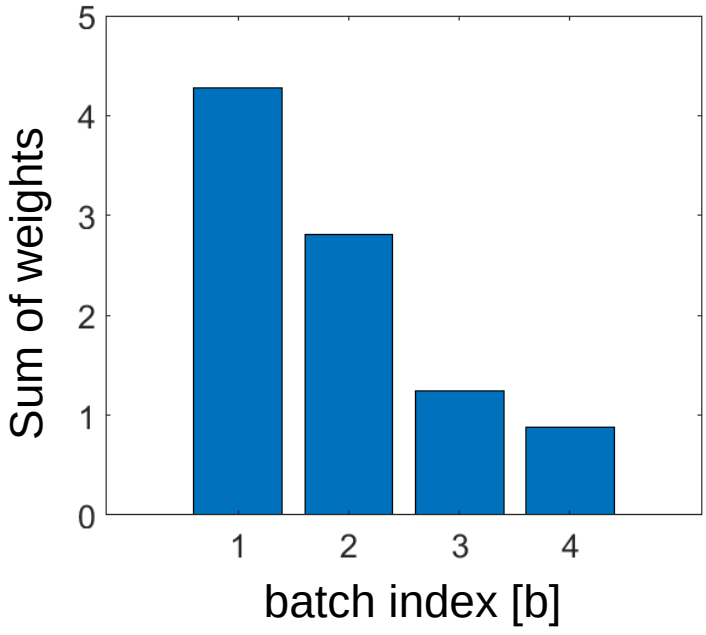


Figure 5.14: Sum of weight values for batches #2 through #4

5.6 Comparison to the D-optimal-only case

Previously in Chapters **5.1**, we explained why one should use MBDOE with two different objective functions in complementary sense. In this chapter, we compare the parameter estimation performance of the combined scheme presented in Chapter **5.2** to the case where only the D-optimal MBDOE is repeatedly used. First, in Figure 5.15, we see that bigger D-optimalities (i.e., the volumes of the confidence region hyper-ellipsoid) have achieved for iterative D-optimal case. However, for variances and correlation indices of individual parameters, the performance of two schemes is reversed. We compared the variances of the parameter #10 and #12, which had shown the largest variances after implementation of the first batch in Chapter **5.3**. Figure 5.16 shows that the variance of parameter 10 decreases sharply when the combined scheme is used, while the variance of the D-optimality case remains large. In the case of parameter #12, the value of the combined scheme is very small from the second arrangement. However, when only the D-optimality is used, the value has decreased after 4th batch. Similar differences can be observed for the parameter correlation index. The progression of the correlation indices between the two parameter pairs (#1, #2) and (#12, #14) are shown in Figure 5.17 (a) and (b). In both cases, the decrease of correlation indices for the combined scheme case is larger than in the D-optimal case, and the sum of the squares of all 91 correlation indices is compared in 5.17 (c). In conclusion, while the iterative D-optimal design is advantageous for the minimizing the overall confidence region, we observe that the combined scheme presented in Chapter **5** shows advantage in improving statistics of individual parameters.

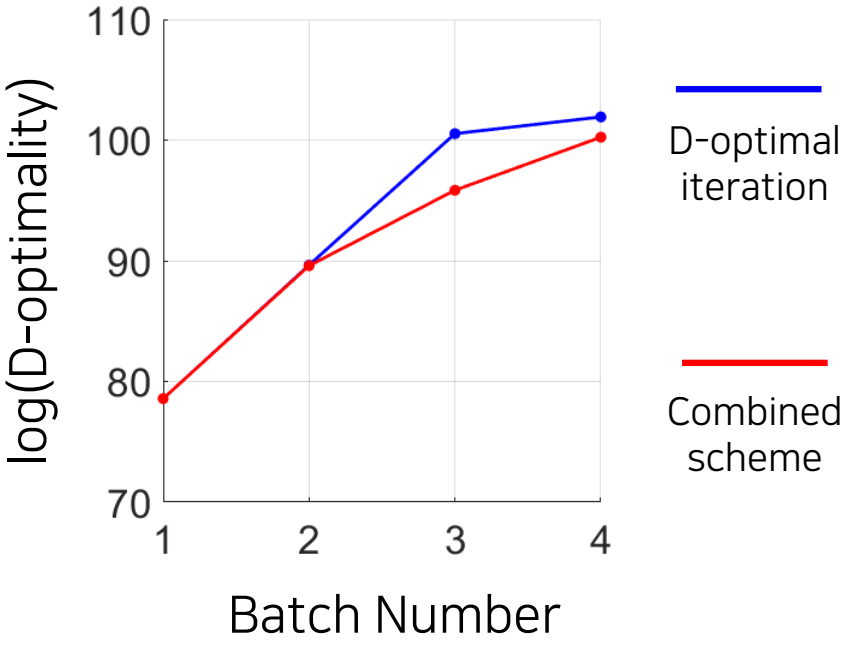
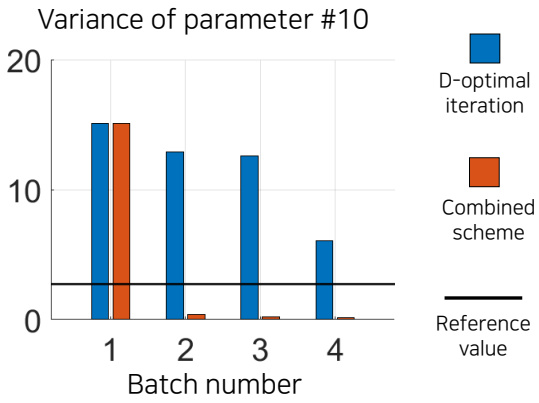
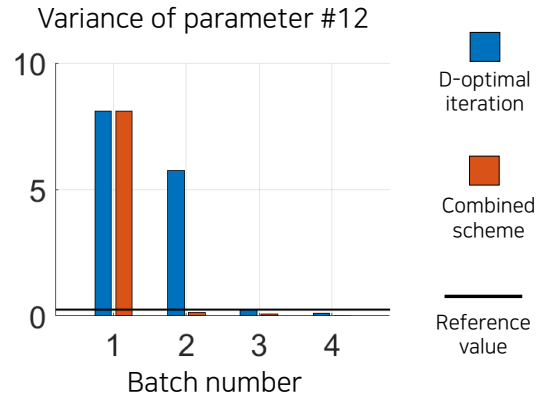


