Experimental Evaluation of an Approach to Online Redesign of Experiments for Parameter Determination

Tilman Barz, Diana C. López Cárdenas, Harvey Arellano-Garcia, and Günter Wozny
Chair of Process Dynamics and Operation, Technische Universität Berlin, 10623 Berlin, Germany

DOI 10.1002/aic.13957
Published online December 3, 2012 in Wiley Online Library (wileyonlinelibrary.com).

The online redesign of experiments for parameter determination of nonlinear dynamic systems has been studied recently by different research groups. In this article, this technique is assessed in a real case study for the first time. The presented algorithm adopts well-known concepts from model-based control. Compared to previous studies, special attention is given to the efficient treatment of the underlying nonlinear and possibly ill-conditioned parameter estimation and experiment design problems. These problems are solved with single shooting and gradient-based nonlinear programming (NLP) solvers. We use an initial value solver, which generates first- and second-order sensitivities to compute exact derivatives of the problem functions. As a special feature, we propose the integration of a local parameter identifiability analysis and a corresponding algorithm that generates well-conditioned problems. The practical applicability is demonstrated by experimental application to a chromatography column system where A, D, and E optimal experiments are performed.
 AIChE J, 59: 1981–1995, 2013
Keywords: optimization, numerical solutions, chromatography

Introduction

The experimental determination of parameters from mechanistic models is accomplished by solving parameter estimation (PE) problems, where model predictions are fitted to measured data. For the most accurate determination of the model parameters, model-based optimum experiment design (ED) techniques have been widely adopted.[1–3] The optimum ED selects conditions that generate the maximum information content in the measured data with respect to the model parameters to be determined. The optimum experiment conditions are defined by the ED variables which are the input actions taken at the process. Their values are computed by minimization of the ED criterion whose definition is based on the sensitivities of model predictions to model parameters.

For general chemical and biological systems, the model is nonlinear in both the ED variables and the parameters. Accordingly, the sensitivity functions used to search for optimum experimental conditions depend on the unknown model parameters. Thus, a computed design is then only locally optimal. Indeed, if the assumed parameter values differ too much from the ''true'' parameter values, the ED can prove to be far from optimal. The robustification of the method for computing an optimal ED for parameter determination of nonlinear systems has been considered a significant challenge by many researchers. In most contributions, uncertainties are assigned to guesses of parameter values. However, other sources of uncertainties can affect an ED, for example, implementation errors which take into account possible deviations from planned conditions during experiment execution.[4,5]

Generally, two different methods can be distinguished for the treatment of uncertainties: a direct method where uncertainties are defined and considered *a priori*, and an indirect method, which relies on an iterative refinement of the experimental design together with a repeated update of the parameter values. In the direct method, uncertainties are considered when formulating the ED problem. Two criteria are usually adopted: first, the expected value criterion, which optimizes the design criterion on average (see Refs. 6 and 7 for examples), second, the minimax criterion, where the ED problem is solved for the worst case parameter values of the uncertain parameter space (see Refs. 6, 8, and 9 for examples). The indirect or ''adaptive'' method uses an iterative procedure where a local ED is generated in each iteration. After an initial parameter guess, the experiment is carried out, parameters are re-estimated from the measurement data, and subsequent optimized experiments are computed, conducted, and analyzed. This sequential solution of ED and PE problems is then continued until the estimated parameters meet predefined significance specifications (see Refs. 4 and 10, for examples).

As an extension to the adaptive design, the idea of an online model-based redesign of experiments was first presented by Ref. 1. Although the methodology was initially developed for nonphysical models, it was studied recently by various authors for mechanistic models also.[5,11–15] In contrast to the off-line adaptive design, in the online redesign of experiments, new measurement data are exploited before designing a new experiment. Hence, the experiment is conducted for one or several measurement sampling points, the

Correspondence concerning this article should be addressed to T. Barz at tilman.barz@tu-berlin.de.

parameter values are updated, and the experiment is redesigned. This procedure is then repeated until the end of the experiment. It should be noted that in Refs. 5 and 11–15, the realization of this technique differs in its algorithmic implementation. Some authors adopt concepts from model predictive control (MPC) and recursive state and PE methods. Those implementation schemes are also referred to as receding horizon ED (see Refs. 11, 14, and 15). However, the validation has always involved "numerical experiments" and, thus, some critical points relating to an online implementation, such as computational burden or robustness and convergence issues have not been covered completely.

In this contribution for the first time, an algorithm for online model-based redesign of experiments is assessed in a real case study. The adopted algorithm is in line with the implementation schemes presented by Refs. 11, 14, and 15. However, for practical applicability, the efficient solution of highly nonlinear and possibly ill-conditioned problems needs to be guaranteed. On the latter point, the algorithm is extended with a local parameter identifiability analysis. In doing so, at each time instant the parameters are identified which can be reliably estimated from the available measurements. The consideration of these selected parameters then produces well-conditioned PE and ED problems. The following assumptions are made regarding the process:

- It is a stable dynamic process.
- Continuous measurement devices are used which provide sampled data at equal intervals.
- The process state variables are not restricted (e.g., safety restrictions or product specifications are not considered).

The article is organized as follows: in section "Problem Formulation," a brief review of the general formulation of PE and ED problems is given. Additionally, we present an algorithm for parameter subset selection (SsS) which is used for a local parameter identifiability analysis and to guarantee well-conditioned PE and ED problems. In section "Online Parameter Estimation and Adaptive Redesign of Experiments," the algorithmic implementation of the framework is presented, which comprises the discretization of the problem according to the measurement sampling instants, measures for the reduction of the computational effort, the identifiability analysis, and the formulation of appropriate stopping criteria. A case study is then described, a process model is presented and details of the numerical solution of all problems are given. The experimental results are then presented and discussed.

Problem Formulation

We consider systems of differential algebraic equations (DAEs), taking into account that most of the arising dynamic models in chemical engineering can be expressed as follows

$$\frac{d}{dt}\left[g^{\mathrm{lhs}}(x(t),\theta)\right] = g^{\mathrm{rhs}}(x(t),u(t),\theta,t) \tag{1a}$$

$$0 = h(x(t),\theta,t) \tag{1b}$$

where $x(t) \in \mathbb{R}^{\mathrm{Nx}}$ denotes the dependent state variables, $u(t) \in \mathbb{R}^{\mathrm{Nu}}$ the time-varying input actions or ED variables, $\theta \in \mathbb{R}^{\mathrm{Np}}$ the unknown parameters, and $t \in [t_0, t_f] \subseteq \mathbb{R}$ the time as independent variable. Initial conditions are given as $x(t_0) = x_0$. With the definition $M(x(t),\theta) = \partial g^{lhs}(x(t),\theta)/\partial x$ and writing Eq. 1a as $M(x(t),\theta)\cdot \partial x/\partial t = g^{rhs}(x(t),u(t),\theta,t)$, it can be seen that Eq. 1 denote a set of linear implicit differential and algebraic equations, respectively.

A vector of predicted response variables $y(t) \in \mathbb{R}^{\mathrm{Ny}}$ is considered, whose elements are functions of the solution x^* (obtained from integration of Eq. 1); these functions are defined through the constant matrix $A \in \mathbb{R}^{\mathrm{Ny}\times\mathrm{Nx}}$

$$y(u,\theta,t) = A \cdot x^*(u,\theta,t) \tag{2}$$

The continuous variables $u(t)$ in Eq. 1 are transformed into a finite dimensional vector of variables which for sake of simplicity are referred to as u. Thus, the predicted responses in Eq. 2 are functions of the independent variables u, θ, and t. Usually not all state variables are sampled and collected during the experiment, and so the dimension of the corresponding predicted responses $y(u,\theta,t)$ is also smaller than the dimension of the state variables $x^*(u,\theta,t)$, the former being a subset of the latter $y(u,\theta,t) \subseteq x^*(u,\theta,t)$ and $\mathrm{Ny} \le \mathrm{Nx}$. Additionally, there exists a discrete set of time instances $t_k \in \mathbb{R}^{\mathrm{Nm}}$ for the collection of the predicted response variables $y(u,\theta,t_k)$, with

$$y(u,\theta,t_k) = \left[y_1(u,\theta,t_k), y_2(u,\theta,t_k), \cdots, y_{\mathrm{Ny}}(u,\theta,t_k)\right]^{\mathrm{T}} \quad \forall t_k \in \{t_1,\cdots,t_{\mathrm{Nm}}\} \tag{3}$$

The model parameters θ are estimated by solving the following unconstrained PE problem[16]

$$\hat{\theta} = arg\min_{\theta} \quad \Phi^{ML}(u,\theta) \tag{4}$$

where $\hat{\theta}$ is an unbiased estimator of the true parameter values and Φ^{ML} denotes the maximum likelihood (ML) criterion whose definition reads

$$\Phi^{ML}(u,\theta) = (Y(u,\theta) - Y^m)^{\mathrm{T}} \cdot \Sigma_y^{-1} \cdot (Y(u,\theta) - Y^m) \tag{5}$$

In Eq. 5, $Y(u,\theta) = \left[y(u,\theta,t_1)^{\mathrm{T}},\cdots,y(u,\theta,t_{\mathrm{Nm}})^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{\mathrm{Ny}\cdot\mathrm{Nm}}$ is the collection of response variables predicted for all discrete times t_k. The corresponding experimentally obtained data are assembled in the measurement vector $Y^m = \left[y^m(t_1)^{\mathrm{T}},\cdots,\ y^m(t_{\mathrm{Nm}})^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{\mathrm{Ny}\cdot\mathrm{Nm}}$, where $y^m(t_k) = [y_1^m(t_k), y_2^m(t_k),\cdots,y_{\mathrm{Ny}}^m(t_k)]^{\mathrm{T}}\ \forall t_k \in \{t_1,\cdots,t_{\mathrm{Nm}}\}$. Finally, $\Sigma_y \in \mathbb{R}^{\mathrm{Ny}\cdot\mathrm{Nm}\times\mathrm{Ny}\cdot\mathrm{Nm}}$ is the covariance matrix from experimental data.

The Cramér–Rao lower bound $\Sigma_{\hat{\theta}} \in \mathbb{R}^{\mathrm{Np}\times\mathrm{Np}}$ for the covariance matrix of an estimator $\hat{\theta}$ is obtained from the inverse of the Fisher Information Matrix (FIM) $F \in \mathbb{R}^{\mathrm{Np}\times\mathrm{Np}}$ [16]

$$\Sigma_{\hat{\theta}}(u,\hat{\theta}) \ge F^{-1}(u,\hat{\theta}) = S(u,\hat{\theta})^{\mathrm{T}} \cdot \Sigma_y^{-1} \cdot S(u,\hat{\theta}) \tag{6}$$

where the sensitivity matrix $S = \partial Y/\partial\theta \in \mathbb{R}^{\mathrm{Ny}\cdot\mathrm{Nm}\times\mathrm{Np}}$ is evaluated at u and $\hat{\theta}$.

Finally, with the aim of optimizing the information content of an experiment, the input actions u are adjusted by solving the ED problem[3]

$$u^*(t) = arg\min_{u} \quad \Phi^{ED}(u,\hat{\theta}) \tag{7}$$

where u^* is the optimal vector of input variables of the experiment and Φ^{ED} is one of the following scalar criteria

$\Phi^{ED} \in \{\Phi^A, \Phi^D, \Phi^E\}$ which are applied to the parameter covariance matrix $\Sigma_{\hat{\theta}}$

$$\Phi^A(u,\hat{\theta}) = \frac{1}{\mathrm{Np}} \cdot \mathrm{tr}\left(\Sigma_{\hat{\theta}}(u,\hat{\theta})\right) \tag{8a}$$

$$\Phi^D(u,\hat{\theta}) = \left(\det\left(\Sigma_{\hat{\theta}}(u,\hat{\theta})\right)\right)^{\frac{1}{\mathrm{Np}}} \tag{8b}$$

$$\Phi^E(u,\hat{\theta}) = \max\left(\lambda\left(\Sigma_{\hat{\theta}}(u,\hat{\theta})\right)\right) \tag{8c}$$

Parameter SsS

Testing nonlinear parametric model structures for identifiability and distinguishability are classical elements in model identification.[17,18] In this contribution, the determination of parameters of a preselected model structure is considered. However, errors in the structural form of the model might exist, for example, resulting from certain assumptions and simplifications. A local identifiability analysis is adopted to determine the subset of the parameter space which can be reliably estimated from the available measurements. Generally, all parameters with low or nonexisting sensitivities (columns of S with values equal or near to zero), or linearly dependent parameters (nearly identical columns of S or columns which are linear combinations of others) are not identifiable. Both cases reflect near indeterminacy in the parameter estimates, caused by having more parameters than can be reliably estimated from available measurements.[19] This situation is undesirable, because S is singular or "almost" singular from a numerical point of view leading to ill-conditioned PE and ED problems. Thus, the motivation for applying the parameter SsS is to fix ill-conditioned parameters at prior estimates and consider reduced-order and well-conditioned PE and ED problems.

