Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnline CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 55-7 (2022) 934–939

Model predictive control guided with optimal experimental design for pulse-based parallel cultivation [1]

Jong Woo Kim [*,2] **Niels Krausch** [*] **Judit Aizpuru** [*] **Tilman Barz** [**] **Sergio Lucia** [***] **Ernesto C. Martínez** [****] **Peter Neubauer** [*] **Mariano N. Cruz Bournazou** [*,†]

[*] *Technische Universität Berlin, Chair of Bioprocess Engineering, Strasse des 17. Juni 135, 10623 Berlin, Germany*

[**] *AIT Austrian Institute of Technology GmbH, Center for Energy, Giefinggasse 4, 1210 Vienna, Austria*

[***] *Technische Universität Dortmund, Department of Biochemical and Chemical Engineering, Emil-Figge-Strasse 70, 44227 Dortmund, Germany*

[****] *INGAR (CONICET-UTN), Avellandeda 3657, S3002GJC Santa Fe, Argentina*

[†] *DataHow AG, Zürichstrasse 137, 8600 Dübendorf, Switzerland*

Abstract: Optimal experimental design for parameter precision attempts to maximize the information content in experimental data for a most effective identification of parametric model. With the recent developments in miniaturization and parallelization of cultivation platforms for high-throughput screening of optimal growth conditions massive amounts of informative data can be generated with few experiments. Increasing the quantity of the data means to increase the number of parameters and experimental design variables which might deteriorate the identifiability and hamper the online computation of optimal inputs. To reduce the problem complexity, in this work, we introduce an auxiliary controller at a lower level that tracks the optimal feeding strategy computed by a high-level optimizer in an online fashion. The hierarchical framework is especially interesting for the operation under constraints. The key aspect of this method are discussed together with an *in silico* study considering parallel glucose limited bacterial fed batch cultivations.

Keywords: Interaction between design and control, Bio-applications, Batch process modeling and control, Design of experiments

1. INTRODUCTION

Obtaining a mechanistic model of the microbial system is crucial to the effective, consistent, and reliable bioprocess development (Neubauer et al., 2013). A mechanistic model is described by the set of parameters, which characterize the kinetics of the key metabolic pathways (Anane et al., 2019). The parameters are estimated by fitting the parameter to the experimental data with the maximum likelihood estimation objective. The accuracy of the mechanistic model is significantly affected by the amount of experimental data available. Nonetheless, it is difficult to obtain the satisfactory amount of data especially for a bioprocess. This is because the cultivation experiment is typically costly and time consuming. Moreover, the key states of the cultivation are mainly measured through the at-line channels, which are significantly scarcer than those from the online sensors.

Recently, high-throughput (HT) technology allows for obtaining massive experimental data, which accelerates bioprocess development (Cruz Bournazou et al., 2017). Liquid handling station supports for automatizing, parallelizing, and miniaturizing the experimental facilities to perform the laborious cultivation experiments (Hemmerich et al., 2018). In the down-scaled cultivation based on the mini-bioreactors, the heterogeneous (or oscillating) operating conditions in the large-scale bioreactors should be simulated (Anane et al., 2019). Moreover, the microfluidic device for the continuous feeding on the milliliter scale reactor is technically difficult to achieve (Faust et al., 2014). For these reasons, pulse-based glucose feeding has been incorporated for the HT bioprocess development (Hans et al., 2020). It is crucial to exactly capture the sharp change of states from the pulse-feed by an exact parameter estimation, because the state values right after the pulse-feed are often locate at the boundary of the operating conditions, activating the corresponding inequality constraints.

Although the data amount is amplified thanks to the HT technique and thereby resolve issues rooted from the

[1] This work was supported by the German Federal Ministry of Education and Research through the Program International Future Labs for Artificial Intelligence (Grant number 01DD20002A).

[2] Corresponding author, E-mail: jong.w.kim@tu-berlin.de

10.1016/j.ifacol.2022.07.564

scarcity of data, it is still crucial to design the experiment that can maximize the information content of the data given the experimental facility. The purpose of the optimal experimental design (OED) is to search for the feeding strategy that maximizes the information content of the measurement, quantified by the Fisher information matrix (FIM) (Franceschini and Macchietto, 2008). Through the sequential procedure, which consists of 1) computing new experimental strategy, 2) obtaining data, and 3) estimating the parameters, the parametric uncertainties become progressively reduced (Martinez et al., 2009). The online OED following such procedure has been applied for dynamic systems such as batch fermentation, liquid chromatography, and oxidative reaction (Galvanin et al., 2009; Telen et al., 2014; Barz et al., 2016).