Figure 5.15: Comparison of D-optimality values of iterative D-optimal design case(blue) and the case using the combined scheme(red).

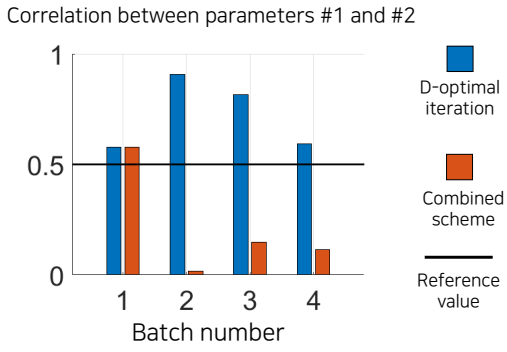


(a)

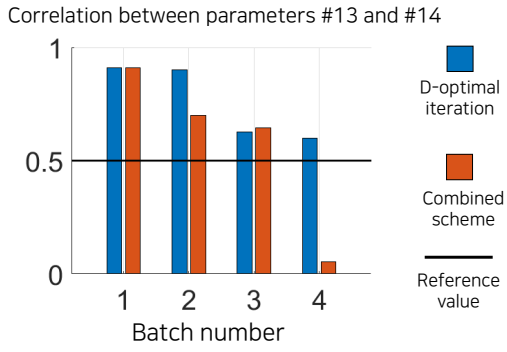


(b)

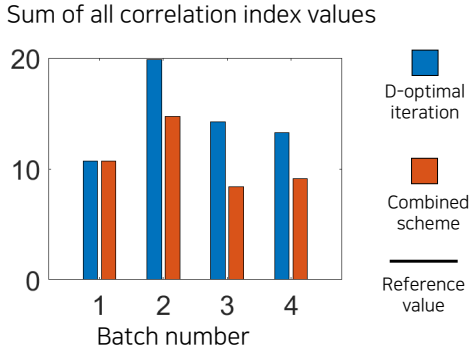
Figure 5.16: (a) Progression of the variance of the parameter #10. (b) Progression of the variance of the parameter #12.



(a)



(b)



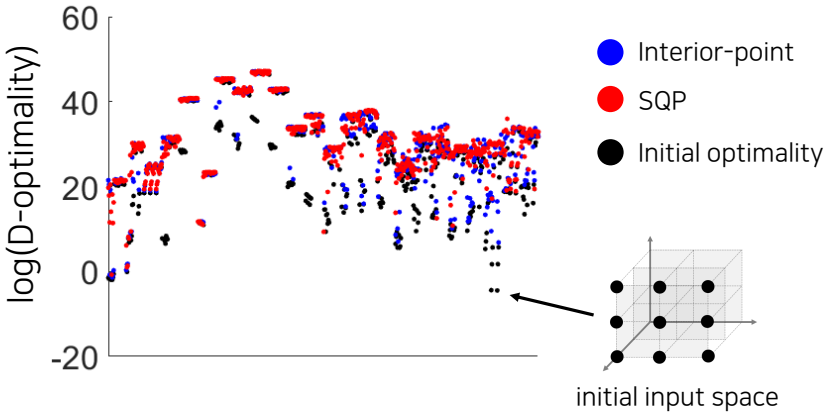
(c)

Figure 5.17: (a) Progression of the correlation index between the parameters #1 and #2. (b) Progression of the correlation index between the parameters #12 and #14. (c) Progression of the sum of squares of all correlation indices.