The identifiability analysis of parameters is performed locally based on their corresponding assumed values. The algorithm is based on the sensitivity matrix $S(u,\hat{\theta})$. All corresponding criteria are adapted from the ill-conditioned parameter selection approach based on the analysis of the Hessian matrix $H_\theta = S^{\mathrm{T}} \cdot S$, proposed in Refs. 19 and 20. Related and alternative algorithms as well as applications can be found, for example, in Refs. 19–24.

A rank-revealing factorization is done by the singular value decomposition (SVD) of $S \in \mathbb{R}^{\mathrm{Ny\cdot Nm}\times\mathrm{Np}}$

$$S = U \cdot \Sigma \cdot V^{\mathrm{T}} \tag{9}$$

where $U \in \mathbb{R}^{\mathrm{Ny\cdot Nm}\times\mathrm{Ny\cdot Nm}}$ is a real or complex unitary matrix, $V^{\mathrm{T}} \in \mathbb{R}^{\mathrm{Np}\times\mathrm{Np}}$ conjugate transpose of V is a real or complex unitary matrix, and $\Sigma \in \mathbb{R}^{\mathrm{Ny\cdot Nm}\times\mathrm{Np}}$ is a rectangular diagonal matrix with nonnegative real numbers on the diagonal. The diagonal entries $\Sigma_{i,i}$ are the singular values $\sigma_i \geq 0$ of matrix S such as $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_{\mathrm{Np}}$. A criterion for the nearness to singularity of S is the condition number $\kappa(S)$, which is the ratio of the largest to smallest singular values, see (Eq. 10)

$$\kappa(S) = \frac{\sigma_1}{\sigma_{\mathrm{Np}}} \tag{10}$$

A "very high" condition number of S indicates an almost singular sensitivity matrix. According to Ref. 23, an upper bound $\kappa^{max} \approx 1000$ is defined and parameter identifiability either of the original or a reduced problem is given when $\kappa(S) \leq \kappa^{max}$ holds. Additionally, the collinearity index $\gamma(S)$ (see Eq. 11) is considered as singularity measurement.

$$\gamma(S) = \frac{1}{\sigma_{\mathrm{Np}}} \tag{11}$$

It measures the degree of near-linear dependence of the S columns. $\gamma(S)$ equals one if the columns of S are orthogonal and reaches infinity if the columns are linearly dependent. In Ref. 21, an empirically found threshold of $\gamma^{max} \approx 10-15$ has been named. Thus, if $\gamma(S) > \gamma^{max}$ the corresponding parameter set is considered as poorly identifiable.

Algorithm for Parameter SsS. In the general algorithm for SsS, the set dimension r of the parameter subset is obtained. The set dimension r can be taken as rank of the sensitivity matrix (quantity of linear independent columns of S). The parameter subset $\theta^{(r)}$ consists of those parameters which are associated with the highest singular values. All parameters associated with the smallest remaining singular values $\theta^{(\mathrm{Np}-r)}$ must be properly chosen and fixed, because these parameters have the linear dependency feature. All reduced problems will then only involve sensitivities with respect to the r remaining or "active" parameters; we denote these sensitivities by $S^{(r)} \in \mathbb{R}^{\mathrm{Ny\cdot Nm}\times r}$, and note that its columns are just a subset of the columns of S, with $S^{(r)} \subseteq S$.

Algorithm: SsS—get $\tilde{r}$ and Rearrange: $\theta \rightarrow \tilde{\theta}$.*

1. For the parameter set θ, compute the SVD of $S(\theta)$, yielding $S = U \cdot \Sigma \cdot V^{\mathrm{T}}$, see (Eq. 9).
2. Evaluate singularity measurements
   - Compute the condition number $\kappa(S)$, see (Eq. 10) and all subcondition numbers $\kappa_j = \sigma_1/\sigma_j$, with $j = 1,\cdots,\mathrm{Np}$ for each σ_j available in diagonal matrix Σ found in the above step.
   - Calculate the collinearity index $\gamma(S)$, see (Eq. 11) and all subcollinearity indices $\gamma_j(S) = 1/\sigma_j$, with $j = 1,\cdots,\mathrm{Np}$.
   - If $\kappa(S) \leq \kappa^{max}$ and $\gamma(S) \leq \gamma^{max}$ the parameter set is identifiable and the algorithm finishes, if not, go to the next step.
3. Determine $\tilde{r}$, such that the maximum number of $\tilde{r}$ singular values σ_j, with $j = 1,\cdots,\tilde{r}$ are found for which $\kappa_j \leq \kappa^{max}$ and $\gamma_j \leq \gamma^{max}$.
   - Build the set Γ:
   $$\Gamma = \{\sigma_j | \kappa_j \leq \kappa^{max} \wedge \gamma_j \leq \gamma^{max}\} \tag{12}$$
   - Determine $\tilde{r}$ as the set dimension of Γ.
4. Determine a permutation matrix by constructing a QR decomposition with column pivoting (QRP) for $S \in \mathbb{R}^{\mathrm{Ny\cdot Nm}\times\mathrm{Np}}$.
   $$S \cdot P = Q \cdot R \tag{13}$$
   where $Q \in \mathbb{R}^{\mathrm{Ny\cdot Nm}\times\mathrm{Ny\cdot Nm}}$ is an orthogonal matrix, $R \in \mathbb{R}^{\mathrm{Ny\cdot Nm}\times\mathrm{Np}}$ is an upper triangular matrix with decreasing diagonal elements and $P \in \mathbb{R}^{\mathrm{Np}\times\mathrm{Np}}$ is a permutation matrix which orders the columns of S according to linear independency, which means that the first columns of $S \cdot P$ are the largest independent set of columns of S.
5. Use P to reorder the parameter vector θ according to $\tilde{\theta} = P^{\mathrm{T}} \cdot \theta$.
6. Make the partition $\tilde{\theta} = \left[\tilde{\theta}^{(\tilde{r})\mathrm{T}},\ \tilde{\theta}^{(\mathrm{Np}-\tilde{r})\mathrm{T}}\right]^{\mathrm{T}}$ with $\tilde{\theta}^{(\tilde{r})}$ containing the first $\tilde{r}$ elements of $\tilde{\theta}$.

*Note that the superscript $\tilde{\{\}}$ is used to indicate the result of the SsS.

7. Fix $\tilde{\theta}^{(\mathrm{Np}-\tilde{r})}$ to *a priori* estimate $\theta^{(\mathrm{Np}-\tilde{r})}$.
8. Solve reduced-order PE or ED problem, considering $\tilde{\theta}^{(\tilde{r})}$ only.

Online PE and Adaptive Redesign of Experiments

The quality (information content) of an ED for PE depends strongly on the accuracy of the assumed parameter values (initial parameter guess). For nonlinear models, uncertainties in the initial parameter guess can lead to a suboptimal ED. To tackle this, we adopt an adaptive online ED, where the procedure iterates between local solutions of the ED and PE problem for one or several samples. The result is the iterative refinement of both the parameter estimates and the input actions. This idea has been discussed recently and theoretical applications are reported.[5,11,13–15] However, the reported algorithms differ and some points related to an experimental implementation have not been addressed. A major challenge is certainly the real-time solution of highly nonlinear and possibly ill-conditioned problems. With regard to ill-conditioned problems, we propose to merge the parameter identifiability analysis and the adaptive method for ED and PE together. A detailed description of the proposed algorithm is presented below and special attention is given to the following points:

- **Sampling and discretization of input actions**

We consider continuous measurement devices. The sampling is realized at equal intervals. Accordingly, there is neither a motivation for a reduction of the number of measurements (reduction of costs of an experiment) nor the need for an optimal placement of individual sampling time points (seek for most sensitive points). The sampling times are selected such that the measured data reflects the dynamics of the system and that the interval between measurements is long enough to compute an update of the parameter estimates and planned input actions. The input actions are discretized according to the sampling points leading to a trajectory of input actions.

- **Reduction of the computational effort using a restricted time horizon and one step iterations**

The computation of future optimal input actions means the solution of a possibly large and complex nonlinear programming (NLP) problem (ED problem), particularly because of its sensitivity-based objective. A well-known concept to reduce the complexity of the ED problem and, thus, the computational effort is the restriction of the considered future time horizon as well as the number of input actions. In addition, the number of NLP iteration steps performed at each sampling time can be limited to one.[25] These restrictions may lead to a suboptimal ED and improvements of the solution algorithms to increase computation speed are always preferable. Besides that, the complexity of the ED problem is not affected by the consideration of past experiments as the related information is included as constant contribution in the problem formulation (see below).

- **Parameter identifiability**

Usually, the assumption is that all parameters to be determined are identifiable. Thus, it is assumed that well-conditioned PE and ED problems are solved in any time instant. This is not always true, especially at the beginning of an experiment, where few measurements are available. Therefore, a parameter SsS is performed to identify the parameter which can be reliably estimated from the available measurements at each time instant. The consideration of these selected parameters then produces well-conditioned PE and ED problems, which are robust and convergent.

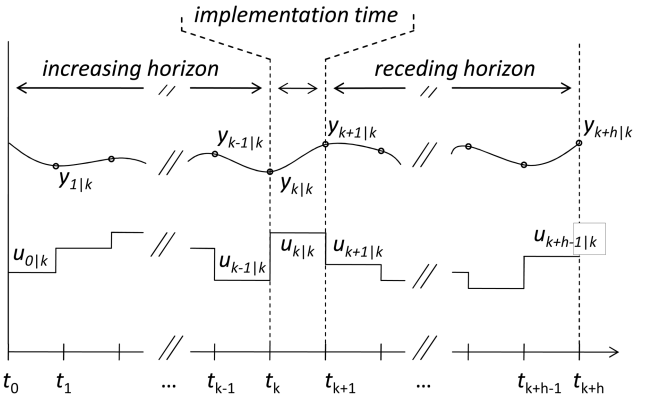

Figure 1. Discretization grids and time horizons used in the online algorithm.

- **Stopping criteria**

The typical objective of ED is to reduce the experiment duration until a specified accuracy of the determined parameters is reached. However, a different objective could be the maximization of the number of model parameters which can be identified with a specific experimental set-up.

Finite time horizon schemes

The implementation of the online PE and optimum ED is based on the adaptation of known concepts from MPC (see, e.g., Ref. 26). We consider a repeated update of current parameter estimates (solution of the PE problem) and planned input actions (solution of the optimum ED problem). Limited time horizons are defined and measurement data are assumed to be taken at equally spaced sampling times. Moreover, input actions are implemented as piece-wise constant trajectories which match the time grid defined by the measurement sampling times.

During an experiment run, for any sampling time instant t_k, a new measurement point $y_k^m = y^m(t_k)$ is obtained. Considering all prior measurements at instants $t_1, \cdots, t_{k-1}$, the actual measurement vector reads $Y_k^m = [y_1^m, y_2^m, \cdots, y_{k-1}^m | \, y_k^m]^{\mathrm{T}}$. According to the growing number of elements in the measurement vector with an ongoing experiment time, the time horizon $[t_0, t_k]$ is referred to here as *increasing horizon*. In Figure 1, a discretization grid with elements of length $\Delta t = (t_k - t_{k-1})$ is shown which meets the measurement sampling instants t_k. In every instant t_k, all past measurements Y_k^m are used to update the current parameter estimate θ_k. This is done by fitting the vector of simulated response variables $Y_k^- = [y_{1|k}, y_{2|k}, \cdots, y_{k|k}]^{\mathrm{T}\dagger}$ to the measurement vector Y_k^m (solution of problem Eq. 14). The simulated response variables Y_k^- are obtained from the solution of Eq. 1 for $t = [t_0, t_k]$ and taking into account the discrete input actions in all past intervals within the increasing horizon $U_k^- = [u_{0|k}, u_{1|k}, \cdots, u_{k-1|k}]^{\mathrm{T}}$ as well as the initial states $x_0 = x(t_0)$.