A critical issue for the online implementation of the OED is the existence of non-identifiable parameters that make the parameter estimation ill-posed. Especially at the beginning of an experiment (i.e., batch phase) the scarcity of at-line measurements of bioprocesses might cause a huge error for the initial parameter guesses, and therefore leads to the failure of the model-based approaches for the entire experiment (López Cárdenas et al., 2015; Barz et al., 2016). Running parallel experiments is advantageous to alleviate this ill-posedness issue in the sense that the amounts of at-line data are multiplied by the number of parallel reactors. Parameter estimation of the parallel experiments has been successfully performed in (Anane et al., 2019; Hans et al., 2020).

Nonetheless, a new problem arises at the OED side from incorporating multiple reactors simultaneously. The online computation or numerical solution of the re-design problem becomes intractable as the number of design variables increases proportionally to the number of parallel bioreactors. Concerning that the typical time interval for online decision of the cultivation is in the order of few minutes, full parallel OED should be relaxed. There exist only a few studies for the parallel OED. One approach is to decompose to the sequential OED for the individual systems by maximizing k^{th} largest singular value of the FIM of k^{th} system (Galvanin et al., 2007). This is based on the assumption that the individual optimizer constitutes the global optimizer. Another approach rather focuses on the stabilization of the numerical solution of the parallel OED problem by the subset selection method (Cruz Bournazou et al., 2017; Barz et al., 2018). However, previous studies have not considered the constrained nonlinear programming formulation based on the full discretization of the dynamic systems and sensitivity. Such method can benefit from the efficient gradient-based solver on providing the analytic derivative (Bauer et al., 2000). In addition, the parallel OED becomes flexible in incorporating important process constraints as well as being extended to the robust setting or data-driven methods (Körkel et al., 2004; Lucia and Paulen, 2014; Telen et al., 2018).

This paper extends the work of Barz et al. (2018), in which the adaptive optimal design of 4 parallel fed-batch mini-bioreactors was performed using single shooting optimization. We add an oxygen constraint to the OED which is important for the cell growth by preventing the oxygen limitation condition. To avoid the computational intractability of the constrained OED under the full discretization method, we separate the problem into the unconstrained OED and the additional constrained controller. The auxiliary controller is responsible for tracking the optimal experimental strategy of each bioreactor computed by the unconstrained OED under the oxygen constraint. Auxiliary controllers now concern the dynamics of each single bioreactor, hence they become independent from each other and easily parallelizable. Moreover, the objective function of the tracking problem is convex and the size of the problem is considerably reduced. The parallel OED is solved with much lower frequency, while the auxiliary controllers are solved in every decision time steps. This hierarchical structure enables the online implementation of the parallel constrained OED by experiencing minimal approximation, as demonstrated similarly in the adaptive optimization of bioprocess (Kim et al., 2021). Proposed methods are demonstrated using the 8 parallel cultivation experiments base on the *in silico* setting.

2. MACRO-KINETIC GROWTH MODEL

The robotic facility of the HT platform is able to conduct eight parallel cultivations. The following measurements are considered:

- Online measurements: Dissolved oxygen tension (DOT) and pH are recorded every 30 seconds online.
- Atline measurements: Biomass, substrate, acetate, and product concentrations are analyzed from samples taken every 120 minutes.
- Pipette actions: Glucose solution 200 g/L is added every 10 min in each reactor. The amount of pulse addition of the glucose solution is the decision variable.

The cultivation is divided into two phases; batch and fed-batch. In the batch phase, the biomass grows by consuming the substrate which initially exists in the reactor. Fed-batch starts as soon as the substrate is depleted, which can be detected by observing the steep increase in the DOT signal.

The macro-kinetic growth (MKG) model is described by a set of ordinary differential equations (ODEs) of six state variables, biomass X, substrate S, acetate A, dissolved oxygen tension measurement DOT_m, product P, and the reactor medium volume V. The governing equations for the biomass, substrate, acetate, and product concentrations are expressed as