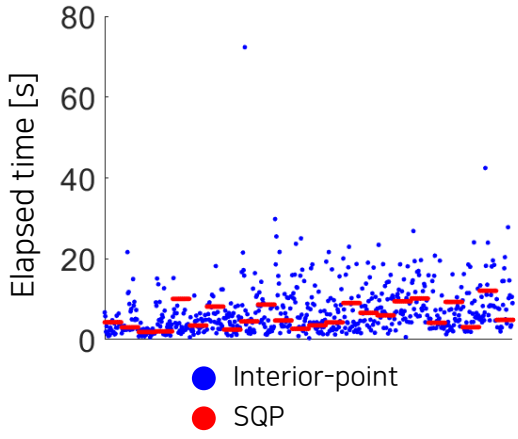
5.7 Remarks

5.7.1 Choice of the solution method

In actual implementing the proposed algorithm, the calculation result of the MBDOE can be different depending on the solution method. In addition, the performance of the proposed MBDOE scheme can also be influenced by the convergence performance and the computation speed of the algorithm. Therefore, it is necessary to find the optimal solver by comparing the performance of solution methods in solving the proposed MBDOE. To do this, we first solved the online reduced MBDOE in Chapter $\mathbf{5.4}$ using two representative optimization solvers, the interior-point method and the SQP method. At each time step, the $3^3 = 27$ initial points were given. Since there are 24 time steps to solve MBDOE, making a total number of initial points(i.e., total number of solving MBDOE) to be 648. For each of these initial points, we compared the improvement of objective function from the initial point, and the elapsed time to reach the solution, as shown in Figures 5.19 (a) and (b). Their average values are shown in the Table 5.2.



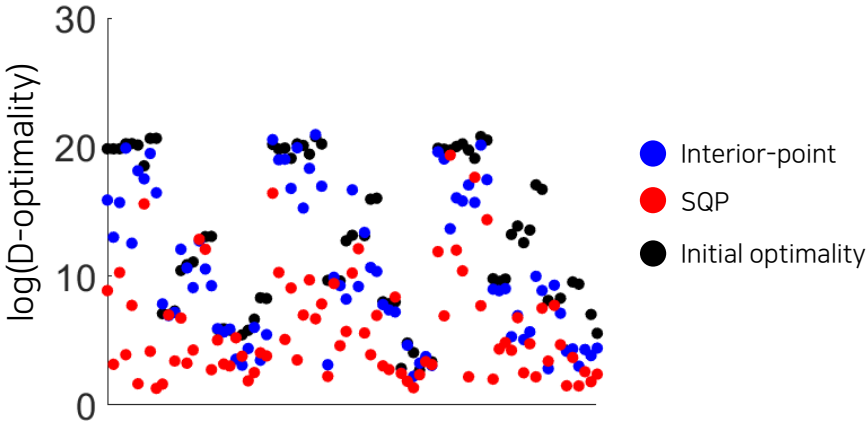
(a)



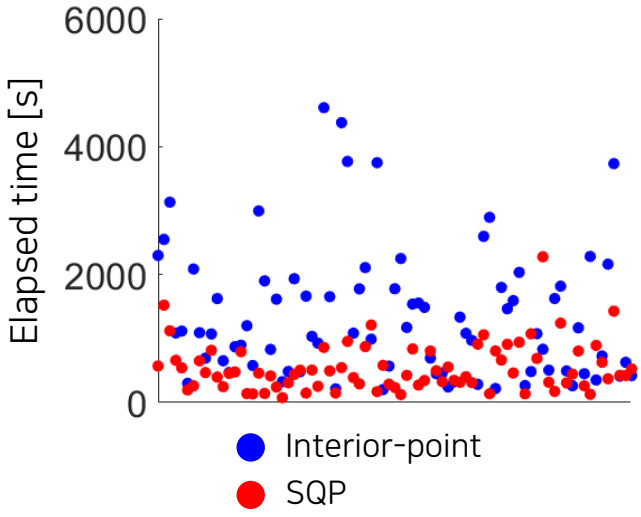
(b)

Figure 5.18: (a) Comparison of the optimization(maximization) performance of on-line reduced MBDOE by interior-point method and SQP. (b) Comparison of the computation time for solving on-line reduced MBDOE by interior-point method and SQP.

We confirmed that SQP is superior method in terms of both optimization performance and computation time. The same analysis is done for successive complementary MBDOEs. In this case, we compared the performance of the two solvers for the 81 initial points used in the first successive complementary MBDOE calculation in Chapter 5.5. As shown in table 5.3, the SQP method has better optimization performance and also, shorter computation time for this case as well. Therefore, we recommend SQP as the solution method for solving the proposed MBDOEs.