In the same way, for any sampling instant t_k, a prediction or *receding horizon* $[t_{k+1}, t_{k+h}]$ is considered, where the predicted outputs $Y_k^+ = [y_{k+2|k}, y_{k+3|k}, \cdots, y_{k+h|k}]^{\mathrm{T}}$ are obtained from the solution of Eq. 1 for $t = [t_{k+1}, t_{k+h}]$ and taking into account the future discrete input actions $U_k^+ = [u_{k+1|k}, u_{k+2|k}, \cdots, u_{k+h-1|k}]^{\mathrm{T}}$ and the initial states $x_0 = x(t_{k+1})$. All future input actions U_k^+ are updated for each t_k by the

†The notation $y_{i|j}$ indicates the values of the predicted variable vector at sampling instant i which is calculated at instant j.

solution of an ED problem and based on the current parameter estimate.

According to the iterative implementation strategy, the estimation of the parameters as well as the generation of new input actions are repeated at each time instant t_k. Thus, all corresponding computations have to be finished within one sampling interval Δt. In Figure 1, this sampling interval is denoted as *implementation time*. It has to be noted that during the implementation time, the corresponding input actions $u_{k|k}$ are implemented so these are not considered in the ED problem formulation (update of the future input actions in the time horizon $[t_{k+1}, t_{k+h}]$). Moreover, the predicted output $y_{k+1|k}$ is not considered in the PE problem (update of the current parameter estimate using measurement from the time horizon $[t_0, t_k]$). However, initial states $x_0 = x(t_{k+1})$ are needed for the definition of the ED problem. The initial states $x(t_{k+1})$ are obtained from a simulation step, solving (Eq. 1) for $t = [t_k, t_{k+1}]$ with $u_{k|k}$, $\hat{\theta}_k$ and $x_0 = x(t_k)$.

Discrete formulation of the PE and ED problem

In the adaptive online ED for PE at each sampling instant t_k, a PE as well as ED problem is solved to improve the current parameter estimate $\hat{\theta}_k$ and update the planned input actions U_k^+, respectively. The discrete formulation of the PE problem in Eqs. 4 and 5 reads

$$\hat{\theta}_k = arg\min_{\theta_k} \quad \Phi_k^{ML}(U_k^-, \theta_k) \tag{14}$$
$$\text{with} \quad \Phi_k^{ML} = \left(Y_k^-(U_k^-, \theta_k) - Y_k^m\right)^{\mathrm{T}} \cdot \Sigma_{y,k}^{-1} \cdot \left(Y_k^-(U_k^-, \theta_k) - Y_k^m\right)$$

where $Y_k^- \in \mathbb{R}^{\mathrm{Ny}\cdot k}$ is the vector of past outputs and Y_k^m and $\Sigma_{y,k}$ are the measurement vector and the covariance matrix of the experimental errors, respectively. Moreover, $\hat{\theta}_k \in \mathbb{R}^{\mathrm{Np}}$ is the parameter estimate at the kth time instant and $U_k^- \in \mathbb{R}^{\mathrm{Nu}\cdot k}$ are the input actions which have been implemented in all past intervals. The discrete formulation of the ED problem in Eqs. 7 and 8 reads

$$U_k^{+*} = arg\min_{U_k^+} \quad \Phi_k^{ED}(U_k^+, u_{k|k}, U_k^-, \hat{\theta}_k) \tag{15}$$
$$\text{with} \quad \Phi_k^{ED} \in \{\Phi^A, \Phi^D, \Phi^E\}$$

where $U_k^{+*} \in \mathbb{R}^{\mathrm{Nu}\cdot(h-1)}$ is the optimal input vector at the kth time instant, Φ_k^{ED} is one of the criteria described in Eq. 8 evaluated in the kth time instant, and $\hat{\theta}_k$ is the parameter estimate at time instant k. The calculation of the FIM F_k for the approximation of $\Sigma_{\hat{\theta}_k}$(Eq. 6) is calculated as

$$F_k(U_k^+, u_{k|k}, U_k^-, \hat{\theta}_k) = F_k^+(U_k^+, \hat{\theta}_k) + F_{k|k}(u_{k|k}, \hat{\theta}_k) + F_k^-(U_k^-, \hat{\theta}_k) \tag{16}$$

where the first term F_k^+ is a nonconstant contribution in Eq. 15 which depends on the future input actions U_k^+ and the second and third term $F_{k|k}$ and F_k^- are constant contributions which depend on currently implemented and past input actions, respectively. Anyhow, all contributions have to be updated for the current parameter estimate $\hat{\theta}_k$.

Algorithm

In Figure 2, the algorithmic implementation for the online redesign of experiments for parameter determination is presented. The iteration counter k represents the kth time or sampling instant. **Initialization:** before the start of an experiment, an initial parameter vector estimate $\theta_{k=0} \in \mathbb{R}^{\mathrm{Np}}$, as well as an initial ED has to be chosen. The latter is defined by the following variables: the length of the receding time horizon h, an initial trajectory of input actions $U_{k=0}^+ = [u_{1|0}, u_{2|0}, \cdots, u_{h-1|0}]^{\mathrm{T}} \in \mathbb{R}^{\mathrm{Nu}\cdot(h-1)}$ for the receding horizon, and finally, a first input action $u_{0|0} \in \mathbb{R}^{\mathrm{Nu}}$. **Start experiment:** the first input actions $u_{0|0} \in \mathbb{R}^{\mathrm{Nu}}$ are implemented. **Get measurement:** for every instant t_k, the actual measurement y_k^m is used to update the current measurement vector $Y_k^m = [y_1^m, y_2^m, \cdots, y_{k-1}^m | y_k^m]^{\mathrm{T}} \in \mathbb{R}^{\mathrm{Ny}\cdot k}$. **Implement next action:** the input actions $u_{k|k-1} \in \mathbb{R}^{\mathrm{Nu}}$ are implemented. They are active for the current implementation time $t = [t_k, t_{k+1}]$. Note that the values of $u_{k|k-1}$ have been calculated during the last implementation time $t = [t_{k-1}, t_k]$. **Do parameter SsS:** based on the last parameter estimate $\theta_{k-1} \in \mathbb{R}^{\mathrm{Np}}$, the vector of simulated response variables $Y_k^-(U_k^-, \theta_{k-1}) \in \mathbb{R}^{\mathrm{Ny}\cdot k}$ as well as the corresponding sensitivity matrix $S_k^-|_{\theta=\theta_{k-1}} \in \mathbb{R}^{\mathrm{Ny}\cdot k \times \mathrm{Np}}$ are generated. Thereafter, a parameter SsS is carried out, see section "Parameter Subset Selection (SsS)." The result is a new ordered vector of parameter estimates $\tilde{\theta}_k = [\tilde{\theta}_k^{(\tilde{r}_k)^{\mathrm{T}}}, \tilde{\theta}_k^{(\mathrm{Np}-\tilde{r}_k)^{\mathrm{T}}}]^{\mathrm{T}}$, where the $\tilde{r}_k$ first elements of $\tilde{\theta}_k \in \mathbb{R}^{\mathrm{Np}}$ construct the identifiable parameter subset at instant t_k, being the so-called active parameters $\tilde{\theta}_k^{(\tilde{r}_k)} \in \mathbb{R}^{\tilde{r}_k}$. All remaining parameters $\tilde{\theta}_k^{(\mathrm{Np}-\tilde{r}_k)} \in \mathbb{R}^{\mathrm{Np}-\tilde{r}_k}$ are not identifiable and, thus, considered as constant variables. **Exists an identifiable parameter subset?: yes:** new parameter estimates as well as an update of the input actions are calculated (solution of the PE and ED problem), if parameter identifiability is given for at least a subset of the parameter space ($\tilde{r}_k > 0$). **No:** if the dimension of the active parameter subset equals zero ($\tilde{r}_k = 0$), the current loop is suspended and the algorithm waits until the next instant $t = t_{k+1}$. The parameter estimates remain unchanged and the input actions are taken from prior solutions of the ED problem. **Are additional input actions available?: yes:** Go to the next loop with $k = k + 1$. **No:** if the condition $\tilde{r}_k > 0$ is not fulfilled repeatedly, there might be no more input actions available. This happens when all $h - 1$ prior actions defined over the receding horizon have been implemented but no new actions have been computed. The experiment has to be stopped unsuccessfully. This problem would most likely occur directly after the start of an experiment. To avoid this, the initial ED might be computed off-line and its suitability should be verified by an parameter identifiability analysis. **Solve the PE problem for the identifiable parameter subset:** the values of the identifiable parameters $\tilde{\theta}_k^{(\tilde{r}_k)}$ are updated to $\hat{\tilde{\theta}}_k^{(\tilde{r}_k)}$ by fitting the simulated response variables Y_k^- to the corresponding measured variables Y_k^m for all sampling instants $t_1, \cdots, t_k$, see Eq. 14. The nonidentifiable parameters $\theta_k^{(\mathrm{Np}-\tilde{r}_k)^{\mathrm{T}}}$ are held constant. **Stopping criterion satisfied?:** there exist several possible criteria to decide if the experiment can be stopped or needs to be continued. A possible criterion might be the accuracy of the active parameter subset, which can be evaluated using an ED criterion (see Eq. 8). However, the quality of the fitting as well as the number of identifiable parameters might be of interest as well. The selection of the stopping criterion certainly depends on the specific process model and its use. **Yes:** the experiment can be stopped. **No:** the experiment needs to be

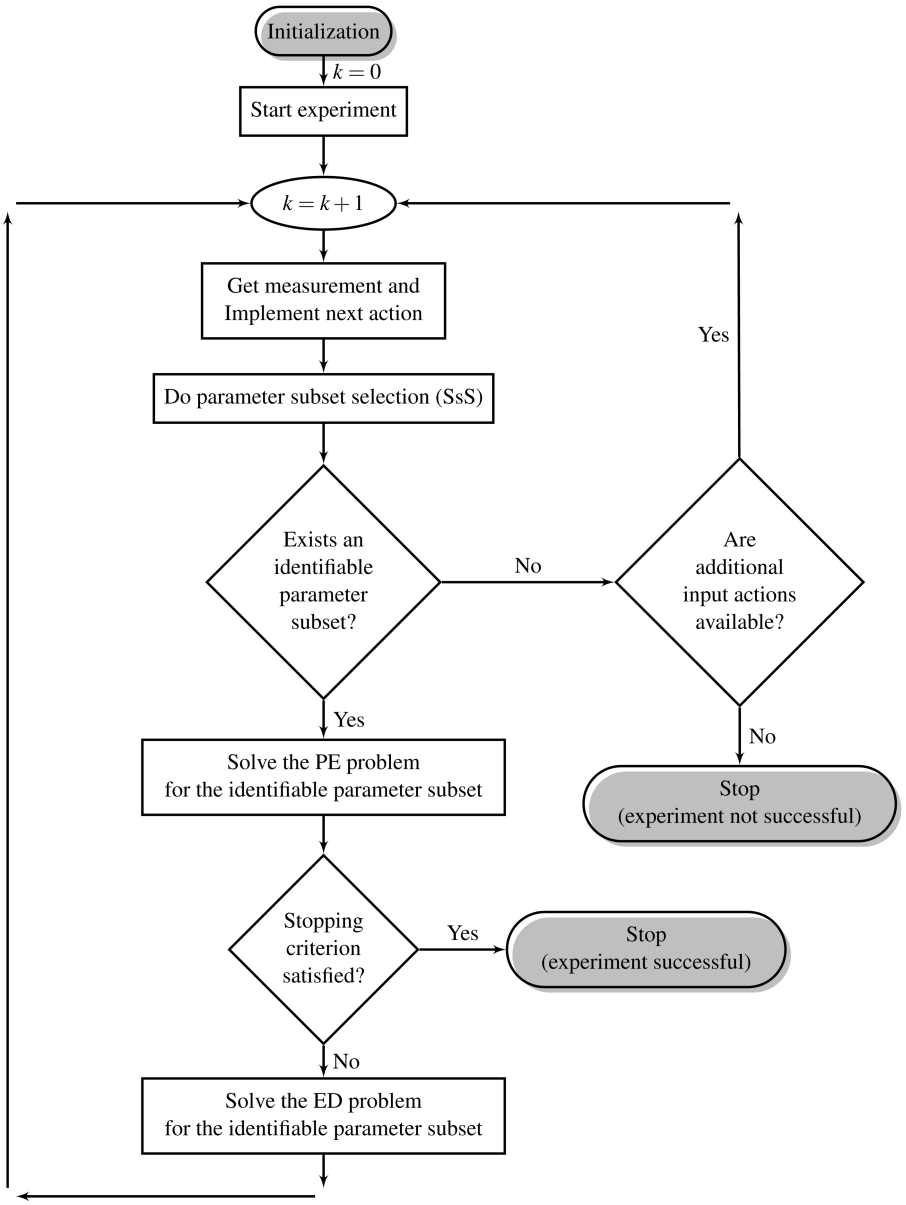

Figure 2. Algorithm for the online redesign of experiments for parameter determination.

continued. **Solve the ED problem for the identifiable parameter subset:** the future input actions $U_k^+ \in \mathbb{R}^{\mathrm{Nu}\cdot(h-1)}$ are updated by solving an ED problem whose definition is based on the analysis of the different contributions to the FIM F_k in the past ($[t_0, t_k]$), the current ($[t_k, t_{k+1}]$), and the future time horizon ($[t_{k+1}, t_{k+h}]$), see also Eqs. 16 and 15. Note that the first computed input actions $u^*_{k+1|k}$ of U_k^{+*} will be implemented in the next iteration ($k = k + 1$) as $u_{k|k-1}$. Finally, the iteration counter k moves to the next time instant, $k = k + 1$.