$$\frac{dX}{dt} = -\mu X + \frac{X}{V}F_\lambda \tag{1}$$

$$\frac{dS}{dt} = -q_S X + \frac{S}{V}F_\lambda \tag{2}$$

$$\frac{dA}{dt} = q_A X + \frac{A}{V}F_\lambda \tag{3}$$

$$\frac{dP}{dt} = q_P X + \frac{P}{V}F_\lambda \tag{4}$$

where μ, q_S, q_A, and q_P are the specific growth rate $(g/(g{\cdot}h))$, specific substrate uptake rate $(g/(g{\cdot}h))$, specific acetate production rate $(g/(g \cdot h))$, and specific product formation rate $(g/(g \cdot h))$, respectively; F_λ (L/h) is the evaporation rate. Dissolved oxygen tension (DOT) is modelled by the algebraic equation considering equilibrium oxygen concentration in the reactor medium as

$$DOT = DOT^* - \frac{q_O X H}{k_{la}} \tag{5}$$

where k_{la} (h^{-1}) denotes the volumetric oxygen transfer coefficient; DOT^* (%) denotes the saturation concentration of DOT; q_O $(g/(g \cdot h))$ denotes the specific oxygen uptake rate; H $(mol/(m^3 \cdot Pa))$ denotes the Henry constant. The oxygen sensor has first order delay, which is written as

$$\frac{dDOT_m}{dt} = k_p(DOT - DOT_m) \tag{6}$$

where k_p (h^{-1}) represent the time constant.

The parameter vector θ comprises of the physical parameters of the MKG model. The reader is referred to Anane et al. (2017) for the detailed description of the MKG model based on the acetate cycling and glucose partitioning. We distinguish the global parameters and local reactor-dependent parameters as θ_g and θ_l, respectively as;

$$\begin{aligned} \theta_g &= \left\{ \begin{array}{l} q_{S,max}, q_m, q_{Ap,max}, q_{Ac,max}, \\ Y_{XS,em}, Y_{AS,of}, Y_{XA}, Y_{OS}, Y_{OA}, Y_{PS}, \\ K_S, K_{qS}, K_{i,SA}, K_A, K_{i,AS}, d_{S,ox,P} \end{array} \right\} \\ \theta_l &= \{k_{la}, k_p\} \end{aligned} \tag{7}$$

In the fed-batch phase, the glucose feed is given in pulse-type rather than continuous-type. This results in instantaneous jumps in the process variables, which are governed by the mass balance. Denote t and t^+ the times just before and after the pulse input occurs, respectively. The mass balance is described by

$$X(t^+) = X(t) - \frac{X(t)}{V(t^+)}\Delta v(t) \tag{8a}$$

$$S(t^+) = S(t) - \frac{S(t) - S_f}{V(t^+)}\Delta v(t) \tag{8b}$$

$$A(t^+) = A(t) - \frac{A(t)}{V(t^+)}\Delta v(t) \tag{8c}$$

$$DOT_m(t^+) = DOT_m(t) \tag{8d}$$

$$P(t^+) = P(t) - \frac{P(t)}{V(t^+)}\Delta v(t) \tag{8e}$$

$$V(t^+) = V(t) + \Delta v(t) \tag{8f}$$

where $\Delta v(t)$ (L) denotes the amount of pulse-feed at time t; S_f (g/L) denotes the substrate concentration in the pulse-feed.

3. MODEL BASED OPTIMAL EXPERIMENTAL DESIGN

3.1 Problem description and formulation

The states x, manipulated variables u, and measured variables y comprise of

$$\begin{aligned} x &= [X, S, A, DOT_m, P, V] \\ y &= [X, S, A, DOT_m, P] \\ u &= [\Delta v] \end{aligned} \tag{9}$$

The high-throughput experiment is characterized by following discrete variables:

$$\begin{aligned} \mathcal{R} &= \{(row) | row \in \{A, B, \ldots, H\}\} \\ \mathcal{U} &= \{10k \text{ (min)} | k \in \mathbb{N}\} \\ \mathcal{M}_{r,y} &= \left\{ \begin{array}{ll} \{120k \text{ (min)} | k \in \mathbb{N}\}, & y \in \{X, S, A, P\} \\ \{30k \text{ (sec)} | k \in \mathbb{N}\}, & r \in \mathcal{R}, \quad y = DOT_m \end{array} \right. \end{aligned} \tag{10}$$

where $\mathcal{R}$ is the index set of the mini-bioreactors; $\mathcal{U}$ is the discrete pulse-feeding times; $\mathcal{M}_{r,y}$ is the measurement times of the reactor r for the measured variable y. The collection of all-time elements of $\mathcal{M}_{r,y}$ is denoted as $\mathcal{M} = \bigcup_{y,r \in \mathcal{R}} \mathcal{M}_{r,y}$.