(a)



(b)

Figure 5.19: (a) Comparison of the optimization(maximization) performance of successive complementary MBDOE by interior-point method and SQP. (b) Comparison of the computation time for solving successive complementary MBDOE by interior-point method and SQP.

Method	Interior-point	SQP
Improvement of optimality	4.16	5.23
Computation time [s]	7.03	5.78

Table 5.2: Comparison of the two solution methods for solving reduced online MBDOE

Method	Interior-point	SQP
Improvement of optimality	2.26	6.77
Computation time [s]	1345	543

Table 5.3: Comparison of the two solution methods for solving reduced successive complementary MBDOE

Nomenclature

α	Ratio of the reference variance value $\bar{v}_{ii}$ with regard to the nominal parameter magnitude θ_i^{nom}
$\mathbf{f}$	Dynamic equations of the states vector
$\mathbf{h}$	State-output relation function
$\mathbf{x}$	State variables vector
$\mathbf{y}$	Output variables vector
$\boldsymbol{\epsilon}$	Measurement error vector
$\boldsymbol{\phi}$	Design variables vector
$\boldsymbol{\theta}$	Model parameters vector
γ	Relative weight parameter between the parameter variance weights and correlation weights
$\mathcal{F}$	A matrix-to-scalar function measuring the MBDOE optimality
μ	Specific growth rate
π	Lipid product formation rate
π_m	Maximum lipid production rate
ρ	Nitrogen substrate consumption rate
ρ_m	Maximum uptake rate

Σ	Covariance matrix for the measurement error
σ_{ij}	Covariance between the measurement y_i and y_j
b	Batch number index
C	Correlation matrix
C_b	Correlation matrix obtained from the cumulative information matrix M_b
c_{ij}	Correlation index between parameters θ_i and θ_j
D	Dilution factor for bioreactor feed
D_{eff}	D-efficiency, ratio of the logarithms of D-optimality values between different MBDOEs
H_p	Prediction horizon used for constructing M_k
k	Discrete time index
K_I	Half saturation constant of light for growth
K_v	Proportional constant of carbon source for lipid consumption
K_π	Half saturation constant for oil production
K_{S1}	Half saturation constant of nitrogen source for uptake
K_{S2}	Half saturation constant of carbon source for growth
L	Intracellular lipid concentration
l	Mass portion of lipid product inside the cell
l_0	Minimum lipid quota for supporting growth

M	Fisher's information matrix
M_k	Fisher's information matrix composed from S_{k+1}, ..., S_{k+H_p}
m_p	Orthogonal magnitude value for the parameter θ_p
mu_m	Maximum growth rate
N_P	Number of parameters, of dimension of the parameter vector
N_r	Number of subset parameters
N_ϕ	Dimension of the design variables
N_{sp}	Number of sampling instants during a batch operation
N_x	Dimension of the state variables
Q	Intracellular nitrogen concentration
q	Mass portion of nitrogen inside the cell
q_m	Maximum quota of nitrogen above which uptake rate stops
q_o	Minimum nitrogen quota for supporting growth
S	Substrate concentration in the medium
S_C	Carbon(glucose) substrate concentration inside the bioreactor
S_C^i	Concentration of the carbon substrate in the inlet feed
S_N	Nitrogen(glycine) substrate concentration inside the bioreactor
S_N^i	Concentration of the nitrogen substrate in the inlet feed
S_{in}	Substrate concentration of the inlet feed

T	Time difference between sampling instants for on-line MB-DOE
t	Continuous time index
u_C	Carbon source inlet feedrate
u_I	Illumination intensity
u_N	Nitrogen source inlet feedrate
V	Variance matrix
V	Volume of the bioreactor medium
v	Intracellular lipid consumption rate
v_m	Maximum lipid consumption rate
V_b	Variance matrix obtained from the cumulative information matrix M_b
v_{ii}	Variance value for parameter θ_i
v_{ij}	Covariance value between parameters θ_i and θ_j
X	Biomass concentration in the medium
Y_b	Measurement vector obtained from the batch b
Y_{ls}	Yield coefficient of substrate to lipid
Y_{xs}	Yield coefficient of substrate to biomass
$\bar{c}_{ij}$	Reference(threshold) correlation index for the parameter pair θ_i and θ_j