All computations have to be finished within the implementation time $\Delta t = t_{k+1} - t_k$, see Figure 1. Most demanding are certainly the nonlinear optimization problems, which require several iterations until a prespecified convergence criterion is satisfied. To avoid this, we take over the idea of a reduction of the calculation time by performing just one iteration per instant k and initializing the PE and ED problem by their respective prior solution.[25] This approach has been applied because the sampling instants are frequent compared to the dynamics of the system used in this case study. Thus, consecutive problems and their solution do not differ much. Given these small differences in the solution of consecutive optimization problems, the results obtained after only one iteration are sufficiently close to the optimal solutions. For the same reasons, we assume that the identifiable parameter subset does not change after the update of the parameter values (solution of the PE problem). Thus, the same identifiable parameter subset is used in the PE and ED problem.

Case Study

Chromatography is an important general-purpose separation process. Most applications are reported from pharmaceutical industry for the separation of fine chemicals at the preparative scale, for example, enantiomers, proteins, and peptides.[27] The hardware used in liquid chromatography typically has the capability for feedback control, including the pumps, the mixing of concentration gradients, and online concentration analysis. Recently, the model-based optimization, monitoring, and

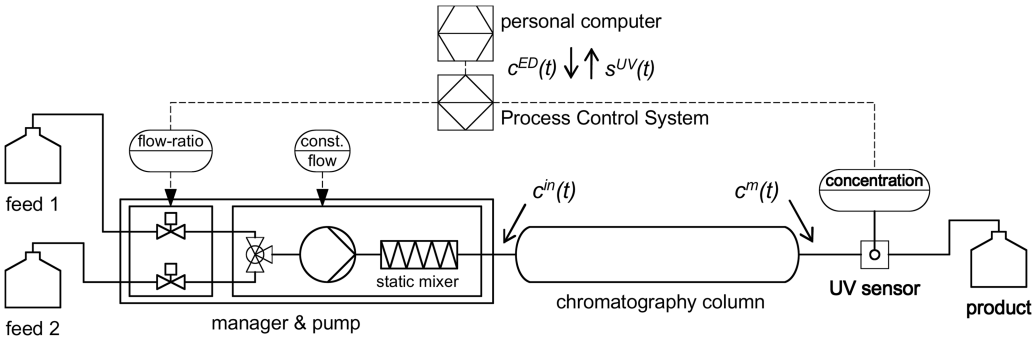

Figure 3. Flow diagram and instrumentation.

control of chromatography processes has gained the attention, especially the simulated moving bed process, see for example, Refs. 28 and 29. Central to the modeling of these processes are the thermodynamics of the phase equilibrium as well as dispersion and mass-transfer phenomena. The determination of the corresponding parameters are the subject of this case study.

Experiments

Experimental Set-Up. The laboratory scale high-performance liquid chromatography (HPLC) system is built up of a Smartline Manager 5050 with low-pressure gradient module operated in combination with a Smartline Pump 1050 (both from Knauer GmbH, Berlin, Germany). The pump heads are made from stainless steel and designed for a maximum flow of 10 mL/min. The system is used to deliver gradient mixtures from two different feeds and is equipped with a serial interface RS-232 for external communication. The separation column is a Kronlab Laboratory Scale Glass Column ECO10/120V0V (YMC Europe GmbH, Dinslaken, Germany) with an internal diameter of 10 mm, a maximal column length of 120 mm, with two adjustable plungers, and a frit porosity of 10 μm.

A continuous measurement of the liquid protein concentrations at the column outlet is realized using the Smartline ultraviolet (UV) Detector 2500 equipped with a deuterium lamp (Knauer GmbH, Berlin, Germany). The sensor is operated at a wavelength of 280 nm, which has been shown to give sound results.[30–33] The analog output signal is scaled to a maximum value of 2 AU and is used for data processing in the ABB Freelance Controller AC 700F using the Analog Input/Output Module AX 722F (ABB, Zurich, Switzerland). The errors of the output signal are assumed to be Gaussian and uncorrelated with a standard deviation $\sigma = 0.05$ AU.

All numerical computations are implemented in the programming environment Matlab 2011b (MathWorks, Natick, MA). Communication between Freelance controller and Matlab software is established using an Open Process Control software interface. The Matlab Instrument Control Toolbox is used for communication with the Smartline Manager and the Pump over serial interface RS-232.

Materials. The column is packed with the strong anion exchanger, Source 30Q, product of GE-Healthcare (Munich, Germany, product code: 17-1275-01). The length of the packed adsorbent bed is 11.2 cm, its diameter 1.0 cm, and the resulting total volume 8.80 mL. It is composed of rigid, monodispersed, spherical resin particles with a diameter of 30 μm. For the pore diameter, values of $\sim$10–150 nm were found[34] and for the intraparticular porosity a value of $\varepsilon_{\text{int}} = 0.589$.[33] The value for the total porosity $\varepsilon_{\text{tot}} = 0.76$ was determined by pulse experiments under nonbinding conditions (see Refs. 30–32 for detailed information). The obtained value shows good agreement with the value of $\varepsilon_{tot} = 0.74$ published in Ref. 30. The external or interstitial porosity $\varepsilon_{ext} = 0.4$ is then obtained from the volume balance, with $\varepsilon_{\text{ext}} = (\varepsilon_{\text{tot}} - \varepsilon_{\text{int}})/(1 - \varepsilon_{\text{int}})$. The columns total ionic resin capacity $\Lambda = 706$ mM (mM = mmol Cl^-/L solid phase) was measured by saturating the column with $NaNO_3$ following the procedure in Refs. 30 and 31. The protein bovine serum albumin (BSA) was purchased from Sigma-Aldrich Chemie GmbH (Taufkirchen, Germany) as pure lyophilized powder (product code: A7960) with a purity $\geq$98%.

The eluent was prepared in 20 mM (mmol/l) bis-tris-propan buffer at a pH = 7.0. The Cl^--ion concentration of the solvent has been implemented by adding at first a fixed amount of 5M HCl. It should be noted that the pH is then below 7 (pH = 6). Subsequently, an additional amount of NaCl is added to reach the desired total salt (Cl^-) ion concentration. The pH was adjusted adding 5M NaOH and the desired eluent is then obtained. This procedure was previously used and validated.[30,34,35] It should be noted that the stated ion concentration refers always to the total Cl^--ions in the eluent.

Operating Conditions. All experiments were carried out at constant temperature of 23°C and a feed flow rate of 2 mL/min with an eluent pH value of 7 with a chloride concentration (CL^-) of 120 mM and variable feed concentrations of BSA varying from 0 to 0.2 mM.

Process model

In this case study, the ED is defined by the protein concentration of the eluent stream which is given into the column. Figure 3 shows the flow diagram and instrumentation. A protein feed concentration $c^{ED}(t)$ is computed online (personal computer) and is realized by the process control system and the ratio controller (manager) which brings the liquid from the two different feeds (feed 1 and 2) together. Before the feed mixture $c^{in}(t)$ enters the chromatography column, it passes a static mixer which is part of the pump. The liquid protein concentration at the column outlet $c^{m}(t)$ is measured by a UV sensor which outputs a signal $s^{UV}(t)$ which is then processed by the personal computer. The process model comprises the manager and the pump, the chromatography column, and the UV sensor. The equations of each of these process units are given in Appendix A. The connection of the units by streams and the respective variables are shown in Figure 4.

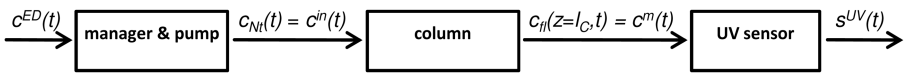

Figure 4. Units of the process model and connecting variables.

For detailed information see Appendix A.

Formulation of the online PE and ED problem

The PE problem is defined for the determination of the chromatography column parameters, namely the axial dispersion coefficient D_{ax}, the mass-transfer parameter β, and the adsorption parameters a, b, as well as the calibration curve parameters m_0, m_1, m_2. Parameters for the dynamic behavior of the manager and pump are not part of the PE problem. The dynamics of the manager and pump do not change for specific components processed by the chromatography system and, thus, the corresponding parameters have been determined independently in prior experiments.

According to the definitions in sections "Problem Formulation" and "Online Parameter Estimation and Adaptive Redesign of Experiments," the parameters θ_k, the input actions $u_{j|k}$, and measured states y_k^m are defined as

$$\theta_k \equiv \left([D_{ax}, \beta, a, b, m_0, m_1, m_2]^{\mathrm{T}} \right)_k \tag{17}$$

$$u_{j|k} \equiv \left(c^{ED}\right)_{j|k}; \qquad \text{with} \quad 0.0 \text{ mM} \le u_{j|k} \le 0.2 \text{ mM}$$

$$y_k^m \equiv \left(s^{UV}\right)_k$$

The initial parameter guess $\theta_{k=0}$ is defined as shown in Table 1. Measurements $y_k^m = y^m(t_k)$ are taken at equally spaced sampling times t_k, with $t_k - t_{k-1} = 15$ s. The receding time horizon including the time for implementation is $t_{k+h} - t_k = 10$ min. The PE and ED problems are both solved at every instant t_k to update the parameter values θ_k and future input actions U_k^+ (see section "Algorithm"). In the ED problem, all input actions have been held constant for more than one sampling time (four times for 1 min and once for 5 min), such that

$$\begin{aligned} u_{k+1|k} &= u_{k+2|k} = \cdots = u_{k+4|k} \\ u_{k+5|k} &= u_{k+6|k} = \cdots = u_{k+8|k} \\ u_{k+9|k} &= u_{k+10|k} = \cdots = u_{k+12|k} \\ u_{k+13|k} &= u_{k+14|k} = \cdots = u_{k+16|k} \\ u_{k+17|k} &= u_{k+18|k} = \cdots = u_{k+20|k} \\ u_{k+21|k} &= u_{k+39|k} = \cdots = u_{k+40|k} \end{aligned} \tag{18}$$

The specific discretization of the receding horizon ED problem is depicted in Figure 5. Finally, the overall experiment time has been set to $t_f = 1$ h.

Numerical solution

The process model described in Appendix A is given as nonlinear partial DAE system. The dynamic PE and ED problems are solved by single shooting. For this, the partial DAE system is discretized in space first and the resulting differential algebraic equation system is then integrated using an initial value solver. For an efficient solution, exact first and second-order gradients are generated.

Solution of the Dynamic Process Model. Eq. A2 is efficiently discretized in axially (z) direction by the Galerkin finite element method together with second-order polynomials.[36] The result is an implicitly defined differential equation system which has to be solved for each discrete point simultaneously with (Eqs. A4 and A5). The solution of the PE problem (Eq. 4) demands numerous simulations with different parameter values. Thus, a stable numerical solution is crucial without oscillations even for relatively sharp concentration profiles in time and space. For the axial discretization of Eq. A2, the number of discrete elements is set to Nde = 35. Together with the 12 linear differential equations of the manager and pump model and the nonlinear algebraic equation for the sensor calibration, the total number of equations is then $\mathrm{Neq} = (\mathrm{Nde} \cdot 2 + 1) \cdot 2 + 12 + 1 = 155$.