Denote differential equations of the MKG model (Eqs. (1)-(6)) as $f \in \mathbb{R}^{n_x}$, algebraic equations of mass balance due to the pulse-feed (Eqs. (8)) as $f_d \in \mathbb{R}^{n_x}$, and output functions for reactor r as $h_r \in \mathbb{R}^{n_y}$. Then the parallel cultivation setup of the mini-bioreactors can be described in the compact form as:

$$\begin{aligned} \dot{x}_r(t) &= f(x_r(t), \theta_r), \quad t \in [t_0, t_f] \setminus \mathcal{U} \\ x_r(t^+) &= f_d(x_r(t), u_r(t)), \quad t \in \mathcal{U} \\ y_r(t) &= h_r(x_r(t)) \quad t \in \mathcal{M} \\ x_r(t_0) &= x_{0,r}, \quad \forall r \in \mathcal{R} \end{aligned} \tag{11}$$

where the subscript r indicates that variables x, u, and y belong to the reactor r; t_0 and t_f are the initial and final cultivation time, respectively; t^+ is the time after which pulse-feed is made. The parameter vector for the individual reactor $r \in \mathcal{R}$ is denoted as θ_r, and the parameter vector that contains the global parameters and the local parameters for the entire reactors is denoted as $\theta \in \mathbb{R}^{n_\theta}$. Following shows the definition:

$$\begin{aligned} \theta_r &= [\theta_g, \theta_{l,r}], \quad r \in \mathcal{R} \\ \theta &= [\theta_g, \{\theta_{l,r} | r \in \mathcal{R}\}] \end{aligned} \tag{12}$$

Each bioreactor has the initial condition $x_{0,r}$. The output function is given by $h_r(x_r(t)) = x_r(t)$ if $t \in \mathcal{M}_{r,y}$ and not defined elsewhere.

3.2 Parameter estimation and Fisher information matrix

Parameters are estimated based on the maximum likelihood estimation method by incorporating previous measurements. Measurements at time instance t_i are assumed to follow the normal distribution with the variance-covariance matrix $\Sigma_i \in \mathbb{R}^{n_y \times n_y}$. The correlations between measurements are not considered, hence Σ_i is a diagonal matrix. Parameters are estimated by minimizing the objective function defined as:

$$\|\hat{\theta}^- - \theta\|^2_{R_0^{-1}} + \sum_{r \in \mathcal{R}} \sum_{t_i \in \mathcal{M}} \|h_r(x_r(t_i)) - y_{r,i}\|^2_{\Sigma_{t_i}^{-1}} \tag{13}$$

where R_0 denotes a weighting constant; $y_{r,i}$ is the measurement of the reactor r at time t_i. The residual vector is scaled by $|\mathcal{M}_{r,y}|$, the number of measurements on the reactor r and output variable y. Here, Tikhonov regularization that penalizes the parameters deviating from the previous estimates (i.e., $\hat{\theta}^-$) is utilized to prevent the ill-conditioning issue (Barz et al., 2016). The optimal parameter vector $\hat{\theta}$ is evaluated by solving the optimization problem Eq. (13) subject to the model equations in Eqs. (11).

If $\hat{\theta}$ is the unconstrained optimizer of Eq. (13), then the Cramer-Rao inequality provides the lower bound of the parameter covariance as

$$\begin{aligned} F(t_0) &= \sum_{r \in \mathcal{R}} \sum_{t_i \in \mathcal{M}} s_r(t_i)^T \Sigma_i^{-1} s_r(t_i) \\ C_{[0,t_0]}(\hat{\theta}) &\approx F(t_0)^{-1} \end{aligned} \tag{14}$$

where $s_r(t_i) \in \mathbb{R}^{|\mathcal{M}| \times n_\theta}$ is the sensitivity matrix of the output of the bioreactor r with respect to the parameter vector, given as $s_r(t_i) = \dfrac{\partial y_r}{\partial \theta}$; $F(t_0)$ is Fisher information matrix; $C_{[0,t_0]}(\hat{\theta})$ is the parameter covariance matrix computed at time t_0.

