Optimal Experiment Design for Nonlinear Dynamic (Bio)chemical Systems Using Sequential Semidefinite Programming

Dries Telen and Filip Logist
Dept. of Chemical Engineering, BioTeC and OPTEC, KU Leuven, W. de Croylaan 46, 3001 Leuven, Belgium

Rien Quirynen
Dept. of Electrical Engineering, SCD and OPTEC, KU Leuven, Kasteelpark Arenberg 10, 3001 Leuven, Belgium

Boris Houska
Dept. of Automation, Shanghai Jiao Tong University, 800 Dong Chuan Road, Shanghai 200240, China

Moritz Diehl
Dept. of Electrical Engineering, SCD and OPTEC, KU Leuven, Kasteelpark Arenberg 10, 3001 Leuven, Belgium

Jan Van Impe
Dept. of Chemical Engineering, BioTeC and OPTEC, KU Leuven, W. de Croylaan 46, 3001 Leuven, Belgium

DOI 10.1002/aic.14389
Published online March 3, 2014 in Wiley Online Library (wileyonlinelibrary.com)

Optimal experiment design (OED) for parameter estimation in nonlinear dynamic (bio)chemical processes is studied in this work. To reduce the uncertainty in an experiment, a suitable measure of the Fisher information matrix or variance–covariance matrix has to be optimized. In this work, novel optimization algorithms based on sequential semidefinite programming (SDP) are proposed. The sequential SDP approach has specific advantages over sequential quadratic programming in the context of OED. First of all, it guarantees on a matrix level a decrease of the uncertainty in the parameter estimation procedure by introducing a linear matrix inequality. Second, it allows an easy formulation of E-optimal designs in a direct optimal control optimization scheme. Finally, a third advantage of SDP is that problems involving the inverse of a matrix can be easily reformulated. The proposed techniques are illustrated in the design of experiments for a fed-batch bioreactor and a microbial kinetics case study. © 2014 American Institute of Chemical Engineers *AIChE J*, 60: 1728–1739, 2014
Keywords: optimal experiment design, semidefinite programming, parameter estimation, dynamic optimization, bioprocesses

Introduction

Dynamic process models play an important role in the analysis, control, and optimization of (bio)chemical processes. Taking and analyzing measurements is often a costly and time consuming practice. In the last decades, optimal experiment design (OED) has gained increasing attention to limit the experimental burden.[1–5] The goal is to design an excitation such that as much information as possible is obtained, see[4] for a review of the state of the art.

In OED for dynamic systems, most often a time-varying input is determined that maximizes the information content or minimizes the uncertainty in the experiment. This usually leads to the optimization of a scalar function of the Fisher information matrix[2] or variance–covariance matrix,[3,6] respectively. Several design criteria have been proposed in the literature.[2] Given the dynamic nature, this approach results in a challenging class of dynamic optimization problems.[1] These dynamic optimization/optimal control problems can be solved by direct methods in which the original infinite dimensional problem is reformulated as a finite dimensional nonlinear program via discretization of the controls and/or states. In[3,7] the specific numerical aspects of OED for nonlinear dynamic systems are addressed.

When a specific criterion is selected, it is not sure that the designed experiment will increase the information content as measured by the other criteria.[8] Furthermore, using, for example, the E-criterion in a direct optimal control formulation can be troublesome. The E-criterion involves the minimization of the largest eigenvalue of the variance–covariance matrix or the maximization of the smallest eigenvalue of the Fisher information matrix. In direct optimal control formulations, usually a Newton type method is used in each step, requiring the computation of first- or even second-order derivatives. However, the maximal eigenvalue function is in general nonsmooth.

To guarantee a decrease in the uncertainty (i.e., an improvement in all criteria at the same time), a sequential semidefinite program (SDP) approach is proposed in this article. SDP involves the optimization of a linear objective function subject to linear matrix inequalities. To ensure a decrease in uncertainty, a matrix inequality is added. The

Correspondence concerning this article should be addressed to J. V. Impe at jan.vanimpe@cit.kuleuven.be.

variance–covariance matrix is convexified by linearization in each iteration, similar to the approach presented in.[9,10] An additional advantage is that the minimization of the maximum eigenvalue can be cast in a SDP which avoids the nonsmoothness problem. Furthermore, the A-criterion involves the minimization of the trace of the inverse of the Fisher information matrix. Using reformulations from the field of convex optimization, this problem can be stated as a linear matrix inequality without the need of computing derivatives throughout a matrix inverse. The proposed algorithms are illustrated by a fed-batch bioreactor and a microbial kinetics case study. The article is structured as follows. In the second section, the mathematical formulation of OED for parameter estimation in dynamic systems is discussed. The third section introduces the concept of SDP and proposes the extension to OED for nonlinear dynamic systems. In the fourth section, the case studies and their numerical implementation are presented. The fifth section discusses the obtained results. The conclusions are formulated in the sixth section.

OED for Parameter Estimation in Dynamic Systems

In this section, the used mathematical formulations are introduced. The first subsection discusses the formulation of nonlinear dynamic systems. The second and third subsections describe the way how information content and uncertainty is quantified for OED by either using the Fisher information matrix or a variance–covariance matrix approach. In the fourth subsection, the novel insight between the two approaches is discussed, while the different design criteria for OED are elaborated in the fifth subsection. The sixth subsection concludes with the optimization problem formulation.

Nonlinear dynamic systems

The dynamic evolution of many (bio)chemical processes in a time interval $[0, t_f]$ can be described by differential equations

$$\dot{y}(t)=g(y(t),p,u(t)) \text{ with } y(0)=y_0 \tag{1}$$

Here, $y(t) \in \mathbb{R}^{n_y}$ is the state vector, $p \in \mathbb{R}^{n_p}$ the unknown but time-invariant parameter vector and $u(t) \in \mathbb{R}^{n_u}$ the control vector which is the degree of freedom in the OED procedure. Measurements are assumed to be $\eta(t)=z(t)+\epsilon \in \mathbb{R}^{n_z}$, where the relation between the measured quantities and the states may in general be nonlinear, $z(t)=h(y(t))$, too. Here, ε denotes the measurement error which is assumed to have a Gaussian distribution $\epsilon \in \mathcal{N}(0, V(t))$ with zero mean and a variance–covariance matrix $V(t) \in \mathbb{R}^{n_z \times n_z}$.

In practice, the initial value for the state vector and the true value for the parameters follow Gaussian distributions $y(0) \in \mathcal{N}(\eta_0, Q_y), p \in \mathcal{N}(\pi_0, Q_p)$, with known positive semidefinite variance–covariance matrices $Q_y \in \mathbb{R}^{n_y \times n_y}$, $Q_p \in \mathbb{R}^{n_p \times n_p}$, and given expectations $\eta_0 \in \mathbb{R}^{n_y}, \pi_0 \in \mathbb{R}^{n_p}$.