The linear implicit DAE system of the type shown in Eq. 1 is numerically solved using the initial value solver sDACl, a sparse DAE solver based on the orthogonal collocation on finite elements method.[37,38] The integration is performed using an orthogonal collocation discretization along with the element-wise solution of the discretized nonlinear equation systems. First- and second-order sensitivities can be generated which are exact for the approximate system solution calculated by sDACl.

Use of First and Mixed Second-Order Sensitivities. The analysis of the parameter accuracy (Eq. 6) as well as the formulation of the objective function in problem (Eq. 7) relies on the accurate determination of the sensitivity matrix $S(u, \theta)$. Moreover, gradient-based optimization methods are applied for the efficient solution of the PE and ED problem. The performance of these methods depends heavily on the accuracy of the derivative information of the problem functions, here the objective functions in Eqs. 4 and 7. We provide exact first-order derivatives. Algorithms for the solution of nonlinear PE problems accept residual vectors for the problem definition and the corresponding first-order derivatives are defined by the matrix $S(u, \theta)$. In contrast, the ED problem (Eq. 7) is defined by a scalar valued objective function $\Phi^{ED}(u, \theta)$ and first-order derivatives are defined as $\partial \Phi^{ED}(u, \theta)/\partial u$. As the evaluation of $\Phi^{ED}(u, \theta)$ is based on $S(u, \theta)$ (see (Eqs: 6–8), second-order derivative information is needed to generate the first-order problem derivatives $\partial \Phi^{ED}(u, \theta)/\partial u$. Strictly speaking, mixed second-order derivatives $H_{u,\theta} \in \mathbb{R}^{\mathrm{Ny} \cdot \mathrm{Nm} \cdot \mathrm{Np} \times \mathrm{Nu}}$ (see Eq. 19) have to be generated

$$H_{u,\theta}(u, \theta) = \frac{\partial S(u, \theta)}{\partial u} = \frac{\partial^2 \hat{y}(u, \theta)}{\partial u \partial \theta} \tag{19}$$

Table 1. Initial Parameter Guess $\theta_{k=0}$ for all Conducted Experiments

Variable	D_{ax} (cm^2/min)	β (cm/min)	a (−)	b (−)	m_0 (−)	m_1 (−)	m_2 (−)
Value	0.020	1.0E-6	0.0	1.0	−0.9	0.8	0.02

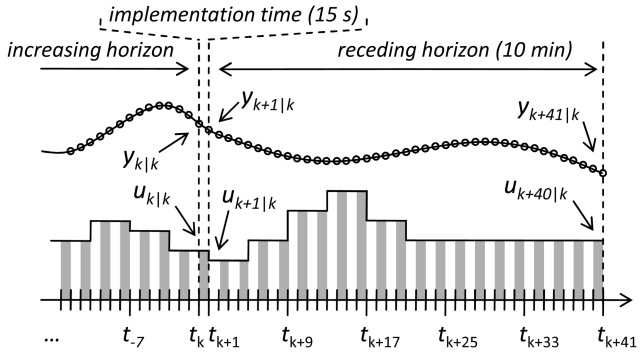

Figure 5. Discretization of the receding horizon ED problem.

For a detailed description of the computation of $\partial\Phi^{ED}(u,\theta)/\partial u$ for the design criteria in Eq. 8 using $\boldsymbol{S}$ and $H_{u,\theta}$, the interested reader is referred to Appendix B.

Computation. All computations were carried out on a 32-bit Linux platform with an Intel(R) Core(TM)2 (CPU 6600 @ 2.40-GHz) computer with 4-GB RAM. Parallel programming was not used. The C++ implementation of the integrator sDACl[38] was interfaced with Matlab using mex functions. The PE and ED problems were solved using Matlab's Optimization Toolbox solvers *lsqnonlin/ trust-region-reflective* and *fmincon/sqp*, respectively. To limit the computational effort, the number of maximum iterations of both solvers was restricted to one. The solvers were restarted at each new sampling time point t_k and the results from the last solver call were used to define the initial values of the PE and ED problem. This repeated one-step solution approach was verified by off-line computations at different time points of the experiment run. They showed that the solutions of the PE and ED problem obtained from a solver call with no limit on the number of maximum iterations are very close to the solutions calculated repeatedly online. This is because in the online solution the problems of adjacent sampling intervals do not differ much and the repeated one-step solution approach refines prior solutions rather than recalculating them (see also Ref. 25). Computation times for one system integration and sensitivity generation were ~1.3 s, which was sufficient for the solution of the PE, ED, and SsS problem within the implementation time $\Delta t = t_{k+1} - t_k = 15s$.

Experimental Results

A total number of six experiments (E1, …, E6) were carried out, they differ in the way the ED is calculated. In contrast to the algorithm depicted in Figure 2, no specific termination criterion was used. Instead, for all experiments, the total experiment duration was set to $t_f = 60$ min.

The first experiment E1, see Figure 6, was conducted to verify the online implementation of the SsS and PE algorithm, the ED problem was not solved. In Figure 6, the first plot (counted from above) shows the fitting of the measured signal from the UV sensor s^{UV}: the predicted output variables $y_{k|f}$ (continuous line) and the measured variables y_k^m (circles), for all time instants $k = 1, 2, \cdots, f$ and $t_f = 60$ min. The second plot shows the ED design variables c^{ED} which were generated by an uniform sampling. Depicted are the realized discrete input actions $u_{k|k-1}$, with $k = 0, 1, \cdots, f-1$. Note that, according to the definitions in (Eq. 18), these actions were held piece-wise constant for 1 min. The third plot shows the estimated parameter values $\theta_{i,k}$, with $i = 1, 2, \cdots, \mathrm{Np}$ and $k = 0, 1, \cdots, f$. The fourth plot shows the subset dimension r_k with $k = 0, 1, \cdots, f$ which is the number of active parameters used in the PE and ED problems and the fifth plot shows when the individual parameters were active. The sixth shows the average residual of the PE problem calculated as weighted L_1 norm $1/k \cdot \sum_{i=1}^{k} \mathrm{abs}\ (y_i^m - y_{i|f})$.

In Figure 6, it can be seen that the response of the measured variable y_k^m to changes in the planned input $u_{k|k-1}$ is dominated by a relative large dead time of around $\Delta k = 20 (\Delta t = 5$ min). However, after the first $k = 20 (t_k = 5$ min) experiment run, the first two parameters θ_2, θ_5 and after $k = 40 (t_k = 10$ min) already four parameters $\theta_2, \theta_3, \theta_5, \theta_6$ are identifiable and, thus, selected for PE. Although the number of these active parameters does not change until the end of the experiment $k = f = 240 (t_f = 60$ min), their state (active/ non-active) does change, see θ_5 and θ_7, approximately around $k \approx 60 (t_k \approx 15$ min). At this time, most parameter values do change significantly (see third plot in Figure 6) to minimize the residual from the PE problem (see sixth plot in Figure 6).

Results for D- and E-optimal experiments (Experiment E3 and E4, respectively) are shown in Figures 7 and 8, where the SsS, PE, and ED problem are solved online. In both figures, the seventh plot (counted from above) shows the value of the ED criterion Φ_k^{ED} with $k = 1, 2, \cdots, f$. It can be seen that for both design criteria the input actions $u_{k|k-1}$ adopt their external values and pulses of different length are formed. There are $r_k = 6$ active parameters in E3, compared with $r_k = 5$ in E4. Additionally, at instant $k = 140$

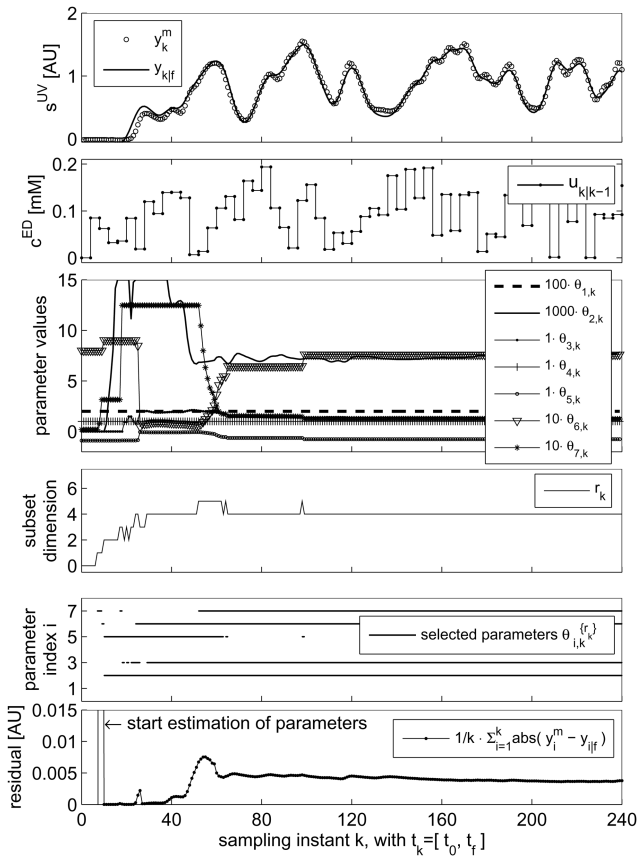

Figure 6. Experiment E1: ED by uniform sampling (an ED problem was not solved).

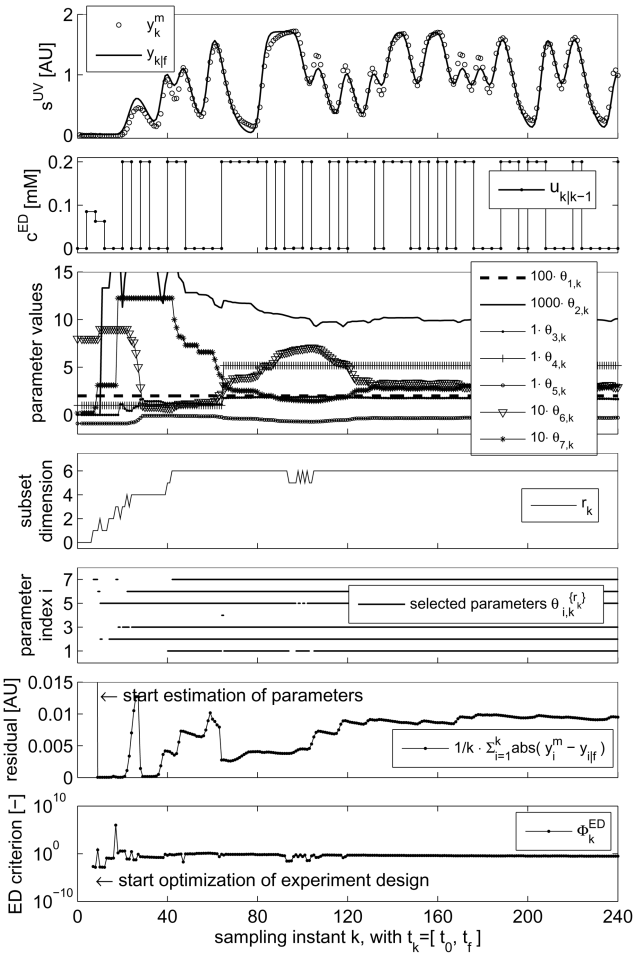

Figure 7. Experiment E3: D-optimal design.

($t_k = 35$ min), there are only $r_k = 3$ active parameters in E4, whereas in E3, we find $r_k = 6$. Moreover, especially between $k = 60$ and $k = 160$, there are large differences in the estimated parameter values between the two experiments. The reason for these large differences is certainly the limited number of active parameters, $r_k = 3$, in the PE problem of E4. Accordingly, the fitting is poor (see the residuals between $k = 60$ and $k = 160$ in the sixth plot in Figure 8). The $r_k = 3$ active parameters are used to compensate all other $\mathrm{Np} - r_k = 4$ nonactive and imprecisely estimated parameters. The different parameter values in E3 and E4 also lead to different solutions of the online ED problem which then causes a different process behavior (compare the first and second plot in Figures 7 and 8 between $k = 60$ and $k = 200$).

The final results at $k = f = 240 (t_f = 60$ min) of all six experiments (E1, …, E6) are summarized in Table 2. The ED in E1 was generated by an uniform sampling of the decision variables $u_{k|k-1}$ (an ED problem was not solved). E2, E3, and E4 were carried out with an online solution of the ED problem using A, D, and E-optimality criteria. In E2, E3, and E4, the initial ED variables $U_{k=0}^{+}$ (see the algorithm in Figure 2) were defined using the ED variables of E1. E5 and E6 were carried out to study the influence of the initial ED variables $U_{k=0}^{+}$ on the final result. To do so, the A-optimal experiment E2 was repeated twice (E5, E6). This means E2, E3, and E4 were initialized with the same $U_{k=0}^{+}$, whereas E2, E5, and E6 were initialized with different $U_{k=0}^{+}$ generated by an uniform sampling.