$\bar{v}_{ii}$	Reference(threshold) variance value for the parameter θ_i
$\mathbf{c}_b$	Vector of correlation indices obtained from the cumulative information matrix M_b
$\mathbf{r}_i$	i-th row of the sensitivity matrix S_k
$\mathbf{s}_p^{(k)}$	Projected vector of the sensitivity vector $\mathbf{s}_p$ at k-th iteration
$\mathbf{s}_p$	p-th column vector of the sensitivity matrix
$\mathbf{t}_{sp}$	Sampling instants
$\mathbf{t}_{sw}$	Control-switching instants
$\mathbf{u}$	Time-varying input variables
$\mathbf{U}[k]$	Array of time-varying input variables from time instants k through $(k + H_p - 1)$
$\mathbf{v}_b$	Vector of parameter variances obtained from the cumulative information matrix M_b
$\mathbf{w}$	Time-invariant input variables
$\mathbf{w}_b^c$	Vector of correlation weight coefficients calculated after batch b
$\mathbf{w}_b^v$	Vector of variance weight coefficients calculated after batch b
$\mathbf{x}(0)$	Initial state variables
$\mathbf{y}[k]$	Measurements obtained at time instant k
$\boldsymbol{\phi}[k]$	Vector of design variables, to be determined at time instant k
$\boldsymbol{\phi}^*$	Optimal experimental design calculated by MBDOE

$\boldsymbol{\phi}_{LB}$ Lower bounds for the design variables $\boldsymbol{\phi}$

$\boldsymbol{\phi}_{UB}$ Upper bounds for the design variables $\boldsymbol{\phi}$

$\boldsymbol{\theta}_{LB}$ Lower bounds for the parameters $\boldsymbol{\theta}$

$\boldsymbol{\theta}_{UB}$ Upper bounds for the parameters $\boldsymbol{\theta}$

$\hat{\mathbf{y}}[k]$ Model-predicted values for the measurements obtained at time instant k

$\hat{\boldsymbol{\theta}}$ Parameter estimate

$\hat{\boldsymbol{\theta}}[k]$ Real-time parameter estimate vector, updated at time instant k

θ_i i-th parameter

$\tilde{\boldsymbol{\theta}}[k]$ Parameter subset selected at time instant k

S_k Sensitivity matrix with regard to the measurement $\mathbf{y}[k]$

S_t Sensitivity matrix with regard to the measurement $\mathbf{y}_t$

$w^c_{b,ij}$ Weight coefficient given for parameter pair θ_i and θ_j calculated after batch b

$w^v_{b,ii}$ Weight coefficient given for parameter θ_i calculated after batch b

Bibliography

[1] D. Rippin, "Batch process systems engineering: a retrospective and prospective review," *Computers & chemical engineering*, vol. 17, pp. S1–S13, 1993.

[2] P. Bakonyi, N. Nemestóthy, V. Simon, and K. Bélafi-Bakó, "Review on the start-up experiences of continuous fermentative hydrogen producing bioreactors," *Renewable and Sustainable Energy Reviews*, vol. 40, pp. 806–813, 2014.

[3] J.-S. Chang and Y.-C. Lin, "Fed-batch bioreactor strategies for microbial decolorization of azo dye using a pseudomonasluteola strain," *Biotechnology progress*, vol. 16, no. 6, pp. 979–985, 2000.

[4] S. Dhir, K. J. Morrow Jr, R. R. Rhinehart, and T. Wiesner, "Dynamic optimization of hybridoma growth in a fed-batch bioreactor," *Biotechnology and Bioengineering*, vol. 67, no. 2, pp. 197–205, 2000.

[5] P. L. van den Bosch, O. C. van Beusekom, C. J. Buisman, and A. J. Janssen, "Sulfide oxidation at halo-alkaline conditions in a fed-batch bioreactor," *Biotechnology and bioengineering*, vol. 97, no. 5, pp. 1053–1063, 2007.

[6] U. Rau, L. Nguyen, H. Roeper, H. Koch, and S. Lang, "Fed-batch bioreactor production of mannosylerythritol lipids secreted by pseudozyma aphidis," *Applied microbiology and biotechnology*, vol. 68, no. 5, pp. 607–613, 2005.

[7] H. Leuenberger, "New trends in the production of pharmaceutical granules: the classical batch concept and the problem of scale-up," *European Journal of Pharmaceutics and Biopharmaceutics*, vol. 52, no. 3, pp. 279–288, 2001.

[8] T. M. S. Attia, X. L. Hu, and D. Q. Yin, "Synthesized magnetic nanoparticles coated zeolite for the adsorption of pharmaceutical compounds from aqueous solution using batch and column studies," *Chemosphere*, vol. 93, no. 9, pp. 2076–2085, 2013.

[9] N. T. Abdel-Ghani and S. H. Hussein, "Determination of ambroxol hydrochloride in pure solutions and some of its pharmaceutical preparations under batch and fia conditions," *Il Farmaco*, vol. 58, no. 8, pp. 581–589, 2003.

[10] A. Costa, "Hybrid genetic optimization for solving the batch-scheduling problem in a pharmaceutical industry," *Computers & Industrial Engineering*, vol. 79, pp. 130–147, 2015.