3.3 Optimal experimental design

We formulate the OED with the Fisher information matrices (FIM) for the past and future measurements. Denote the current and future times as t_0 and t_p, respectively, and the collection of the measurement times within $[t_0, t_p]$ as $\mathcal{M}^+$. Under the last estimates $\hat{\theta}$, the Fisher information matrix is additive:

$$\left(C_{[0,t_p]}(\hat{\theta})\right)^{-1} \approx \left(C_{[0,t_0]}(\hat{\theta})\right)^{-1} + \sum_{r\in\mathcal{R}} \sum_{t_i \in \mathcal{M}^+} s_r(t_i)^T \Sigma_i^{-1} s_r(t_i) \tag{15}$$

The exact evaluation of Eq. (15) is impossible, due to the dependency of the FIM on the model parameter values with the measurements, which cannot be known in advance. Therefore, OED approximates the parameters to be fixed to their prior estimates until the prediction horizon t_p (Galvanin et al., 2009).

There exist several scalar metrics for the Fisher information matrix. In this paper, we apply A criterion of the objective function of the OED problem. The definition is given as:

$$\psi_A = \frac{1}{n_\theta} \mathrm{Tr}\left(C_{[0,t_p]}(\hat{\theta})\right) \tag{16}$$

where $\mathrm{Tr}(\cdot)$ stands for the trace of the matrix, respectively.

Dynamic propagation of the sensitivity matrix can be computed from the chain rule as

$$\dot{s}_r(t) = \frac{\partial f}{\partial x} s_r(t) + \frac{\partial f}{\partial \theta}, \quad s_r(0) = \frac{\partial x_r(0)}{\partial \theta} \tag{17}$$

In addition, the relationship between sensitivity matrices before and after the pulse-feed is obtained from differentiating the mass balance equations f_d as

$$s_r(t^+) = \frac{\partial f_d}{\partial x} s_r(t) + \frac{\partial f_d}{\partial \theta} \tag{18}$$

Accommodating the dynamic models and dynamic sensitivity matrices for all $r \in \mathcal{R}$, and the objective function, the OED problem optimizes the manipulated variable $u_r, r \in \mathcal{R}$ in $[t_0, t_p]$. The problem is formulated as

$$\min_{u_r, r\in\mathcal{R}} \quad \psi_A \tag{19a}$$
$$\text{s.t.} \quad \dot{x}_r(t) = f(x_r(t), \theta_r), \quad t \in [t_0, t_p] \setminus \mathcal{U} \tag{19b}$$
$$x_r(t^+) = f_d\left(x_r(t), u_r(t)\right), \quad t \in \mathcal{U} \tag{19c}$$
$$\dot{s}_r(t) = \frac{\partial f}{\partial x} s_r(t) + \frac{\partial f}{\partial \theta}, \quad t \in [t_0, t_p] \setminus \mathcal{U} \tag{19d}$$
$$s_r(t^+) = \frac{\partial f_d}{\partial x} s_r(t) + \frac{\partial f_d}{\partial \theta}, \quad t \in \mathcal{U} \tag{19e}$$
$$x_{\min} \le x_r(t) \le x_{\max}, \quad t \in [t_0, t_p] \tag{19f}$$
$$u_{\min} \le u_r(t) \le u_{\max}, \quad t \in \mathcal{U} \tag{19g}$$
$$x_r(0) = x_{0,r} \tag{19h}$$
$$r \in \mathcal{R} \tag{19i}$$

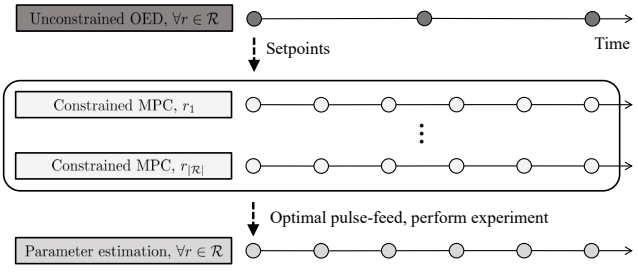

Fig. 1. Hierarchical structure of the parallel OED and auxiliary MPC controllers.

where $x_{\min}$ and $x_{\max}$ are the lower and upper bounds for the state variables; $u_{\min}$ and $u_{\max}$ are the lower and upper bounds for the manipulated variable.

3.4 OED-guided auxiliary model predictive controller

The number of decision variables in the OED problem in Eqs. (19) comprises the model and sensitivity dynamics multiplied by the number of bioreactors as

$$|\mathcal{R}|\left(n_x + n_x n_\theta + n_u\right) n_{col} n_p \tag{20}$$

where n_{col} is the number of collocation points and n_p is the prediction horizon. Because the computational complexity of the optimization is super-linear with respect to the number of decision variables, it is intractable to perform the re-design of 8 bioreactors in a fully online manner. Nevertheless, if the experimental design is not adapted to recent measurements, the process will be operated under the region with low information content. This eventually leads to the violation of the crucial DOT constraint, which is parameter dependent (see Eq. (5)).