The algorithm worked robustly for all adopted ED criteria with no convergence problems during optimization. In terms of the residual of the PE problem, the worst results are obtained for E2, where a limited degree of freedom of $r_k = 4$ active (identifiable) parameters is used to fit the model prediction to the measured signal from the UV sensor y_k^m which represents an A-optimal ED. In contrast, in E1 where the same $r_k = 4$ parameters are active, the residual is much smaller. Moreover, due to the nonoptimal ED in E1, the excitations in the measurement y_k^m are less significant (less steep concentration profiles), and thus, easier to fit (compare also first plot in Figure 6 and first plots in Figures 7 and 8). In Table 2, the estimated parameters and their respective standard deviations are given (standard deviations are depicted in brackets and calculated from the diagonal elements of the parameter covariance matrix Σ_θ). In terms of the accuracy of the estimated parameters, E2 (with the minimum number of active parameters) gives the best results of all conducted experiments and independent of the adopted ED criterion. In terms of the number of active (identifiable) parameters, E3 and E5 are most successful, with $r_f = 6$. However, those experiments with a small number of active parameters are generally better conditioned. This can be seen from the condition numbers $\kappa(S_f^{(r_f)})$ and collinearity indexes $\gamma(S_f^{(r_f)})$ in experiments E2 ($r_f = 4$), E6 ($r_f = 5$), and E5

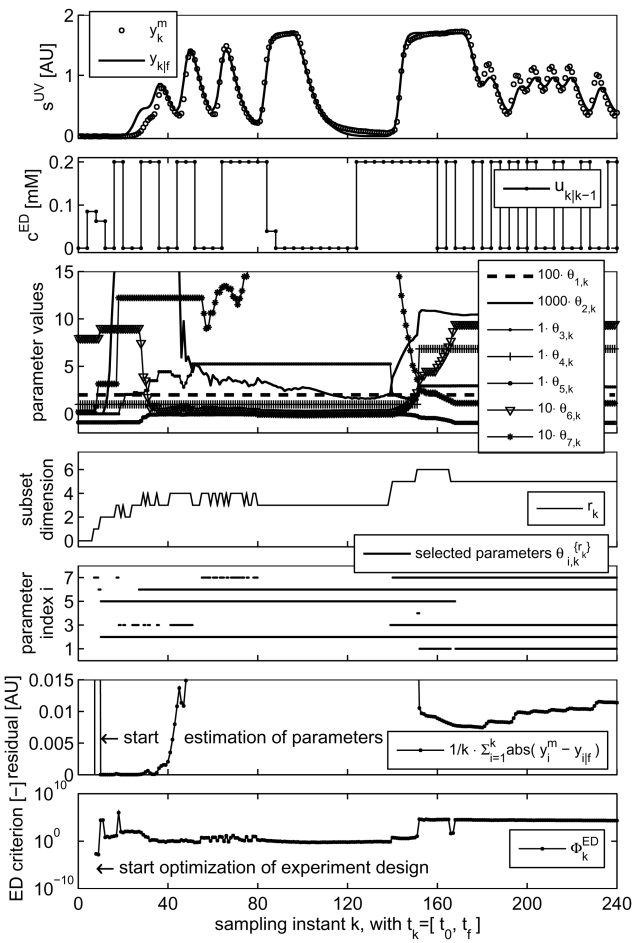

Figure 8. Experiment E4: E-optimal design.

Table 2. Final Results for all Experiments at $k = 240$ ($t_k = 60$ min)

Experiment Label	ED Criterion (Initial ED)	ED Criterion Value Φ_f^{ED}	Subset Dimension r_f	Condition Number $\kappa(S_f^{(r_f)})$	Colinearity Index $\gamma(S_f^{(r_f)})$	Mean Residual PE*	Active/Nonactive × Parameter Values (std. deviations) at $t_f = 60$ min					
							D_{ax} (cm^2/min)	k_{eff} (cm/min)	a (−)	b (−)	m_0 (−)	m_1 (−)	m_2 (−)
E1 (see Fig. 6)	**NONE** ED by uniform sampling	–	4	79.1	7.4	0.0038	0.020†	0.0079 (0.0061)	1.35 (0.33)	1.00†	−0.75†	0.76 (0.01)	0.12 (0.02)
E2	**A-criterion** (initial ED as in E1)	0.04 (A)	4	94.6	5.6	0.0208	0.020†	0.0062 (0.0024)	0.92 (0.20)	1.00†	−0. 57†	0.58 (0.01)	0.16 (0.02)
E3 (see Fig. 7)	**D-criterion** (initial ED as in E1)	0.30 (D)	6	276.6	5.8	0.0095	0.005 (0.179)	0.0101 (0.0047)	1.66 (0.31)	5.17†	−0.29 (0.61)	0.29 (0.60)	0.31 (0.51)
E4 (see Fig. 8)	**E-criterion** (initial ED as in E1)	21008.23 (E)	5	279.9	7.2	0.0114	0.007 (1.14)	0.0095 (0.0110)	2.83 (0.41)	6.84†	−0.94†	0.94 (0.01)	0.11 (0.01)
E5	**A-criterion** (initial ED by uniform sampling)	5711.56 (A)	6	247.5	7.8	0.0103	2.6E-14 (0.022)	0.0102 (0.0058)	2.53 (0.31)	3.57 (3.65)	−0.90†	0.89 (0.01)	0.13 (0.02)
E6	**A-criterion** (initial ED by uniform sampling)	2726.90 (A)	5	364.7	7.3	0.0099	0.013 (1.575)	0.0107 (0.0187)	2.14 (0.31)	5.96†	−0.92†	0.92 (0.01)	0.12 (0.02)

*$1/f \cdot \sum_{i=1}^{f} \mathrm{abs}(y_i^m - y_{i|f})$
†Nonactive (nonidentifiable) parameters were fixed and neglected during PE and ED calculations, respectively.

($r_f = 6$) which have been generated using the same ED criterion. However, the adopted ED criterion also influences on the values of $\kappa(S_f^{(r_f)})$ and $\gamma(S_f^{(r_f)})$ (compare E4 and E6 as well as E3 and E5) as well as the number of active parameters r_k. It can be seen that different numbers of active parameters result for all experiments. From the experimental results, it cannot be concluded if a certain ED criterion is better suited to maximize the number of active parameters. However, for the repeated A-optimal experiments E2, E5, and E6 which have been started using different initial ED variables $U_{k=0}^{+}$, different numbers of active parameters are found with $r_f = 4, 6$, and 5. This is a clear indication of a low robustness of the algorithm with respect to the number of active (identifiable) parameters. From the analysis of the different experiments, it becomes clear that the planned experiments and, thus, also the selection of active parameters are strongly affected by general uncertainties and errors: in the initial parameter guess $\theta_{k=0}$, the measurement y_k^m, and the implementation of the planned input actions $u_{k|f}$.

Summary and Conclusions

A methodology for the online redesign of experiments has been presented and experimentally tested by application to a chromatography column system. The potential of the ED techniques for an improved reliability of the estimated parameters has been widely accepted. In this contribution, special focus is placed on the development of an framework for the online solution of the PE and ED tasks, which are discretized according to the sampling of the measurement data. The result is a repeated step-wise solution of the PE and ED problem, where the current parameter estimate and the planned input actions are updated at every sampling instant. The discrete formulation of the PE problem and especially the ED problem leads to complex NLP problems which have to be solved in a restricted time defined by the sampling instants. Thus, good robustness and convergence properties as well as a limited computational cost for solving these NLP problems are mandatory. To achieve these goals, various issues were addressed. At every sampling instant, a local parameter identifiability analysis is performed. Subsequently, ill-conditioned parameters are fixed at prior estimates and reduced-order and well-conditioned PE and ED problems are defined. These dynamic problems are then efficiently solved by single shooting and using the initial value solver sDACl which solves the dynamic model equations and generates first and second-order sensitivities. Based on this information, the PE and ED problem functions as well as their exact first-order derivatives are computed and efficient iterative gradient-based optimization algorithms are applied. At each sampling instant, one single iteration can be performed. Finally, the complexity of the ED problem is further reduced by the consideration of a receding time horizon strategy, where only a part of all future input actions is optimized.

It has been verified in several experiments with different ED criteria that the algorithm performs well; during an experiment run, the identifiable parameters are selected and set active, their values are estimated and the ED is adapted to maximize the information in the measured response variables with respect to these active parameters.

However, it has also turned out that the algorithm is not robust regarding the number of parameters which can be identified after a certain experiment time. From the comparison of those experiments which were conducted with the same ED criterion, it can be seen that the number of identified

parameters is very sensitive to uncertainties. These uncertainties are mainly represented by: measurement errors which are especially critical in the early PE using small numbers of experimental data, errors in the realization of an optimal planned ED, structural errors in the model, and finally, errors in the initial parameter guess. These uncertainties affect the estimated parameter values during an experiment run and, thus, the local identifiability analysis. The applied standard A-, D-, and E-criteria are aimed at maximizing the information content in the ED with respect to the selected active (identifiable) parameters only. Accordingly, the accuracy of selected active parameters is improved in every time instant, while nonactive parameters are neglected. These nonactive parameters become then active only by chance. This is the reason for the lack of robustness with respect to the number of parameters which can be identified after a certain experiment time. This situation is certainly not desirable, as the online redesign of experiments for parameter determination should provide accurate estimates for as many model parameters as possible in the shortest time. It can be concluded that the ED problem formulation in the presented algorithm should be adapted for the consideration of both identifiable and nonidentifiable parameters and additionally, a suitable regularization method should be applied to accurately solve this problem.

Acknowledgments

The authors gratefully acknowledge technical support provided by the Knauer GmbH (Berlin, Germany). This work is part of the Collaborative Research Center SFB/TR 63 InPROMPT "Integrated Chemical Processes in Liquid Multiphase Systems" coordinated by the Technische Universität Berlin and funded by the German Research Foundation. Diana C. López Cárdenas gratefully acknowledges financial support from the German Academic Interchange Service (DAAD) for a doctoral stipend.

Notation

Latin Letters

a = Langmuir coefficient, –
A = matrix of measurement equation
b = Langmuir coefficient, –
c = liquid fluid concentration of manager and pump, UV sensor, mM
c_{fl}/c_p = liquid fluid/ pore concentration, mM
C = quadratic matrix
d_P = particle diameter, cm
D_{ax} = dispersion coefficient, cm^2/min
F = Fisher Information Matrix
g^{lhs}/g^{rhs} = vector of left/right hand side functions of the differential equations
h = vector of functions of the algebraic equations
H = Hessian
I = identity matrix
K = permutation matrix
l_c = column bed length, cm
m = vector of calibration parameters
M = mass matrix of the differential equation system
M_{sc} = scaling matrix
P = permutation matrix
q = adsorbent concentration, mM
Q = orthogonal matrix
r = set dimension
R = upper triangular matrix
r_P = particle radius, cm
S = sensitivity matrix
s^{UV} = UV signal, AU
t = time, min
TOL = tolerance value
u = velocity of moving liquid, mL/min
u, U = vector of input actions
U = unitary matrix
v = eigenvector
V = unitary matrix
x = vector of state variables
y, Y = vector of predicted/measured response variables
z = axial column coordinate, cm

Greek letters

β = lumped mass-transfer coefficient, cm/min
γ = collinearity index
Γ = active set
Δ = difference
$\varepsilon_{int}/\varepsilon_{ext}/\varepsilon_{tot}$ = internal/external/total porosity,–
θ = vector of parameters
κ = condition number
λ = eigenvalue
Λ = ionic resin capacity, mM
σ = standard deviation
Σ = covariance matrix
Σ = diagonal matrix containing singular values
τ = time constant
Φ = objective function

Additional Indices and Subscripts

$\{\}_f$ = final element/ time instant
$\{\}_h$ = horizon length
$\{\}_i, \{\}_j, \{\}_k$ = indices
$\{\}_u$ = related to input actions
$\{\}_y$ = related to measured variables
$\{\}_\theta$ = related to parameters
$\{\}_0$ = initial