[11] H. Leuenberger and G. Betz, "Granulation process control—production of pharmaceutical granules: the classical batch concept and the problem of scale-up," in *Handbook of Powder Technology*, vol. 11, pp. 705–733, Elsevier, 2007.

[12] C. Gentric, F. Pla, M. Latifi, and J. Corriou, "Optimization and non-linear control of a batch emulsion polymerization reactor," *Chemical engineering journal*, vol. 75, no. 1, pp. 31–46, 1999.

[13] P. Nomikos and J. F. MacGregor, "Monitoring batch processes using multiway principal component analysis," *AIChE Journal*, vol. 40, no. 8, pp. 1361–1375, 1994.

[14] T. Peterson, E. Hernandez, Y. Arkun, and F. Schork, "A nonlinear dmc algorithm and its application to a semibatch polymerization reactor," *Chemical Engineering Science*, vol. 47, no. 4, pp. 737–753, 1992.

[15] D. J. Arriola, E. M. Carnahan, P. D. Hustad, R. L. Kuhlman, and T. T. Wenzel, "Catalytic production of olefin block copolymers via chain shuttling polymerization," *Science*, vol. 312, no. 5774, pp. 714–719, 2006.

[16] R. Zeman and N. Amundson, "Continuous polymerization models—part ii: Batch reactor polymerization," *Chemical Engineering Science*, vol. 20, no. 7, pp. 637–664, 1965.

[17] J.-H. Kim, M. Chen, N. Kishida, and R. Sudo, "Integrated real-time control strategy for nitrogen removal in swine wastewater treatment using sequencing batch reactors," *Water Research*, vol. 38, no. 14-15, pp. 3340–3348, 2004.

[18] M. Sheikhzadeh, M. Trifkovic, and S. Rohani, "Real-time optimal control of an anti-solvent isothermal semi-batch crystallization process," *Chemical Engineering Science*, vol. 63, no. 3, pp. 829–839, 2008.

[19] B. Srinivasan and D. Bonvin, "Real-time optimization of batch processes by tracking the necessary conditions of optimality," *Industrial & engineering chemistry research*, vol. 46, no. 2, pp. 492–504, 2007.

[20] A. Singh, J. F. Forbes, P. Vermeer, and S. Woo, "Model-based real-time optimization of automotive gasoline blending operations," *Journal of process control*, vol. 10, no. 1, pp. 43–58, 2000.

[21] M. Titica, D. Dochain, and M. Guay, "Adaptive extremum seeking control of fed-batch bioreactors," *European Journal of Control*, vol. 9, no. 6, pp. 618–631, 2003.

[22] P. Cougnon, D. Dochain, M. Guay, and M. Perrier, "On-line optimization of fedbatch bioreactors by adaptive extremum seeking control," *Journal of Process Control*, vol. 21, no. 10, pp. 1526–1532, 2011.

[23] J. H. Lee and K. S. Lee, "Iterative learning control applied to batch processes: An overview," *Control Engineering Practice*, vol. 15, no. 10, pp. 1306–1318, 2007.

[24] K. S. Lee, I.-S. Chin, H. J. Lee, and J. H. Lee, "Model predictive control technique combined with iterative learning for batch processes," *AIChE Journal*, vol. 45, no. 10, pp. 2175–2187, 1999.

[25] Y. Wang, D. Zhou, and F. Gao, "Iterative learning model predictive control for multi-phase batch processes," *Journal of Process Control*, vol. 18, no. 6, pp. 543–557, 2008.

[26] Y. Tian, J. Zhang, and J. Morris, "Modeling and optimal control of a batch polymerization reactor using a hybrid stacked recurrent neural network model," *Industrial & engineering chemistry research*, vol. 40, no. 21, pp. 4525–4535, 2001.

[27] J. Rawlings and W. Ray, "The modeling of batch and continuous emulsion polymerization reactors. part i: Model formulation and sensitivity to parameters," *Polymer Engineering & Science*, vol. 28, no. 5, pp. 237–256, 1988.

[28] F. Galvanin, A. Boschiero, M. Barolo, and F. Bezzo, "Model-based design of experiments in the presence of continuous measurement systems," *Industrial & Engineering Chemistry Research*, vol. 50, no. 4, pp. 2167–2175, 2011.

[29] Y. Zhang and T. F. Edgar, "Pca combined model-based design of experiments (doe) criteria for differential and algebraic system parameter estimation," *Industrial & Engineering Chemistry Research*, vol. 47, no. 20, pp. 7772–7783, 2008.

[30] J. Fricke, K. Pohlmann, N. A. Jonescheit, A. Ellert, B. Joksch, and R. Luttmann, "Designing a fully automated multi-bioreactor plant for fast doe optimization of pharmaceutical protein production," *Biotechnology journal*, vol. 8, no. 6, pp. 738–747, 2013.