To overcome such limitation, we propose utilizing an auxiliary model predictive controller (MPC) that tracks the pulse-feed strategy of a single bioreactor computed by the OED and considering the constraints in the auxiliary MPC instead of in the parallel OED. Here, we relax the DOT constraint of the OED problem which is likely to be violated in the large-scale optimization problem. Instead, the DOT constraint is considered in the auxiliary controllers individually. The optimization problem for the auxiliary controller for the reactor $r \in \mathcal{R}$ is formulated as follows:

$$\min_{u_r} \quad \sum_{t\in\mathcal{U}} \left\| u_r(t) - u_r^{oed}(t) \right\|^2 \tag{21a}$$
$$\text{s.t.} \quad \dot{x}_r(t) = f(x_r(t), \theta_r), \quad t \in [t_0, t_p] \setminus \mathcal{U} \tag{21b}$$
$$x_r(t^+) = f_d\left(x_r(t), u_r(t)\right), \quad t \in \mathcal{U} \tag{21c}$$
$$x_{\min} \le x_r(t) \le x_{\max}, \quad t \in [t_0, t_p] \setminus \mathcal{U} \tag{21d}$$
$$u_{\min} \le u_r(t) \le u_{\max}, \quad t \in \mathcal{U} \tag{21e}$$
$$DOT_{lb,r} \le DOT_r(t, \theta_r), \quad t \in [t_0, t_p] \tag{21f}$$
$$x_r(0) = x_{0,r}, \tag{21g}$$

where $u_r^{oed}(t)$ is the pulse-feed strategy computed from the optimization problem of Eqs. (19); $DOT_{lb,r}$ represents the DOT lower bound of the reactor r. The number of decision variables of Eq. (21) is given as

$$\left(n_x + n_u\right) n_{col} n_p \tag{22}$$

Table 1. Size of the optimization problems of the OED and auxiliary MPCs, and their computation times.

	OED	Aux. MPC
Num. decision var.	109728	1556
Num. equality constraints	109576	1537
Num. inequality constraints	4560	570
CPU time (min)	22.7	1.2

which is considerably reduced compared to Eq. (20). The total computational complexity is now linear with respect to the number of reactors. The implementation procedure is illustrated in Fig. 1. Unconstrained OED is computed with lower frequency due to its complexity, while auxiliary MPC controllers for the individual reactors are computed every decision time steps (i.e., 10 min). With this hierarchical structure, the experimental design for the parallel mini-bioreactors with the oxygen constraint can be conducted online.

4. *IN SILICO* RESULTS

The *in silico* cultivation for 8 mini-bioreactors is performed. To generate the *in silico* data, we add uncertainties to the parameters and measurements by adding random uniform noises. 10 % and 5 % with respect to the operation bounds are added to the parameters and measurements, respectively. Do-mpc software, a toolbox developed based on CasADi (Andersson et al., 2019), is utilized to solve the optimization problems (Lucia et al., 2017). The prediction horizon for the OED and auxiliary MPCs are until the end of the batch and 180 min, respectively. The number of decision variables and constraints and computation times of the OED and auxiliary MPCs are written in Table 1. The computation time for the OED exceeds 10 min, and therefore it is solved every 60 min. On the other hand, the computation time for solving 8 auxiliary MPC problems take 1.2 min.

A cultivation result of one of eight reactors operated by the unconstrained OED is presented in Fig. 2. Biomass, substrate, acetate, DOT_m, DOT, and pulse-feed trajectories are plotted. The sensitivities for the four states (i.e., X, S, A, and DOT_m) with respect to 18 kinetic parameters θ_g and θ_1 (Eq. (7)) are shown in Fig. 3. This figure highlights that the magnitude of sensitivity values arise as the pulse-feed is given. The increase of sensitivities are proportional to the feed amount, and therefore the OED objective encourages to feed as much as possible. Because the oxygen uptake is proportional to the biomass amount according to the Eq. (5), the optimal solution of the OED problem activates the DOT constraint.

Figure 4 depicts the cultivation results of one of eight reactors operated by the OED problem with the auxiliary MPC. The MPC is solved every 10 min. MPC tracks the OED result initially where the violation does not happen. After the violation is detected within the prediction horizon, the pulse-feed amount is decreased. The feasibility can be guaranteed by shifting the DOT constraint to the auxiliary controllers.