Additional Superscripts

$\{\}^A, \{\}^D, \{\}^E$ = different experiment design criteria
$\{\}^{ED}$ = experiment design
$\{\}^{in}$ = inlet
$\{\}^m$ = measurement
$\{\}^{max}$ = maximum value
$\{\}^{ML}$ = maximum likelihood
$\{\}^*$ = solution
$\{\}^-$ = refers to past intervals/ time instants
$\{\}^+$ = refers to future intervals/ time instants
$\hat{\{\}}$ = estimator
$\bar{\{\}}$ = normalized
$\tilde{\{\}}$ = update

Value Symbols

Nde = number of discrete elements
Neq = number of equations
Nm = number of measurement time instances
Np = dimension of the parameter vector
Nu = dimension of vector of input actions
Nt = number of linear dynamic systems
Nx = dimension of state vector
Ny = dimension of vector of predicted response variables

Abreviations

AU = Arbitrary Units
DAE = Differential Algebraic Equations
ED = Experiment Design
ETDM = Equilibrium Transport Dispersive Model
FIM = Fisher Information Matrix
HPLC = High Performance Liquid Chromatography
ML = Maximum Likelihood
mM = milli Molarity (millimoles per liter)
MPC = Model Predictive Control
NLP = Nonlinear Programming
PE = Parameter Estimation
QRP = QR decomposition with column Pivoting
sDACl = sparse Differential Algebraic equation solver based on orthogonal Collocation on finite elements
SMB = Simulated Moving Bed
SsS = Subset Selection
SVD = Singular Value Decomposition
UV = Ultraviolet

Literature Cited

1. Mehra R. Optimal input signals for parameter estimation in dynamic systems–survey and new results. *IEEE Trans Automat Contr*. 1974;19(6):753–768.
2. Walter E, Pronzato L. Qualitative and quantitative experiment design for phenomenological models—a survey. *Automatica*. 1990;26(2):195–213.
3. Franceschini G, Macchietto S. Model-based design of experiments for parameter precision: State of the art. *Chem Eng Sci*. 2007;63(19):4846–4872.
4. Barz T, Arellano-Garcia H, Wozny G. Handling uncertainty in model-based optimal experimental design. *Ind Eng Chem Res*. 2010;49(12):5702–5713.
5. Galvanin F, Barolo M, Pannocchia G, Bezzo F. Online model-based redesign of experiments with erratic models: a disturbance estimation approach. *Comput Chem Eng*. 2012;42:138–151.
6. Asprey SP, Macchietto S. Designing robust optimal dynamic experiments. *J Process Control*. 2002;12(4):545–556.
7. Walter E. *Identifiability of parametric models*. Elmsford, NY: Pergamon Press, Inc., 1987.
8. Dette H, Melas VB, Pepelyshev A, Strigul N. Robust and efficient design of experiments for the Monod model. *J Theor Biol*. 2005;234(4):537–550.
9. Körkel S, Kostina E, Bock H, Schlöder J. Numerical methods for optimal control problems in design of robust optimal experiments for nonlinear dynamic processes. *Optim Methods Softw*. 2004;19:327–338.
10. Körkel S, Bauer I, Bock HG, Schlöder JP. *A sequential approach for nonlinear optimum experimental design in DAE systems*. In: Keil F, Mackens W, Vo H, Werther J, editors. *Scientific Computing in Chemical Engineering II: Simulation, Image Processing, Optimization, and Control*. Berlin: Springer, 1999;338–345.
11. Stigter JD, Vries D, Keesman KJ. On adaptive optimal input design: a bioreactor case study. *AIChE J*. 2006;52(9):3290–3296.
12. Körkel S, Arellano-Garcia H. *Online experimental design for model validation*. In: de Brito Alves RM, do Nascimento CAO, Biscaia EC Jr, editors. *10th International Symposium on Process Systems Engineering—PSE 2009*, Vol. 27. Salvador-Bahia, Brazil: Elsevier, 2009;1–8.
13. Galvanin F, Barolo M, Bezzo F. Online model-based redesign of experiments for parameter estimation in dynamic systems. *Ind Eng Chem Res*. 2009;48:4415–4427.
14. Jayasankar B, Huang B, Ben-Zvi A. *Receding horizon experiment design with application in SOFC parameter estimation*. In: Kothare M, Tade M, Wouwer AV, Smets I, editors. *9th International Symposium on Dynamics and Control of Process Systems (DYCOPS 2010)*, Leuven, Belgium. 2010; pp. 527–532.
15. Zhu Y, Huang B. Constrained receding-horizon experiment design and parameter estimation in the presence of poor initial conditions. *AIChE J*. 2011;57(10):2808–2820.
16. Bard Y. *Nonlinear Parameter Estimation*. New York: Academic Press, 1974.
17. Walter E, Pronzato L. On the identifiability and distinguishability of nonlinear parametric models. *Math Comput Simul*. 1996;42(2):125–134.
18. Bellman R, Åström KJ. On structural identifiability. *Math Biosci*. 1970;7(3):329–339.
19. Burth M, Verghese GC, Vélez-Reyes M. Subset selection for improved parameter estimation in on-line identification of a synchronous generator. *IEEE Trans Power Syst*. 1999;14(1):218–225.
20. Velez-Reyes M, Verghese GC. Subset selection in identification, and application to speed and parameter estimation for induction machines. In: *Proceedings of the 4th IEEE Conference on Control Applications*. Albany. 1995;991–997.
21. Brun R, Kühni M, Siegrist H, Gujer W, Reichert P. Practical identifiability of asm2d parameters - systematic selection and tuning of parameter subsets. *Water Res*. 2002;36(16):4113–4127.
22. Yao KZ, Shaw BM, Kou B, McAuley KB, Bacon DW. Modeling ethylene/butene copolymerization with multi-site catalysts: parameter estimability and experimental design. *Polym React Eng*. 2003;11(3):563–588.
23. Grah A. Entwicklung und Anwendung modularer Software zur Simulation und Parameterschätzung in gaskatalytischen Festbettreaktoren. Ph.D. Thesis, Martin-Luther-Universität Halle-Wittenberg, 2004.
24. Gadkar KG, Varner J, Doyle Iii FJ. Model identification of signal transduction networks from data using a state regulator problem. *IEEE Syst Biol*. 2005;2:17–30.
25. Diehl M, Bock HG, Schlöder JP, Findeisen R, Nagy Z, Allgöwer F. Real-time optimization and nonlinear model predictive control of processes governed by differential-algebraic equations. *J Process Control*. 2002;12(4):577–585.
26. Camacho EF, Bordons C. Model predictive control. In: *Advanced Textbooks in Control and Signal Processing*. London, UK: Springer, 2nd ed. 2004.
27. Guiochon G, Felinger A, Shirazi DG, Katti AM. *Fundamentals of Preparative and Nonlinear Chromatography*, 2nd ed. San Diego, USA: Elsevier Inc., 2006.
28. Natarajan S, Lee JH. Repetitive model predictive control applied to a simulated moving bed chromatography system. *Comput Chem Eng*. 2000;24(2–7):1127–1133.
29. Toumi A, Engell S. Optimization-based control of a reactive simulated moving bed process for glucose isomerization. *Chem Eng Sci*. 2004;59(18):3777–3792.
30. Frederiksen SS. Computer aided development and optimisation of chromatographic separations. Ph.D. Thesis, Technical University of Denmark, 2004.
31. Susanto A, Wekenborg K, Hubbuch J, Schmidt-Traub H. Developing a chromatographic column model for bovine serum albumin on strong anion-exchanger Source30Q using data from confocal laser scanning microscopy. *J Chromatogr A*. 2006;1137:63–75.
32. Susanto A. Untersuchung und Modellierung intrapartikularer Stofftransportmechanismen bei der Proteinaufreinigung durch Lonenaustauschchromatographie, Vol. 865 of *VDI Fortschritt-Berichte Reihe 3 Verfahrenstechnik*. Düsseldorf: VDI Verlag, 2006.
33. Wekenborg K, Susanto A, Frederiksen S, Schmidt-Traub H. Nicht-isokratische SMB-Trennung von Proteinen mittels Ionenaustauschchromatographie. *Chemie-Ingenieur-Technik*. 2004;76(6):815–819.
34. Barz T, Löffler V, Arellano-Garcia H, Wozny G. Optimal determination of steric mass action model parameters for β-lactoglobulin using static batch experiments. *J Chromatogr A*. 2010;1217(26):4267–4277.
35. Pedersen L, Mollerup J, Hansen E, Jungbauer A. Whey proteins as a model system for chromatographic separation of proteins. *J Chromatogr B*. 2003;790(1–2):161–173.
36. Gu T. *Mathematical Modeling and Scale-up of Liquid Chromatography*. Berlin, New York: Springer Verlag, 1995.
37. Barz T, Kuntsche S, Wozny G, Arellano-Garcia H. An efficient sparse approach to sensitivity generation for large-scale dynamic optimization. *Comput Chem Eng*. 2011;35(10):2053–2065.
38. Barz T, Kraus R, Zhu L, Wozny G, Arellano Garcia H. Generation of discrete first- and second- order sensitivities for single shooting. *AIChE J*. 2012;58(10):3110–3122.
39. Bauer I, Bock HG, Körkel S, Schlöder JP. Numerical methods for optimum experimental design in DAE systems. *J Comput Appl Math*. 2000;120(1):1–25.
40. Körkel S. Numerische Methoden für Optimale Versuchsplanungsprobleme bei nichtlinearen DAE-Modellen. Ph.D. Thesis, Ruprecht-Karls-Universität Heidelberg, 2002.
41. Fackler PL. Notes on matrix calculus. Tech. rep., North Carolina State University, Raleigh NC, USA, 2005 http://www4.ncsu.edu/~pfackler/MatCalc.pdf. Last accessed on: November 13, 2012.
42. Marlow WH. *Mathematics for Operations Research*. New York: Dover Publications, 1993.

Appendix A: Process Model Equations

Dynamic behavior of the manager and the pump

The desired concentration $c^{ED}(t)$ is computed solving the ED problem online by the personal computer (see Figure 4). The manager and the pump are used to realize this concentration, the output concentration is referred to as $c_{\mathrm{Nt}}(t)$. The dynamic behavior of this system depends mainly on the static mixer which is part of the pump (see Figure 3). It generates a time delay related to the inlet concentration. A

series of linear first-order dynamic systems are used to model $c_{\mathrm{Nt}}(t)$ as a function of $c^{ED}(t)$

$$\tau_1 \cdot \frac{dc_1(t)}{dt} = c^{ED}(t) - c_1(t) \tag{A1}$$

$$\tau_1 \cdot \frac{dc_2(t)}{dt} = c_1(t) - c_2(t)$$

$$\vdots$$

$$\tau_1 \cdot \frac{dc_{\mathrm{Nt}-1}(t)}{dt} = c_{\mathrm{Nt}-2}(t) - c_{\mathrm{Nt}-1}(t)$$

$$\tau_2 \cdot \frac{dc_{\mathrm{Nt}}(t)}{dt} = c_{\mathrm{Nt}-1}(t) - c_{\mathrm{Nt}}(t)$$

In Eq. (A1), $\mathrm{Nt} = 12$ is the number of linear systems and $\tau_1 = 0.0376$ min, $\tau_2 = 0.3$ min are the corresponding time constants.