[31] F. Galvanin, M. Barolo, G. Pannocchia, and F. Bezzo, "Online model-based redesign of experiments with erratic models: a disturbance estimation approach," *Computers & Chemical Engineering*, vol. 42, pp. 138–151, 2012.

[32] V. V. Fedorov and P. Hackl, *Model-oriented design of experiments*, vol. 125. Springer Science & Business Media, 2012.

[33] L. Ljung, "System identification," *Wiley Encyclopedia of Electrical and Electronics Engineering*, pp. 1–19, 1999.

[34] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, and J. Timmer, "Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood," *Bioinformatics*, vol. 25, no. 15, pp. 1923–1929, 2009.

[35] A. Holmberg, "On the practical identifiability of microbial growth models incorporating michaelis-menten type nonlinearities," *Mathematical Biosciences*, vol. 62, no. 1, pp. 23–43, 1982.

[36] E. Walter, Y. Lecourtier, J. Happel, and J.-Y. Kao, "Identifiability and distinguishability of fundamental parameters in catalytic methanation," *AIChE journal*, vol. 32, no. 8, pp. 1360–1366, 1986.

[37] E. Walter and L. Pronzato, "On the identifiability and distinguishability of nonlinear parametric models," *Mathematics and computers in simulation*, vol. 42, no. 2-3, pp. 125–134, 1996.

[38] M. Chappell and K. Godfrey, "Structural identifiability of nonlinear systems: Application to a batch reactor," *IFAC Proceedings Volumes*, vol. 24, no. 3, pp. 865–870, 1991.

[39] P. Li and Q. D. Vu, "Identification of parameter correlations for parameter estimation in dynamic biological models," *BMC systems biology*, vol. 7, no. 1, p. 91, 2013.

[40] A. Raue, V. Becker, U. Klingmüller, and J. Timmer, "Identifiability and observability analysis for experimental design in nonlinear dynamical models," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 20, no. 4, p. 045105, 2010.

[41] S. Asprey and S. Macchietto, "Designing robust optimal dynamic experiments," *Journal of Process Control*, vol. 12, no. 4, pp. 545–556, 2002.

[42] W. DuMouchel and B. Jones, "A simple bayesian modification of d-optimal designs to reduce dependence on an assumed model," *Technometrics*, vol. 36, no. 1, pp. 37–47, 1994.

[43] X. Huan and Y. M. Marzouk, "Simulation-based optimal bayesian experimental design for nonlinear systems," *Journal of Computational Physics*, vol. 232, no. 1, pp. 288–317, 2013.

[44] G. Franceschini and S. Macchietto, "Model-based design of experiments for parameter precision: State of the art," *Chemical Engineering Science*, vol. 63, no. 19, pp. 4846–4872, 2008.

[45] J. Stigter, D. Vries, and K. Keesman, "On adaptive optimal input design," in *2003 European Control Conference (ECC)*, pp. 393–398, IEEE, 2003.

[46] J. Stigter, D. Vries, and K. Keesman, "On adaptive optimal input design: a bioreactor case study," *AIChE journal*, vol. 52, no. 9, pp. 3290–3296, 2006.

[47] F. Galvanin, M. Barolo, and F. Bezzo, "Online model-based redesign of experiments for parameter estimation in dynamic systems," *Industrial & Engineering Chemistry Research*, vol. 48, no. 9, pp. 4415–4427, 2009.

[48] B. R. Jayasankar, B. Huang, and A. Ben-Zvi, "Receding horizon experiment design with application in sofc parameter estimation," *IFAC Proceedings Volumes*, vol. 43, no. 5, pp. 541–546, 2010.

[49] Y. Zhu and B. Huang, "Constrained receding-horizon experiment design and parameter estimation in the presence of poor initial conditions," *AIChE Journal*, vol. 57, no. 10, pp. 2808–2820, 2011.

[50] J. Rathouskỳ and V. Havlena, "Mpc-based approximate dual controller by information matrix maximization," *International Journal of Adaptive Control and Signal Processing*, vol. 27, no. 11, pp. 974–999, 2013.

[51] R. S. Patwardhan and R. B. Goapluni, "A moving horizon approach to input design for closed loop identification," *Journal of Process Control*, vol. 24, no. 3, pp. 188–202, 2014.

[52] C. A. Larsson, C. R. Rojas, and H. Hjalmarsson, "Mpc oriented experiment design," *IFAC Proceedings Volumes*, vol. 44, no. 1, pp. 9966–9971, 2011.

[53] C. A. Larsson, M. Annergren, H. Hjalmarsson, C. R. Rojas, X. Bombois, A. Mesbah, and P. E. Modén, "Model predictive control with integrated experiment design for output error systems," in *2013 European Control Conference (ECC)*, pp. 3790–3795, IEEE, 2013.