The Fisher information matrix approach

The information content of an experiment has to be quantified. A practical method to do so for dynamic systems is based on the Fisher information matrix.[4] The classic time-varying version of the Fisher information matrix, $F(t) \in \mathbb{R}^{n_p \times n_p}$, is defined as the solution of the following differential equation

$$\dot{F}(t)=S(t)^\top D(t)^\top V^{-1}(t) D(t) S(t) \tag{2}$$

$$F(0)=Q_p^{-1} \tag{3}$$

The Fisher information matrix is positive semidefinite and symmetric. Furthermore, $F(t)$ satisfies $F(t) \preceq F(t')$ for all $t, t' \in [0, t_f]$ with $t' \geq t$. As the true values p are unknown, the Fisher information matrix depends in general on the current best estimate. Besides the inverse of the measurement error variance–covariance matrix $V^{-1}(t)$, the sensitivities of the model output with respect to the parameters are present in the Fisher information matrix. These sensitivities, $S(t) \in \mathbb{R}^{n_x \times n_p}$ are computed as the solution of the following variational differential equation

$$\dot{S}(t)=B(t)S(t)+P(t) \tag{4}$$

$$S(0)=\frac{\partial y_0}{\partial p} \tag{5}$$

in which

$$B(t)=\frac{\partial g(y(t),p,u(t))}{\partial y}, \quad P(t)=\frac{\partial g(y(t),p,u(t))}{\partial p}, \quad D(t)=\frac{\partial h(y(t))}{\partial y} \tag{6}$$

Under the assumption of unbiased estimators, the inverse of $F(t_f)$ is the lower bound of the parameter estimation variance–covariance matrix, that is, the Cramér–Rao bound.[11] In the presented definition, it is assumed that measurements are taken continuously. However, the decision whether to measure can be easily incorporated in the Fisher information matrix approach in the following way

$$\dot{F}(t)=w(t)S(t)^\top D(t)^\top V^{-1}(t) D(t) S(t) \tag{7}$$

$$F(0)=Q_p^{-1} \tag{8}$$

in which the additional control function $w(t) \in \{0, 1\}$ is introduced, similar to the approach in.[12] Note that in the current formulation measuring several times at the same time instance is excluded.

A variance–covariance matrix approach

For notational convenience, a simpler but equivalent dynamic process model can be used. The parameters are stacked to the states in the following way: $x(t)=[y(t)^\top, p^\top]^\top$ with $x(t) \in \mathbb{R}^{n_x}$ and $n_x=n_y+n_p$. The time-invariant parameters are added to the dynamic system by adding the trivial differential equation

$$\frac{dp}{dt}=0 \tag{9}$$

$$p(0)=p \tag{10}$$

where p is the parameter vector as previously defined. It is the current best estimate for π_0. The dynamic system formulation subsequently becomes

$$\dot{x}(t)=f(x(t),u(t))=[g(y(t),p,u(t))^\top, 0^\top]^\top \tag{11}$$

$$\text{with } x(0)=[y_0^\top, p^\top]^\top \tag{12}$$

In the classic formulation of OED, the accuracy of estimating the unknown parameter vector p is analyzed. The formulation in this section is based on the dynamic system formulation in (11) to (12) which allows for a more general procedure which can optionally also take joint information about states $y(t)$ and parameters p into account.

In[6] an efficient computational method is proposed to obtain the variance–covariance matrix of the state vector

(and, thus, also of the parameter vector as the stacked approach is used here). The variance–covariance matrix is computed as the solution of a Riccati differential equation. The proposed computational strategy is the following

$$\begin{aligned}\dot{Q}(t)&=A(t)Q(t)+Q(t)A(t)^{\top}-Q(t)C(t)^{\top}V^{-1}(t)C(t)Q(t)\\ Q(t)&=Q_0\,,\end{aligned} \tag{13}$$

in which the following short hands are used

$$A(t)=\frac{\partial f(x(t),u(t))}{\partial x},\quad C(t)=\frac{\partial h(x(t))}{\partial x},\quad Q_0=\begin{pmatrix}Q_{\mathrm{y}} & 0\\ 0 & Q_{\mathrm{p}}\end{pmatrix} \tag{14}$$

Equation 13 yields the desired variance–covariance matrix.

Novel insight in the connection between the classic Fisher information matrix and the proposed variance–covariance matrix

When the focus is only on the parameter accuracy in the variance–covariance approach, the following scaling approach can be used $\Sigma Q(t)\Sigma^{\top}$ where Σ is

$$\Sigma\ =(0\quad I) \tag{15}$$

with $I\in\mathbb{R}^{n_p\times n_p}$ the identity matrix. If y_0 is fixed, then this scaling allows for selecting and optimizing those elements of $Q(t)$ which are associated with the parameter variance–covariance matrix, resulting in a similar formulation as the current practice of OED for parameter estimation in nonlinear dynamic systems.[4] Note that there is a subtle difference between the two suggested approaches. In most dynamic OED approaches, the Fisher information matrix as defined by Eq. 2 is used. However, if the Fisher information matrix is extended to incorporate uncertainty regarding the states as in,[6] the Fisher information matrix $F_{\mathrm{ext}}(t)\in\mathbb{R}^{n_x\times n_x}$ has the following structure

$$F_{\mathrm{ext}}(t)\ =\begin{pmatrix}F_{\mathrm{y}} & F_{\mathrm{y,p}}\\ F_{\mathrm{p,y}} & F_{\mathrm{p}}\end{pmatrix} \tag{16}$$

in which F_{p} is the same as $F(t)$ used in this work, $F_{\mathrm{y,p}}\in\mathbb{R}^{n_y\times n_p}$ is the subblock relating information between the states y and the parameters p, $F_{\mathrm{p,y}}=F_{\mathrm{y,p}}^{\top}$, and $F_{\mathrm{y}}\in\mathbb{R}^{n_y\times n_y}$ is the subblock which computes the Fisher information for the states y. Necessary for this extended Fisher information matrix, are the following extended sensitivity equations, $S_{\mathrm{ext}}(t)\in\mathbb{R}^{n_x\times n_x}$, computed as the solution of

$$\dot{S}_{\mathrm{ext}}(t)=A(t)S_{\mathrm{ext}}(t) \tag{17}$$

$$S_{\mathrm{ext}}(0)=I \tag{18}$$

Between the extended Fisher information matrix, $F_{\mathrm{ext}}(t)$ and the variance–covariance matrix, $Q(t)$ the following relationship was proven in[6]

$$Q(t)=S_{\mathrm{ext}}(t)F_{\mathrm{ext}}(t)^{-1}S_{\mathrm{ext}}(t)^{\top} \tag{19}$$