Finally, we validate that the loss of the information content of the pulse-feed strategy computed by the proposed method is not significant compared to the constrained

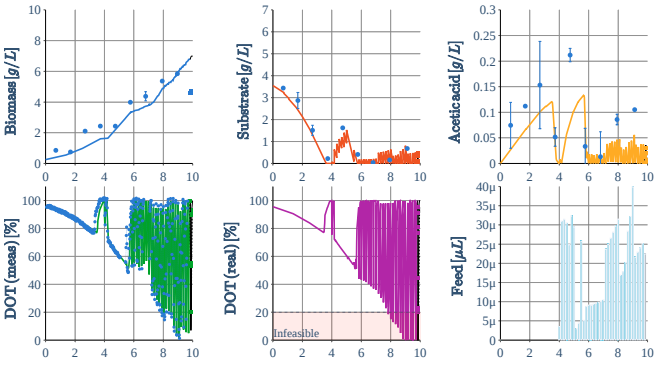

Fig. 2. State trajectory under the pulse-feed computed by the pure OED. The colored solid line represents the state trajectory, the dot represents the measurement.

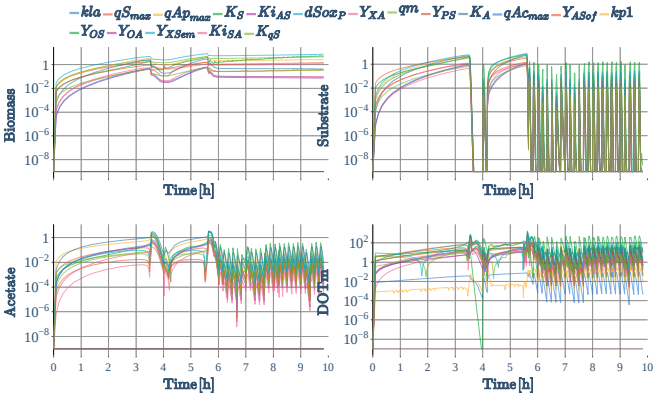

Fig. 3. Dynamic sensitivities for four states with respect to 18 parameters throughout the cultivation experiment.

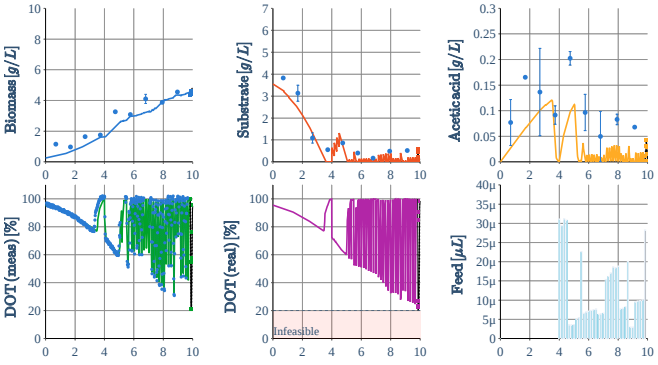

Fig. 4. State trajectory under the pulse-feed computed by OED guided MPC. The colored solid line represents the state trajectory, the dot represents the measurement.

OED. For comparison, we consider the OED problem for a single reactor because the optimization problem is able to be solved online. ψ_A values for three different feeding strategies, proposed hierarchical structure, constrained OED problem, and the OED without constraint, are 1.014, 1.106, and 3.386, respectively. This indicates that due to the approximation the reduction of the information content is unavoidable, however not considerable.

5. CONCLUDING REMARKS

In this study, we introduce the auxiliary controller to enable online implementation of the computationally intractable OED for the parallel cultivation. The auxiliary controller has a simple structure that can be parallelizable, and has a quadratic objective for tracking the results from the OED problem. The additional constraint such as oxygen limitation constraint can be relaxed from the OED problem and then be accounted for to the auxiliary controller. Through the *in silico* study for the cultivation of eight parallel replicates, we demonstrate that the proposed approach yields a closed-loop feeding strategy with near-optimal information content within the feasible region. Future work will focus on extending the strategy to the full high-throughput experiment involving different experimental conditions. Results obtained demonstrate that it is also important to account for the parametric uncertainty by formulating robust OED.

REFERENCES

Anane, E., Lopez Cárdenas, D.C., Neubauer, P., and Cruz Bournazou, M.N. (2017). Modelling overflow metabolism in Escherichia coli by acetate cycling. *Biochemical engineering journal*, 125, 23–30.