Model of the chromatography column

The HPLC column is modeled using the Equilibrium Transport Dispersive Model (ETDM), which is obtained from the consideration of three phases: the moving liquid phase (bulk phase), the stagnant liquid pore phase (pore phase) and the solid adsorbent phase (adsorbed phase). The equation for the bulk phase is given as a parabolic differential equation of second order with constant coefficients (see e.g. Ref. 27)

$$\frac{\partial c_{fl}(z,t)}{\partial t} = D_{ax} \cdot \frac{\partial^2 c_{fl}(z,t)}{\partial z^2} - u \cdot \frac{\partial c_{fl}(z,t)}{\partial z} - \frac{1-\varepsilon_{ext}}{\varepsilon_{ext}} \cdot \frac{3}{r_P} \cdot \beta \cdot (c_{fl}(z,t) - c_p(z,t)) \tag{A2}$$

where $c_{fl}(z,t)$ is the concentration in the moving fluid phase and $c_p(z,t)$ the concentration in the pore phase. Moreover, D_{ax} represents the axial dispersion coefficient, u is the constant velocity of the mobile phase, ε_{ext} is the external porosity, r_P is the adsorbent particle radius, β the lumped mass-transfer coefficient, t is the time, and z is the axial-coordinate along the column. The Danckwerts inflow and outflow boundary conditions are given as

$$\frac{\partial c_{fl}}{\partial z}(z=0,t) = \frac{u}{D_{ax}} \cdot \left(c_{fl}(z=0,t) - c^{in}(t)\right); \quad \frac{\partial c_{fl}}{\partial z}(z=l_c,t) = 0 \tag{A3}$$

where $c^{in}(t) = c_{\mathrm{Nt}}(t)$ is the concentration of the eluent given into the column (see Figure 4) and l_c is the bed length of the column. In the ETDM model a second linear implicit equation (Eq. A4) for the pore phase is considered and, thus, the accumulation in the pore phase is decoupled from the accumulation in the bulk phase

$$\varepsilon_{int} \cdot \frac{\partial c_p(z,t)}{\partial t} + (1-\varepsilon_{int}) \cdot \frac{\partial q(z,t)}{\partial t} = \frac{\partial}{\partial t}\left(\varepsilon_{int} \cdot c_p(z,t) + (1-\varepsilon_{int}) \cdot q(z,t)\right) = \frac{3}{r_P} \cdot \beta \cdot (c_{fl}(z,t) - c_p(z,t)) \tag{A4}$$

In Eq. A4, $c_p(z,t)$ is the concentration in the pore phase, ε_{int} is the internal or intraparticular porosity and $q(z,t)$ is the adsorbent concentration (referred to the particle volume). Finally, the multicomponent Langmuir adsorption isotherm describes the equilibrium between adsorbent phase and pore phase (Eq. A5)

$$q(z,t) = \frac{a \cdot c_p(z,t)}{1 + b \cdot c_p(z,t)} \tag{A5}$$

where a, b are the isotherm parameters. The initial conditions for (Eq. A2 and A4) are given as

$$c_{fl}(z, t=0) = c_{fl,0}(z)\,; \quad c_p(z, t=0) = c_{p,0}(z) \tag{A6}$$

Calibration curve of the UV sensor

Finally the calibration of the UV sensor is given as

$$c^m(t) = m_0 + m_1 \cdot \exp[m_2 \cdot s^{UV}(t)] \tag{A7}$$

where $s^{UV}(t)$ is the signal value from the UV sensor, m_0, m_1, m_2 are the calibration parameters and $c^m(t) = c_{fl}(z = l_c, t)$ is the measured liquid concentration at the column outlet (see Figure 4). It has to be noted that (Eq. A7) is solved for $s^{UV}(t)$ and this predicted signal is then fitted to the experimentally measured UV signal by solving the PE problem.

Appendix B: Derivatives of ED Criteria

In Eq. 8, different criteria can be selected in order to define the objective function Φ^{ED} in Eq. 7. These criteria are applied to the covariance matrix Σ_θ whose definition is given for unconstrained estimation problems, see Eq. 6. We use the normalized FIM $\bar{F}$ which is obtained from the normalized sensitivities $\bar{S}$, see Eq. B1

$$\bar{F}(u,\hat{\theta}) = \bar{S}(u,\hat{\theta})^{\mathrm{T}} \cdot \Sigma_y^{-1} \cdot \bar{S}(u,\hat{\theta})\,; \tag{B1}$$

$$\text{with} \quad \bar{S}_{i,j} = S_{i,j} \cdot \hat{\theta}_j \quad \forall i \in \mathrm{Ny} \cdot \mathrm{Nm}\,,\ j \in \mathrm{Np}$$

Based on this, the normalized covariance matrix $\bar{\Sigma}_{\hat{\theta}}$ then indicates relative rather then absolute uncertainties of the parameters.

For an efficient gradient-based solution of the ED problem (see Eq. 7) the first-order derivatives $\partial\Phi^{ED}/\partial u$ are generated together with the solution of the DAE system in Eq. 1. The rules to calculate these derivatives for all ED criteria follow. The calculation of directional derivatives of ED criteria for constrained parameter estimation problems can be found elsewhere.[39,40]

The applied matrix derivation rules, the system for ordering the results using matrix operations as well as the corresponding notations have been taken from Refs. 41 and 42. The application of the chain rule yields a product of four terms

$$\underbrace{\frac{\partial \Phi^{ED}}{\partial u}}_{\in \mathbb{R}^{1\times \mathrm{Nu}}} = \underbrace{\frac{\partial \Phi^{ED}}{\partial \Sigma_\theta}}_{\in \mathbb{R}^{1\times \mathrm{Np}\cdot\mathrm{Np}}} \cdot \underbrace{\frac{\partial \Sigma_\theta}{\partial F}}_{\in \mathbb{R}^{\mathrm{Np}\cdot\mathrm{Np}\times\mathrm{Np}\cdot\mathrm{Np}}} \cdot \underbrace{\frac{\partial F}{\partial S}}_{\in \mathbb{R}^{\mathrm{Np}\cdot\mathrm{Np}\times\mathrm{Ny}\cdot\mathrm{Nm}\cdot\mathrm{Np}}} \cdot \underbrace{\frac{\partial S}{\partial u}}_{\in \mathbb{R}^{\mathrm{Ny}\cdot\mathrm{Nm}\cdot\mathrm{Np}\times\mathrm{Nu}}} \tag{B2}$$

From the dimensions of the different terms in Eq. B2 it can be seen that the matrices Σ_θ, F, S are transformed to vectors. This is done by the vec() operator which indicates the vectorization of a matrix by stacking its columns.[41] The result for the transformed sensitivity matrix reads

$$\mathrm{vec}(S) = \left(\left(\frac{\partial y_1}{\partial \theta_1}, \cdots, \frac{\partial y_{\mathrm{Ny}}}{\partial \theta_1} \right)\Big|_{t_1}, \left(\frac{\partial y_1}{\partial \theta_1}, \cdots, \frac{\partial y_{\mathrm{Ny}}}{\partial \theta_1} \right)\Big|_{t_2}, \cdots, \left(\frac{\partial y_1}{\partial \theta_1}, \cdots, \frac{\partial y_{\mathrm{Ny}}}{\partial \theta_1} \right)\Big|_{t_{\mathrm{Nm}}}, \right. \quad \text{(B3)}$$

$$\left(\frac{\partial y_1}{\partial \theta_2}, \cdots, \frac{\partial y_{\mathrm{Ny}}}{\partial \theta_2} \right)\Big|_{t_1}, \left(\frac{\partial y_1}{\partial \theta_2}, \cdots, \frac{\partial y_{\mathrm{Ny}}}{\partial \theta_2} \right)\Big|_{t_2}, \cdots, \left(\frac{\partial y_1}{\partial \theta_2}, \cdots, \frac{\partial y_{\mathrm{Ny}}}{\partial \theta_2} \right)\Big|_{t_{\mathrm{Nm}}}, \quad \cdots \quad ,$$

$$\left(\frac{\partial y_1}{\partial \theta_{\mathrm{Np}}}, \cdots, \frac{\partial y_{\mathrm{Ny}}}{\partial \theta_{\mathrm{Np}}} \right)\Big|_{t_1}, \left(\frac{\partial y_1}{\partial \theta_{\mathrm{Np}}}, \cdots, \frac{\partial y_{\mathrm{Ny}}}{\partial \theta_{\mathrm{Np}}} \right)\Big|_{t_2}, \cdots, \left. \left(\frac{\partial y_1}{\partial \theta_{\mathrm{Np}}}, \cdots, \frac{\partial y_{\mathrm{Ny}}}{\partial \theta_{\mathrm{Np}}} \right)\Big|_{t_{\mathrm{Nm}}} \right)^{\mathrm{T}}$$

The calculation of the first derivative term in Eq. B2 is given in Eq. B4. According to the different design criteria in (Eq. 8) they read

$$\frac{\partial \Phi^A}{\partial \Sigma_\theta} = \frac{\partial}{\partial \Sigma_\theta} \left(\frac{1}{\mathrm{Np}} \cdot \mathrm{tr}(\Sigma_\theta) \right) = \frac{1}{\mathrm{Np}} \cdot \mathrm{vec}(I_{\mathrm{Np}})^{\mathrm{T}} \quad \text{(B4a)}$$

$$\frac{\partial \Phi^D}{\partial \Sigma_\theta} = \frac{\partial}{\partial \Sigma_\theta} \left((\det(\Sigma_\theta))^{\frac{1}{\mathrm{Np}}} \right) = \frac{1}{\mathrm{Np}} \cdot (\det(\Sigma_\theta))^{\frac{1}{\mathrm{Np}}} \cdot \mathrm{vec}(\Sigma_\theta^{-\mathrm{T}})^{\mathrm{T}} \quad \text{(B4b)}$$

$$\frac{\partial \Phi^E}{\partial \Sigma_\theta} = \frac{\partial}{\partial \Sigma_\theta} (\max(\lambda(\Sigma_\theta))) = (v^{\mathrm{T}} \cdot v)^{-1} \cdot v^{\mathrm{T}} \cdot (I_{\mathrm{Np}} \otimes v^{\mathrm{T}}) \quad \text{(B4c)}$$

In Eq. B4b, $\Sigma_\theta^{-\mathrm{T}} = (\Sigma_\theta^{-1})^{\mathrm{T}}$ and in Eq. B4c, v denotes the eigenvector which corresponds to the maximum eigenvalue of the parameter covariance matrix $\max(\lambda(\Sigma_\theta))$. The derivative of a matrix with respect to its inverse is

$$\frac{\partial \Sigma_\theta}{\partial F} = \frac{\partial}{\partial F} (F + C)^{-1} = -\left(F^{-\mathrm{T}} \otimes F^{-1} \right) \quad \text{(B5)}$$

where $C \in \mathbb{R}^{\mathrm{Np} \times \mathrm{Np}}$ is considered to be an additional constant matrix, used here to represent the information content of previous experiments, see also the definition in Eq. 16. The derivative of the FIM F reads (see Refs. 41 and 42)

$$\frac{\partial F}{\partial S} = \frac{\partial}{\partial S} \left(S^{\mathrm{T}} \cdot \Sigma_y^{-1} \cdot S \right) = \left(S^{\mathrm{T}} \cdot \Sigma_y^{-1} \otimes I_{\mathrm{Np}} \right) \cdot K + \left(I_{\mathrm{Np}} \otimes S^{\mathrm{T}} \cdot \Sigma_y^{-\mathrm{T}} \right) \quad \text{(B6)}$$

where K is the $(\mathrm{Ny} \cdot \mathrm{Nm} \cdot \mathrm{Np} \times \mathrm{Ny} \cdot \mathrm{Nm} \cdot \mathrm{Np})$ permutation matrix that transforms $[\mathrm{vec}(S)]$ into $[\mathrm{vec}(S^{\mathrm{T}})]$, see Ref. 41.

For a scaling of the sensitivity matrix S with the parameter values (see Eq. B1), the following element-wise multiplication of two matrices is considered $\bar{S} = S \circ M_{sc}$, being $M_{sc} = \theta^{\mathrm{T}} \otimes I_{\mathrm{Ny} \cdot \mathrm{Nm} \times 1}$ the scaling matrix. The derivatives of F with respect to $\bar{S}$ are then calculated in the same way as in Eq. B6 and additionally an element-wise multiplication of the result is performed.

$$\frac{\partial F}{\partial S} = \frac{\partial F}{\partial \bar{S}} \circ \frac{\partial \bar{S}}{\partial S} = \frac{\partial}{\partial \bar{S}} \left(\bar{S}^{\mathrm{T}} \cdot \Sigma_y^{-1} \cdot \bar{S} \right) \circ \left(\mathrm{vec}(M_{sc}) \otimes I_{\mathrm{Np} \cdot \mathrm{Np} \times 1} \right) \quad \text{(B7)}$$

Finally, the last derivative term in Eq. B2 is defined by mixed second-order directional derivatives

$$\frac{\partial S}{\partial u} = \frac{\partial}{\partial u} \left(\frac{\partial y}{\partial \theta} \right) = \frac{\partial^2 y}{\partial u \partial \theta} = H_{u,\theta} \quad \text{(B8)}$$

They correspond to second-order sensitivities of the DAE system and have to be generated together with the solution of Eq. 1. The dimension of $H_{u,\theta}$ is shown in Eq. B9. The ordering results directly from the derivation of the vectorized matrix S whose definition is given in Eq. B3

$$H_{u,\theta} = \frac{\partial (\mathrm{vec}(S))}{\partial u} \in \mathbb{R}^{\mathrm{Ny} \cdot \mathrm{Nm} \cdot \mathrm{Np} \times \mathrm{Nu}} \quad \text{(B9)}$$

Manuscript received July 12, 2012, and revision received Oct. 11, 2012.