However, the solution of the extended sensitivities has a particular form

$$S_{\mathrm{ext}}(t)\ =\begin{pmatrix}S_{\mathrm{y}} & S_{\mathrm{y,p}}\\ 0 & I\end{pmatrix} \tag{20}$$

in which $S_{\mathrm{y,p}}=S(t)$ and $S_{\mathrm{y}}\in\mathbb{R}^{n_y\times n_y}$ are the sensitivities of the states with respect to themselves, if this connection is inserted in Eq. 19 and subsequently computed, the following expression is obtained

$$Q(t)=\begin{pmatrix}S_{\mathrm{y}}f_{11}S_{\mathrm{y}}+S_{\mathrm{y,p}}f_{21}S_{\mathrm{y}}+S_{\mathrm{y}}f_{12}S_{\mathrm{y,p}}+S_{\mathrm{y,p}}f_{22}S_{\mathrm{y,p}} & S_{\mathrm{y}}f_{12}+S_{\mathrm{y,p}}f_{22}\\ f_{21}S_{\mathrm{y}}+f_{22}S_{\mathrm{y,p}} & f_{22}\end{pmatrix} \tag{21}$$

in which f_{11},f_{12},f_{21}, and f_{22} are the block matrices of the inverse of the extended Fisher information matrix

$$F_{\mathrm{ext}}^{-1}=\begin{pmatrix}f_{11} & f_{12}\\ f_{21} & f_{22}\end{pmatrix} \tag{22}$$

So, by exploiting formulas for the inverse of block matrices, it can be shown that there exists a simple connection between the subblock matrix governing the parameters of the extended Fisher information matrix and the subblock matrix governing the parameters in the variance–covariance approach proposed in[6]

$$\Sigma Q(t)\Sigma^{\top}=\left(F_{\mathrm{p}}(t)-F_{\mathrm{p,y}}(t)F_{\mathrm{y}}^{-1}(t)F_{\mathrm{y,p}}(t)\right)^{-1} \tag{23}$$

From this equation, it can be inferred that the used variance–covariance matrix approach is very similar to the classic Fisher information matrix. The variance–covariance matrix approach, however, has the additional advantage of taking uncertainty regarding the states into account which is apparent in the expression as $-F_{\mathrm{p,y}}(t)F_{\mathrm{y}}^{-1}(t)F_{\mathrm{y,p}}(t)$. For an in depth discussion on the proposed variance–covariance matrix, the interested reader is referred to.[6]

Design criteria

The goal is to design an experiment such that the information content is maximal or the variance–covariance is minimal. Because optimizing a matrix is not possible, several design criteria have been suggested[2,4,13,14] in the literature. These criteria are typically scalar functions, $\Phi(\cdot)$ of the variance–covariance matrix or Fisher information matrix. Some well known and widely used criteria are the following

- A-optimal designs minimize the mean of the asymptotic variances of the parameter estimates. This boils down to minimizing the trace of the variance–covariance matrix, that is, $\Phi(Q(t_{\mathrm{f}}))=\mathrm{Tr}\,(Q(t_{\mathrm{f}}))$ is chosen if the variance–covariance matrix is used or $\Phi(F(t_{\mathrm{f}}))=\mathrm{Tr}\,(F(t_{\mathrm{f}})^{-1})$ is used when the computation is based on a Fisher information matrix approach.
- D-optimal designs minimize the geometric mean of the eigenvalues of $Q(t_{\mathrm{f}})$. This is equivalent to minimizing the determinant of the variance–covariance matrix, that is, choose $\Phi(Q(t_{\mathrm{f}}))=\mathrm{Det}\,(Q(t_{\mathrm{f}}))$. Formulated in the Fisher information matrix approach, this yields $\Phi(F(t_{\mathrm{f}}))=-\mathrm{Det}\,(F(t_{\mathrm{f}}))$. Note that the determinant is scaling invariant.[4]
- E-optimal designs aim at minimizing the maximum eigenvalue of $Q(t_{\mathrm{f}})$, that is, choose $\Phi(Q(t_{\mathrm{f}}))=\lambda_{\max}(Q(t_{\mathrm{f}}))$ or maximizing the minimal eigenvalue of $F(t_{\mathrm{f}})$, that is, $\Phi(F(t_{\mathrm{f}}))=-\lambda_{\min}(F(t_{\mathrm{f}}))$. Geometrically, this means minimizing the largest uncertainty axis of the joint confidence region.
- M-optimal designs minimize the maximal diagonal element of the variance–covariance matrix, that is, choose $\Phi(Q(t_{\mathrm{f}}))=\max_i Q_{ii}(t_{\mathrm{f}})$ or in the Fisher information

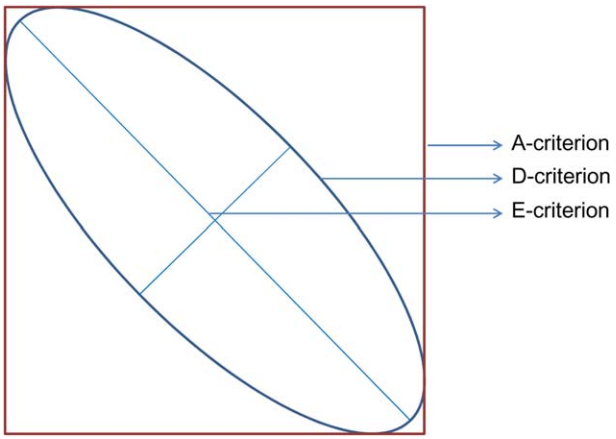

Figure 1. Illustration of the geometric meaning of the different OED-criteria for a two parameter case.

Minimizing the enclosing frame is the A-criterion, minimizing the volume/area is the D-criterion, and minimizing the largest uncertainty axis is the E-criterion. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

matrix approach, that is, choose $\Phi(F(t_f)) = \max_i F_{ii}^{-1}(t_f)$. The M-criterion tries to minimize directly the uncertainty of the most uncertain parameter by selecting the corresponding diagonal element. This criterion is very analogous to the E-criterion as both work directly on the most uncertain parameter.

Figure 1 illustrates the geometric interpretation of different criteria for a two parameter case.