Anane, E., Sawatzki, A., Neubauer, P., and Cruz-Bournazou, M.N. (2019). Modelling concentration gradients in fed-batch cultivations of E. coli - towards the flexible design of scale-down experiments. *Journal of Chemical Technology & Biotechnology*, 94(2), 516–526.

Andersson, J.A.E., Gillis, J., Horn, G., Rawlings, J.B., and Diehl, M. (2019). CasADi – A software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation*, 11(1), 1–36. doi: 10.1007/s12532-018-0139-4.

Barz, T., Bournazou, M.N.C., Körkel, S., Walter, S.F., et al. (2016). Real-time adaptive input design for the determination of competitive adsorption isotherms in liquid chromatography. *Computers & Chemical Engineering*, 94, 104–116.

Barz, T., Sommer, A., Wilms, T., Neubauer, P., and Cruz Bournazou, M.N. (2018). Adaptive optimal operation of a parallel robotic liquid handling station. *IFAC-PapersOnLine*, 51(2), 765–770.

Bauer, I., Bock, H.G., Körkel, S., and Schlöder, J.P. (2000). Numerical methods for optimum experimental design in DAE systems. *Journal of Computational and Applied mathematics*, 120(1-2), 1–25.

Cruz Bournazou, M.N., Barz, T., Nickel, D., Lopez Cárdenas, D., Glauche, F., Knepper, A., and Neubauer, P. (2017). Online optimal experimental re-design in robotic parallel fed-batch cultivation facilities. *Biotechnology and bioengineering*, 114(3), 610–619.

Faust, G., Janzen, N.H., Bendig, C., Römer, L., Kaufmann, K., and Weuster-Botz, D. (2014). Feeding strategies enhance high cell density cultivation and protein expression in milliliter scale bioreactors. *Biotechnology journal*, 9(10), 1293–1303.

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

Galvanin, F., Barolo, M., and Bezzo, F. (2009). Online model-based redesign of experiments for parameter estimation in dynamic systems. *Industrial & Engineering Chemistry Research*, 48(9), 4415–4427.

Galvanin, F., Macchietto, S., and Bezzo, F. (2007). Model-based design of parallel experiments. *Industrial & engineering chemistry research*, 46(3), 871–882.

Hans, S., Haby, B., Krausch, N., Barz, T., Neubauer, P., and Cruz Bournazou, M.N. (2020). Automated Conditional Screening of Multiple Escherichia coli Strains in Parallel Adaptive Fed-Batch Cultivations. *Bioengineering*, 7(4), 145.

Hemmerich, J., Noack, S., Wiechert, W., and Oldiges, M. (2018). Microbioreactor systems for accelerated bioprocess development. *Biotechnology journal*, 13(4), 1700141.

Kim, J.W., Park, B.J., Oh, T.H., and Lee, J.M. (2021). Model-based reinforcement learning and predictive control for two-stage optimal control of fed-batch bioreactor. *Computers & Chemical Engineering*, 154, 107465.

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

López Cárdenas, D.C., Barz, T., Körkel, S., and Wozny, G. (2015). Nonlinear ill-posed problem analysis in model-based parameter estimation and experimental design. *Computers & Chemical Engineering*, 77, 24–42.

Lucia, S. and Paulen, R. (2014). Robust nonlinear model predictive control with reduction of uncertainty via robust optimal experiment design. *IFAC Proceedings Volumes*, 47(3), 1904–1909.

Lucia, S., Tătulea-Codrean, A., Schoppmeyer, C., and Engell, S. (2017). Rapid development of modular and sustainable nonlinear model predictive control solutions. *Control Engineering Practice*, 60, 51–62.

Martinez, E.C., Cristaldi, M.D., and Grau, R.J. (2009). Design of dynamic experiments in modeling for optimization of batch processes. *Industrial & engineering chemistry research*, 48(7), 3453–3465.

Neubauer, P., Cruz, N., Glauche, F., Junne, S., Knepper, A., and Raven, M. (2013). Consistent development of bioprocesses from microliter cultures to the industrial scale. *Engineering in Life Sciences*, 13(3), 224–238.

Telen, D., Logist, F., Quirynen, R., Houska, B., Diehl, M., and Van Impe, J. (2014). Optimal experiment design for nonlinear dynamic (bio) chemical systems using sequential semidefinite programming. *AIChE Journal*, 60(5), 1728–1739.

Telen, D., Nimmegeers, P., and Van Impe, J. (2018). Uncertainty in optimal experiment design: comparing an online versus offline approaches. *IFAC-PapersOnLine*, 51(2), 771–776.