[54] T. A. N. Heirung, B. Foss, and B. E. Ydstie, "Mpc-based dual control with online experiment design," *Journal of Process Control*, vol. 32, pp. 64–76, 2015.

[55] D. Telen, B. Houska, M. Vallerio, F. Logist, and J. Van Impe, "A study of integrated experiment design for nmpc applied to the droop model," *Chemical Engineering Science*, vol. 160, pp. 370–383, 2017.

[56] F. Galvanin, C. C. Ballan, M. Barolo, and F. Bezzo, "A general model-based design of experiments approach to achieve practical identifiability of pharmacokinetic and pharmacodynamic models," *Journal of pharmacokinetics and pharmacodynamics*, vol. 40, no. 4, pp. 451–467, 2013.

[57] 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," *International journal of food microbiology*, vol. 54, no. 1-2, pp. 27–38, 2000.

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

[59] D. J. Pritchard and D. W. Bacon, "Prospects for reducing correlations among parameter estimates in kinetic models," *Chemical Engineering Science*, vol. 33, no. 11, pp. 1539–1543, 1978.

[60] G. Franceschini and S. Macchietto, "Novel anticorrelation criteria for design of experiments: Algorithm and application," *AIChE journal*, vol. 54, no. 12, pp. 3221–3238, 2008.

[61] 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.

[62] G. Franceschini and S. Macchietto, "Anti-correlation approach to model-based experiment design: application to a biodiesel production process," *Industrial & Engineering Chemistry Research*, vol. 47, no. 7, pp. 2331–2348, 2008.

[63] J. Pinto, M. Lobao, and J. Monteiro, "Sequential experimental design for parameter estimation: a different approach," *Chemical engineering science*, vol. 45, no. 4, pp. 883–892, 1990.

[64] E. Walter and L. Pronzato, *Identification of parametric models from experimental data*. Springer Verlag, 1997.

[65] L. Zullo, "Computer aided design of experiments," *An engineering approach (Ph. D. thesis). UK: The University of London*, 1991.

[66] S. J. Yoo, J. H. Kim, and J. M. Lee, "Dynamic modelling of mixotrophic microalgal photobioreactor systems with time-varying yield coefficient for the lipid consumption," *Bioresource technology*, vol. 162, pp. 228–235, 2014.

[67] Y. Chu and J. Hahn, "Generalization of a parameter set selection procedure based on orthogonal projections and the d-optimality criterion," *AIChE Journal*, vol. 58, no. 7, pp. 2085–2096, 2012.

[68] 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.

[69] Y. Chu and J. Hahn, "Parameter set selection for estimation of nonlinear dynamic systems," *AIChE journal*, vol. 53, no. 11, pp. 2858–2870, 2007.

초 록

회분식 및 반회분식 반응기 모델은 매우 복잡하고 비선형성이 크기 때문에, 파라미터 추정이 매우 어렵다. 모델에 대한 구조가 알려져 있는 상태라면, 파라미터 추정을 위해서 모델 기반 실험계획법(MBDOE)를 사용할 수 있다. 하지만 이 MBDOE를 회분식 반응기의 파라미터 추정에 적용할 경우 여러 가지의 치명적인 문제점이 발생하게 된다. 첫 번째, MBDOE의 결과가 초기 파라미터 추정치에 따라 달라진다. 두 번째, 문제 자체의 크기가 너무 커서 한정된 시간 안에 믿을 만한 해를 구하기가 불가능하다. 세 번째, 파라미터들간의 상관성 때문에 수치적으로 안정된 MBDOE 계산을 수행 하는 것이 어렵다. 본 논문에서는 이러한 기존의 MBDOE 기법의 문제점들을 해결하는 두 가지의 새로운 MBDOE 기법을 제안한다. 첫 번째 MBDOE는 기존의 온라인 MBDOE를 그 크기가 큰 모델에도 효율적으로 적용 가능한 형태로 개선하여 초기 파라미터에 대한 의존성 문제, 계산 시간 문제와, sensitivity matrix의 불안정성 문제를 해결한다. 두 번째로 제안한 MBDOE는 기존의 anti-correlation MBDOE을 더 개선시켜서 반복 실험에 적당하고 수치적으로 안정한 형태로 발전시킨다. 마지막으로, 이렇게 제안된 두 가지의 방법론을 반회분식 미세조류 모델의 파라미터 추정 문제에 실제로 적용하여, 알고리즘의 사용 방법을 실제적으로 증명하고, 적용 과정에서 발생할 수 있는 다양한 문제들에 대해 탐구하였다.

주요어 : 반회분식 공정, 생물공정, 파라미터 추정, 모델기반 실험계획법

학번 : 2013-20962