Problem formulation

To design an optimal experiment for (bio)chemical processes, a dynamic optimization problem has to be solved. The optimization problem formulated in the Fisher information matrix approach can be expressed as

$$\min_{u(\cdot),y(\cdot),S(\cdot),F(\cdot)} \Phi(F(t_f)) \tag{24}$$

subject to

$$\dot{y}(t) = g(y(t), p, u(t)) \tag{25}$$
$$y(0) = y_0 \tag{26}$$
$$\dot{S}(t) = B(t)S(t) + P(t) \tag{27}$$
$$S(0) = \frac{\partial y_0}{\partial p} \tag{28}$$
$$\dot{F}(t) = S(t)^\top D(t)^\top V(t)^{-1} D(t) S(t) \tag{29}$$
$$F(0) = Q_p^{-1} \tag{30}$$
$$0 \geq c_p(y(t), p, u(t)) \tag{31}$$
$$0 \geq c_t(y(t_f)) \tag{32}$$

The variance–covariance approach yields a similar formulation

$$\min_{u(\cdot),x(\cdot),Q(\cdot)} \Phi(Q(t_f)) \tag{33}$$

subject to

$$\dot{x}(t) = f(x(t), u(t)) \tag{34}$$
$$x(0) = x_0 \tag{35}$$
$$\dot{Q}(t) = A(t)Q(t) + Q(t)A(t)^\top - Q(t)C(t)^\top V(t)^{-1} C(t) Q(t) \tag{36}$$
$$Q(0) = Q_0 \tag{37}$$
$$0 \geq c_p(x(t), u(t)) \tag{38}$$
$$0 \geq c_t(x(t_f)) \tag{39}$$

In the above formulation, c_p and c_t indicate the path constraints and the terminal constraints, respectively. This optimization problem is infinite dimensional. Direct optimal control approaches discretize such problems to end up with finite dimensional nonlinear programming problems. Two different approaches exist for direct optimal control. Direct sequential methods as single shooting (e.g.,[15–17]) discretize the control functions only, whereas the direct simultaneous approaches discretize both state and control functions. Within these simultaneous approaches one can distinguish: multiple shooting (e.g.,[18,19]) and orthogonal collocation (e.g.,[20,21]). In this work, the single shooting approach is used as direct optimal control method.

As experimental design inherently has a matrix valued objective function that artificially needs to be scalarized to make it amenable to optimization, OED formulations usually do not ensure that an optimized new experiment is better than an old one under all criteria.[8] However, it is clear that a variance–covariance matrix Q_{new} is better under all meaningful criteria than an old one Q_{old} if and only if $Q_{old} \geqslant Q_{new}$ in the sense of matrix inequalities.

Here, a design criterium Φ is considered meaningful, if we have

$$\Phi(X) \leq \Phi(Y) \text{ for all matrices } X, Y \in \mathbb{S}_+^{n_x} \text{ with } X \preceq Y,$$

with $\mathbb{S}_+^{n_x}$ the set of all positive semidefinite matrices. For this reason, a matrix constraint is included which ensures that an optimized experiment is surely better than an old one, whose matrix is used in the constraint. For the Fisher information matrix approach this leads to the following formulation

$$\min_{u(\cdot),y(\cdot),S(\cdot),F(\cdot)} \Phi(F(t_f)) \tag{40}$$

subject to

$$\begin{gathered} (25) - (32) \\ F(t_f) \geqslant \Sigma Q_{\text{bound}}^{-1} \Sigma^\top \end{gathered} \tag{41}$$

while for the variance–covariance approach the formulation becomes

$$\min_{u(\cdot),x(\cdot),P(\cdot)} \Phi(Q(t_f)) \tag{42}$$

subject to

$$\begin{gathered} (34) - (39) \\ Q_{\text{bound}} \geqslant Q(t_f) \end{gathered} \tag{43}$$

The matrix Q_{bound} denotes a prespecified bound on the variance–covariance matrix, for example, *a priori* knowledge. The addition of the nonlinear matrix inequality results in a nonlinear SDP.[22] How to treat the matrix constraints will be discussed in the third section.

Sequential SDP

This section starts with a brief discussion on SDP in the first subsection. The proposed algorithm for OED using SDP and the different optimization formulations are discussed in the second subsection. The third subsection on numerical and software aspects concludes this section.

Semidefinite programming

In a SDP problem, a linear objective function is minimized subject to a linear matrix inequality. The SDP problem can be formulated as

$$\min_{\xi} c^{\top}\xi \tag{44}$$

subject to

$$K(\xi)\succcurlyeq 0 \tag{45}$$

with

$$K(\xi)=K_0+\sum_{i=1}^{m}\xi_i K_i \tag{46}$$

Here, ξ_i denotes the ith component of the vector $\xi\in\mathbb{R}^m$. The matrices $K_0,\ldots,K_n\in\mathbb{R}^{n\times n}$ are symmetric and the inequality $K(\xi)\succcurlyeq 0$ indicates that $K(\xi)$ is a positive semidefinite matrix, that is, $\forall z\in\mathbb{R}^n : z^{\top}K(\xi)z\geq 0$. This optimization problem is a SDP problem[23] which is convex. The main advantage of convex optimization problems is that any local minimum is also the global minimum. In addition, SDPs and convex optimization problems in general can be solved efficiently, both in the theory and practice.[23–25]

Remark. *An interesting case of SDP is the minimization of the maximum eigenvalue. This corresponds to the E-optimal design criterion in OED. The SDP formulation becomes in this case the following*

$$\min_{\xi,\tau}\ \tau \tag{47}$$

subject to

$$\tau I-K(\xi)\succcurlyeq 0 \tag{48}$$

with

$$K(\xi)=K_0+\sum_{i=1}^{m}\xi_i K_i \tag{49}$$

with $\tau\in\mathbb{R}$. This formulation is used in the following section to construct an iterative algorithm that allows the user to perform E-OED for dynamic systems by sequentially solving appropriate SDP problems.

Proposed algorithm

In this section, the novel OED techniques based on sequential SDP are introduced. First, the approximation of the Fisher information matrix or variance–covariance matrix is discussed. Second, a generic sequential convex programming (SCP) approach is introduced. Note that SDPs are a particular type of convex optimization problems. Subsequently, different SDP problem formulations corresponding to different design criteria are discussed.

The main idea is to approximate the Fisher information matrix (2) or the variance–covariance matrix (13) by a linearization in the following way

$$F_{\mathrm{k}}(t_{\mathrm{f}},u,\bar{u}_{\mathrm{k}})=F(t_{\mathrm{f}},\bar{u}_{\mathrm{k}})+\frac{\partial F(t_{\mathrm{f}},\bar{u}_{\mathrm{k}})}{\partial u}(u-\bar{u}_{\mathrm{k}}) \tag{50}$$

$$Q_{\mathrm{k}}(t_{\mathrm{f}},u,\bar{u}_{\mathrm{k}})=Q(t_{\mathrm{f}},\bar{u}_{\mathrm{k}})+\frac{\partial Q(t_{\mathrm{f}},\bar{u}_{\mathrm{k}})}{\partial u}(u-\bar{u}_{\mathrm{k}}) \tag{51}$$

in which $\frac{\partial F(t_{\mathrm{f}},\bar{u}_{\mathrm{k}})}{\partial u},\frac{\partial Q(t_{\mathrm{f}},\bar{u}_{\mathrm{k}})}{\partial u}$ are the matrix derivative of $F(t_{\mathrm{f}},\bar{u}_{\mathrm{k}}),Q(t_{\mathrm{f}},\bar{u}_{\mathrm{k}})$, respectively. Both derivatives are evaluated in the current point $\bar{u}_{\mathrm{k}}$.

The above formulation is subsequently used in a SCP algorithm. Similar ideas for (robust) optimal control have been proposed in.[9,10,22] To obtain a single shooting approach, the control function $u(t)$ is discretized in N equidistant intervals with a piecewise constant value u_i, with $i=1,\ldots,N$, in each interval. The vector u_{d} containing all u_i as elements is the decision variable of the convex optimization problem.

The SCP algorithm is an iterative method formulated as follows

Algorithm: *SCP for OED.*

Step 1. Find a feasible initial vector $\bar{u}_k$, an initial variance–covariance matrix or Fisher information matrix, Q_p or Q_p^{-1}, an upper bound on the variance–covariance matrix or lower bound on the Fisher information matrix, Q_{bound} or F_{bound}, and set $k=1$.

Step 2. Integrate the dynamic system (34–39) and compute $Q(t_{\mathrm{f}},\bar{u}_k)$ and $\frac{\partial Q(t_{\mathrm{f}},\bar{u}_k)}{\partial u}$ or integrate (25–32) $F(t_{\mathrm{f}},\bar{u}_k)$ and $\frac{\partial F(t_{\mathrm{f}},\bar{u}_k)}{\partial u}$.

Step 3. Solve a convex problem with respect to the discretized control vector u_{d}. The solution of the optimization procedure is the vector u^*. Several optimization formulations can be found in the following subsections.

Step 4. If $||u^*-\bar{u}_k||_2\leq\epsilon$ then terminate, otherwise, $\Delta u=u^*-\bar{u}_k$ update $\bar{u}_{k+1}=\bar{u}_k+\alpha\Delta u, k=k+1$ and go back to Step 2.

The described full-step procedure is obtained for $\alpha=1$. Moreover, to increase the reliability of the algorithm, a backtracking line search algorithm or other techniques to ensure global convergence can be used.[26]

Formulation 1: Guaranteed Decrease on the Variance–Covariance Matrix Level (Guaranteed Increase on the Fisher Information Matrix Level). In OED, one is interested in a design which minimizes the variance–covariance matrix. In the literature some scalar functions, for example, A-criterion is used. An advantage of the presented semidefinite framework is that a decrease of the variance–covariance matrix can be guaranteed by a matrix inequality. Given a control vector $\bar{u}_{\mathrm{k}}$, the following optimization problem is formulated

$$\min_{u_{\mathrm{d}}}\ \mathrm{Tr}\left(Q_{\mathrm{k}}(t_{\mathrm{f}},u_{\mathrm{d}},\bar{u}_{\mathrm{k}})\right) \tag{52}$$

subject to

$$0\succcurlyeq Q_{\mathrm{k}}(t_{\mathrm{f}},u_{\mathrm{d}},\bar{u}_{\mathrm{k}})-Q_{\mathrm{bound}} \tag{53}$$

in which Q_{bound} is a predefined upper bound on the variance–covariance matrix, for example, the result of *a priori* knowledge. If the M-criterion would be considered, the problem can be formulated in the following linear way

$$\min_{u_{\mathrm{d}},\tau}\ \tau \tag{54}$$

subject to

$$0\succcurlyeq Q_{\mathrm{k}}(t_{\mathrm{f}},u_{\mathrm{d}},\bar{u}_{\mathrm{k}})-Q_{\mathrm{bound}} \tag{55}$$

$$\tau\geq Q_{\mathrm{k,ii}}(t_{\mathrm{f}},u_{\mathrm{d}},\bar{u}_{\mathrm{k}})\quad\forall i=1,\ldots,n_x \tag{56}$$

The above problem formulations are semidefinite problems which will be solved in each Step 3 of the proposed algorithm in the second subsection of the third section. Note that the two considered objective functions are linear in the decision variables. If this function is not linear but convex, for example, the logarithm of the determinant of the Fisher information matrix, that is, the D-criterion, the problem requires a dedicated convex

optimization solver able to cope with the objective function and the linear matrix inequalities.[27] Another possibility is that the convex objective function can be linearly approximated in the same way as the variance–covariance matrix.

Formulation 2: Minimizing the Maximal Eigenvalue (Maximizing the Minimal Eigenvalue). An advantage of SDP is that it allows for an easy formulation to incorporate the minimization of the maximal eigenvalue (see the remark in the third section). In OED, the minimization of the maximal eigenvalue is known as the E-criterion. The problem with this formulation for a system described by differential equations is that an accurate computation of the derivatives of the maximal eigenvalue function, $\lambda_{\max}(Q(t_f))$ is needed. As this function is typically not differentiable everywhere, the computation of accurate derivatives is often troublesome. However, by casting this problem in a sequential SDP formulation, the problem of computing the derivatives can be overcome. The sequential semidefinite problem formulation for minimizing the maximal eigenvalue of the variance–covariance matrix is

$$\min_{\tau,u_d} \tau \tag{57}$$

subject to

$$0 \geqslant Q_k(t_f, u_d, \bar{u}_k) - Q_{\text{bound}} \tag{58}$$

$$\tau I - Q_k(t_f, u_d, \bar{u}_k) \geqslant 0 \tag{59}$$

in which u_d and $\bar{u}_k$ are defined as previously, and there is now an additional linear matrix inequality. The SDP defined by Eqs. 57–59 is again iteratively solved in Step 3 of the described algorithm. Note that the formulation using the Fisher information matrix is similar. In addition, it is possible to leave out the guaranteed decrease formulation. The result is that a generic formulation for the minimization of the maximal eigenvalue is obtained which avoids problems of the gradient computation.

Formulation 3: the Presence of an Inverse Matrix in the Optimization Problem. In general, the computation of the inverse of a matrix is computationally expensive and leads to an involved computation of derivatives when used in a classic optimal control approach. In OED, this problem arises when using the A- or the M-criterion. The A-criterion involves the minimization of the trace of the variance–covariance matrix or the minimization of the trace of the inverse of the Fisher information matrix.[2] In some cases, it is cheaper to compute the Fisher information matrix (see[6] for a detailed discussion) and to use the A-criterion, the inverse of this Fisher information matrix needs to be computed. A widely made misconception is that the maximization of the trace of the Fisher information matrix is also called the A-criterion, for example.[4] In SDP, computing the inverse of the Fisher information matrix can be avoided. The optimization problem can first be reformulated in the following way

$$\min_{u_d,X} \operatorname{Tr}(X) \tag{60}$$

subject to

$$F_k(t_f, u_d, \bar{u}_k) - F_{\text{bound}} \geqslant 0 \tag{61}$$

$$X \geqslant F_k(t_f, u_d, \bar{u}_k)^{-1} \tag{62}$$

in which X is an additional matrix decision variable, bounded by the inverse of the Fisher information matrix and F_{bound}^{-1}. Equation 62 can be rewritten as[23]

$$\begin{pmatrix} F_k(t_f, u_d, \bar{u}_k) & I \\ I & X \end{pmatrix} \geqslant 0 \tag{63}$$

which is a linear matrix inequality. The Schur complement of the matrix in 63 turns out to be the inequality used in Eq. 62. This yields the semidefinite optimization problem to be solved in Step 3 of the proposed algorithm. The problem formulation involving the M-criterion is the following

$$\min_{u_d,X,\tau} \tau \tag{64}$$

subject to

$$F_k(t_f, u_d, \bar{u}_k) - F_{\text{bound}} \geqslant 0 \tag{65}$$

$$\begin{pmatrix} F_k(t_f, u_d, \bar{u}_k) & I \\ I & X \end{pmatrix} \geqslant 0 \tag{66}$$

$$\tau \geq X_{ii} \quad \forall i = 1, \ldots, n_p \tag{67}$$

Remark. *All of the above problem formulations are locally convex. So bounds and constraints on the control input u need to be convex too or otherwise be approximated in a convex way. This is also necessary for possible state constraints of the original dynamic systems. These constraints can be taken into account if they are formulated in a convex way or approximated by convex functions. A detailed discussion can be found in.*[9] *Furthermore, if the control action whether to measure or not is included in the OED, the Fisher information matrix approach has the additional advantage that the Fisher information matrix depends linearly on this control action. The optimization formulation including this weighing can be formulated as*[12]

$$\min_{u(\cdot),w(\cdot),y(\cdot),S(\cdot),F(\cdot)} \Phi(F(t_f)) \tag{68}$$

subject to

$$\dot{y}(t) = g(y(t), p, u(t)) \tag{69}$$

$$y(0) = y_0 \tag{70}$$

$$\dot{S}(t) = B(t)S(t) + P(t) \tag{71}$$

$$S(0) = \frac{\partial y_0}{\partial p} \tag{72}$$

$$\dot{F}(t) = w(t)S(t)^{\top} D(t)^{\top} V(t)^{-1} D(t) S(t) \tag{73}$$

$$F(0) = Q_p^{-1} \tag{74}$$

$$0 \geq c_p(y(t), p, u(t)) \tag{75}$$

$$0 \geq c_t(y(t_f)) \tag{76}$$

Numerical and software aspects

In this subsection, the used software and several specific numerical aspects are briefly discussed.

In general, all proposed methods are implemented in MATLAB.[28] To solve the semidefinite problem, a software solution based on a combination of YALMIP[29] and SeDuMi[24] is used iteratively. YALMIP is a modeling language for solving both convex and nonconvex optimization problems. It is a toolbox freely available for MATLAB. The solutions to these optimization problems are computed by external solvers, for example, SeDuMi. SeDuMi is an external solver for optimization problems with linear, quadratic,

and linear matrix inequality constraints. For a detailed description of the subject, the interested reader is referred to.[24]

The dynamic model and the computation of the variance–covariance matrix or Fisher information matrix are implemented and solved in MATLAB. Two computational methods are implemented. First a fourth-order explicit Runge–Kutta integrator is considered. Using a fixed step size allows for the computation of $\partial Q/\partial u$ by finite differences through the internal numerical differentiation approach from.[30] A second integrator is based on implicit Runge–Kutta schemes with efficient sensitivity generation.[31] These implicit integrators are code generated and interfaced with MATLAB through the ACADO Toolkit.[32] To guarantee progress in every iteration a backtracking line search algorithm is implemented.[26] Initially a full step is chosen. However, the reduction parameter of the step in the line search is $\alpha=0.8$. Note that a formulation with an additional regularization term $\frac{\rho}{2}||u-\bar{u}_k||_2^2$ in the objective function, is not used in this article but can be interesting to improve potential convergence problems.

Case Studies

Two different biochemical case studies are introduced in this section. The first subsection discusses a fed-batch bioreactor, whereas in the second subsection the focus is on microbial kinetics in microbiology.

A fed-batch bioreactor model

To benchmark the techniques for OED, a well-mixed fed-batch bioreactor model[33] is used as case study. The dynamic model equations are given by

$$\frac{dC_s}{dt}=-\sigma C_x+\frac{u}{v}C_{s,in}-\frac{u}{v}C_s \tag{77}$$

$$\frac{dC_x}{dt}=\mu C_x-\frac{u}{v}C_x \tag{78}$$

$$\frac{d\mu_{max}}{dt}=0 \tag{79}$$

$$\frac{dK_s}{dt}=0 \tag{80}$$

$$\frac{dv}{dt}=u \tag{81}$$

in which C_s [g/L] is the concentration limiting substrate, C_x [g/L] the biomass concentration, and v [L] the bioreactor volume. Note that the formulation where the unknown parameters are stacked as trivial differential equations is used (Eqs. 11,12), as explained in the second section. The function u [L/h] denotes the volumetric rate of the feed stream, containing a substrate concentration $C_{s,in}$ and is the control function of the case study. The specific growth rate studied in this case is of the monotonic Monod type. The corresponding algebraic relation is given by

$$\mu=\mu_{max}\frac{C_s}{K_s+C_s} \tag{82}$$

The substrate consumption rate is in the case study modeled by an affine dependence which is known as the linear law

$$\sigma=\mu/Y_{X|S}+m, \tag{83}$$

where $Y_{X|S}$ is the yield and m the maintenance factor. The parameter values are given in Table 1. The current best estimate for the parameters μ_{max} and K_s are represented as $\bar{\mu}_{max}$ and $\bar{K}_s$. The initial concentration of substrate and biomass are set to 50 and 1.3125 g, respectively. The initial volume is set to 8 L. The feed rate u is constrained by

Table 1. Parameter Values for the Fed-Batch Bioreactor

$\bar{\mu}_{max}$	0.1 (h^{-1})	$\bar{K}_s$	1 (g/L)
M	0.29 (g/g)	$Y_{X\|S}$	0.47 (g/g)
$\sigma^2_{C_s}$	1×10^{-2}(g^2/L^2)	$\sigma^2_{C_x}$	6.25×10^{-4}(g^2/L^2)

$$0\le u(t)\le 1\mathrm{L/h} \tag{84}$$

It is assumed that the states C_s and C_x can be measured on-line. So the observation function is $h(t)=[C_s,C_x]^\top$. It is assumed that the measurement errors of the states are Gaussian satisfying the modeling assumption. The associated measurement error variance matrix is given as the diagonal matrix $V(t)=\mathrm{diag}\,(\sigma^2_{C_s},\sigma^2_{C_x})^\top$. The initial variances of the states and parameters are

$$Q(0)=\mathrm{diag}\left(10\times\sigma^2_{C_s},10\times\sigma^2_{C_x},0.05(\frac{1}{h})^2,0.5(\frac{g}{L})^2\right)^\top$$

The remaining nondiagonal components of the matrix $Q(0)$ are all 0. Note that the matrix Q is only a 4×4 matrix, as the differential state v is not affected by the uncertainty in the parameters. Note that due to the linearization of the variance–covariance matrix, no reduction in the initial uncertainty of v is anyhow possible, even though a nonlinear observability analysis reveals all states to be observable.

Furthermore, the concentrations of the biomass and substrate can never be negative. For this reason, linearized state constraints are added to the general problem formulation

$$C_{s,k}(t_k,u,\bar{u}_k)=C_{s,k}(t_k,\bar{u}_k)+\sum_{i=1}^{n_u}\frac{\partial C_{s,k}(t_k,\bar{u}_k)}{\partial u_i}(u_i-\bar{u}_{k,i})\ge 0 \tag{85}$$

The constraints for C_x are added in a similar way. For this case study, problem Formulation 1 (minimizing the trace of the variance–covariance matrix) and 2 (minimizing the largest eigenvalue) are considered in the variance–covariance matrix framework. For the complete dynamic OED formulation 15 differential states have to be computed, that is five for the system and 10 for the (symmetric) variance–covariance matrix.

A microbial kinetics model

In this case study, optimal dynamic experiments for estimating the parameters of the Cardinal Temperature Model with Inflection (CTMI)[34] are designed. This CTMI model is a secondary model to the primary growth model of Baranyi and Roberts.[35] This latter model describes the cell density as a function of time, whereas the former incorporates the dependency of the specific growth rate on temperature. The primary model equations are

Table 2. Parameter Values Used for the Design of the Optimal Experiments for the Microbial Kinetics Model

T_{min}	284,48 (K)	T_{opt}	314,0 (K)
T_{max}	319,69 (K)	μ_{opt}	2.397 (h^{-1})
n_{max}	22.55 ln(CFU/mL)	σ_n^2	3.27×10^{-2}ln(CFU/mL)^2

$$\frac{dn}{dt} = \frac{G}{G+1}\mu_{max}(T)[1-\exp(n-n_{max})] \quad (86)$$

$$\frac{dG}{dt} = \mu_{max}(T)G \quad (87)$$

with n [ln(CFU/mL)] the natural logarithm of the cell density and G [−] the physiological state of the cells. The state which can be measured is n, so the observation function $h(t)$ is equal to n. The control input to this system is the temperature profile $T(t)$. The temperature dependency described by the CTMI is given by

$$\mu_{max} = \mu_{opt}\gamma(T) \quad (88)$$

with

$$\gamma(T) = \frac{(T-T_{min})^2(T-T_{max})}{(T_{opt}-T_{min})[(T_{opt}-T_{min})(T-T_{opt})-(T_{opt}-T_{max})(T_{opt}+T_{min}-2T)]} \quad (89)$$

The values of the parameters $p=[\mu_{max}\ T_{min}\ T_{opt}\ T_{max}]^\top$ for the CTMI model[36] are depicted in Table 2. The end time is fixed to 38 h.[36] For model validity reasons the dynamic temperature profiles are constrained to

$$273.15\text{K} \le T(t) \le 318.15\text{K} \quad (90)$$

$$-5\text{K} \le \Delta T \le 5\text{K} \quad (91)$$

The temperature profile is discretized in a piecewise constant manner with the restriction that the temperature can only change 5K in each control action (Eq. 91). In this article, experiments are designed that take all the four parameters into account, simultaneously, which is a significant extension compared to the work presented in.[36] The duration of the microbial lag phase, modeled by the state $G(t)$, is in practice determined by the prior and actual experimental conditions. This means that it cannot be predicted accurately. Therefore, a reduced form of the model of Baranyi and Roberts, that is, without the state $G(t)$ is used in the OED. The model reduces to the logistic growth model which describes exponential growth followed by a stationary phase.[36] Furthermore, for this

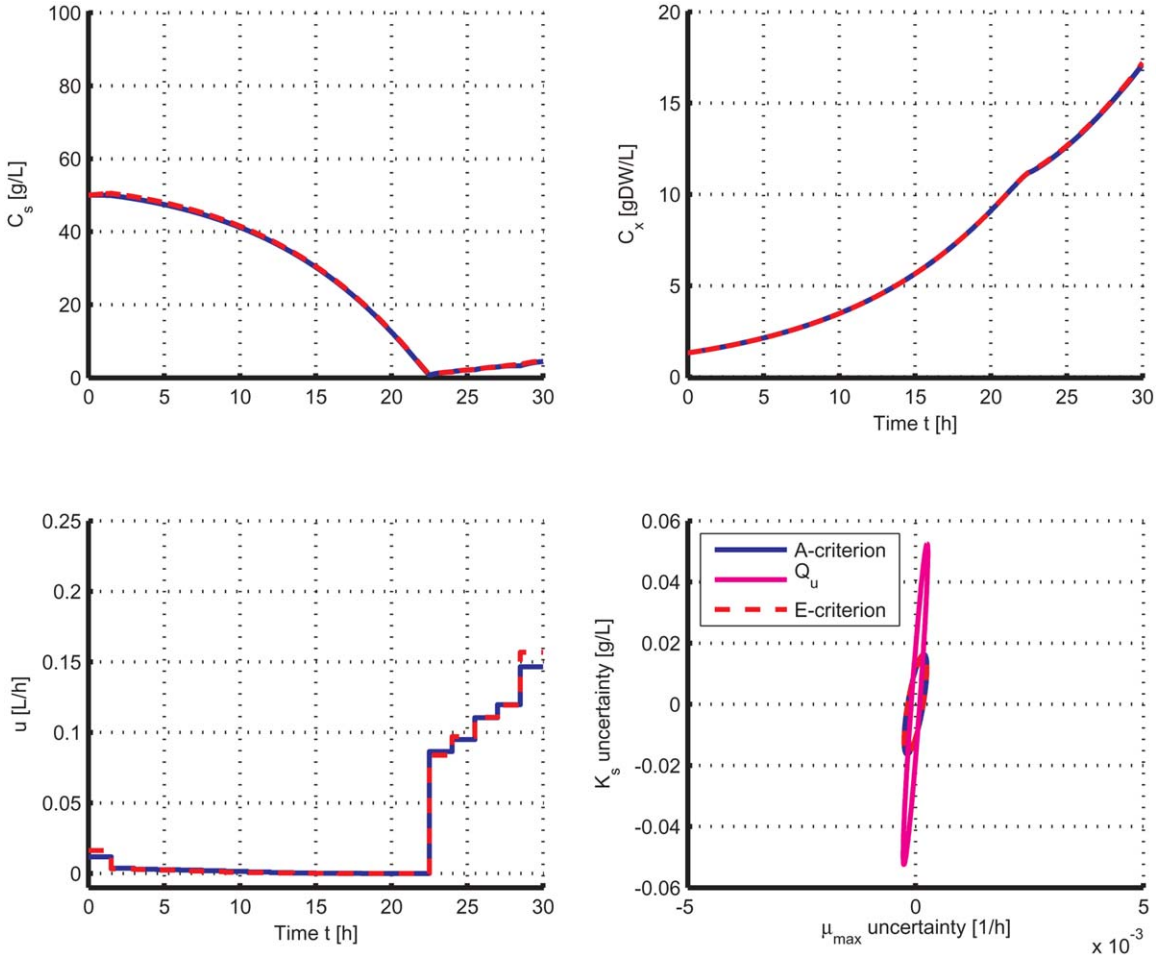

Figure 2. The results for the expected state trajectories, the optimized control input and the joint confidence region for the coinciding cases of minimizing the trace (solid lines), and minimizing the maximum eigenvalue (dashed lines) of the variance-covariance matrix.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

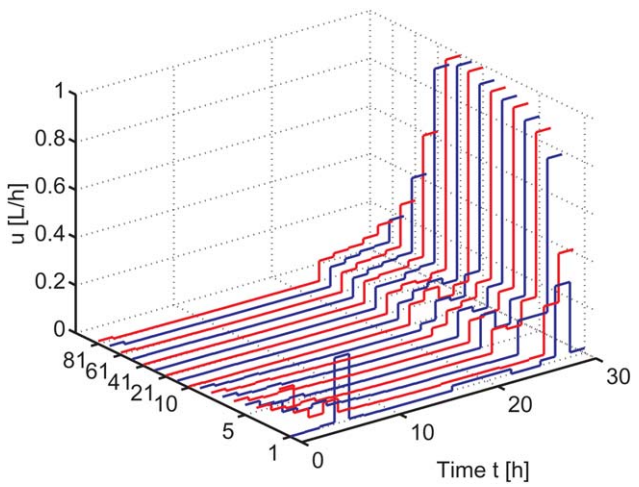

Figure 3. The results for the obtained control input after solving the SDP for the first 11 iterations and subsequently every 10 iterations in the case of the minimization of the trace.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

case study, the Fisher information approach is used because the decision when to measure is also considered in this case study. Problem Formulation 2 and 3 are exploited. The total number of states computed in the case study is 15. The first state originates from the dynamic model, four states for the sensitivity equations and 10 for the symmetric Fisher information matrix. As lower bound on the Fisher information matrix the following constant matrix is taken $F_{\text{bound}} = 0.5I$.

Simulation Results and Discussion

The numerical results are presented and discussed in this section. The fed-batch bioreactor results can be found in the first subsection, those of the microbial kinetics in the second subsection. The experiments presented in this section are performed using ACADO, YALMIP, and SeDuMi from

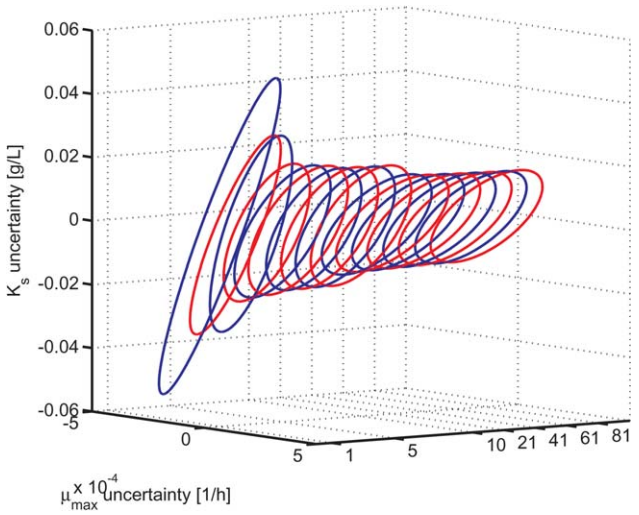

Figure 4. The obtained joint confidence regions after solving the SDP for the first 11 iterations and subsequently every 10 iterations in the case of the minimization of the trace.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Table 3. Overview of the Typical Computational Times for the Two Approaches in the Fed-Batch Bioreactor Case Study

Computational Block/Approach	Finite Differences	Implicit RK
Integration and sensitivity generation	1.5 s	0.037 s
Solution of semidefinite program	0.26 s	0.28 s
Line search and additional integration	0.26 s	0.054 s
Average backtracking steps	1.9	1.6

MATLAB R2011a and this on an ordinary computer (Intel i7-3720QM 6MB cache, 2.60 GHz, 64-bit Ubuntu 12.04).

Fed-batch bioreactor

The obtained state and input profiles for the cases of the minimization of the trace (Formulation 1) and minimization of the maximum eigenvalue of the variance–covariance matrix (Formulation 2) are depicted in Figure 2. From the simulations, it is clear that both objective functions lead to a similar feeding profile. A similar comparison between two objective functions regarding this case study is made in.[6] In the initial part of the experiment, there is a small step in the feeding profile, which gives rise to a small increase in the substrate concentration. Note that the input profile is not zero in the time frame of 1.5 to 22.5 h but is slightly decreasing over this interval. In this time frame, the substrate is consumed while biomass is being formed until 22.5 h and the substrate concentration reaches almost zero. After 22.5 h, there is a feeding phase which leads again to an increase of the substrate and biomass concentration. In Figure 2 the expected, linearized joint confidence regions are also displayed. Q_{bound} denotes the upper bound on the variance–covariance matrix which is passed to the algorithm. This Q_{bound} is the variance–covariance obtained by integrating the initial guess for the control input. From Figure 2, it is difficult to assess the decrease of uncertainty. The joint

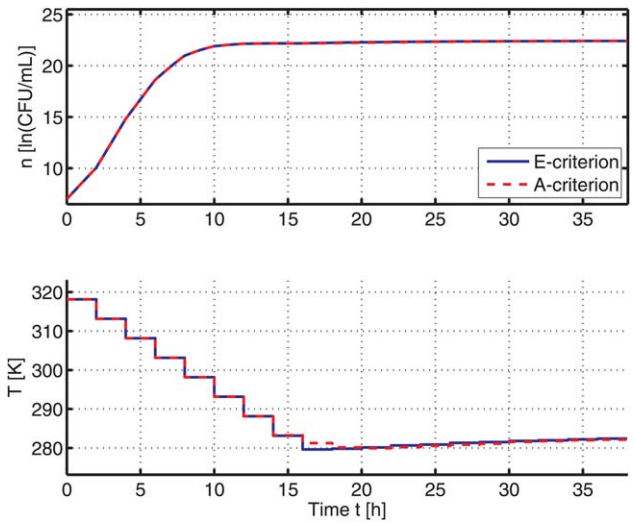

Figure 5. Resulting states and control inputs for the optimization of the E- and A-criterion for the microbial kinetics case study.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

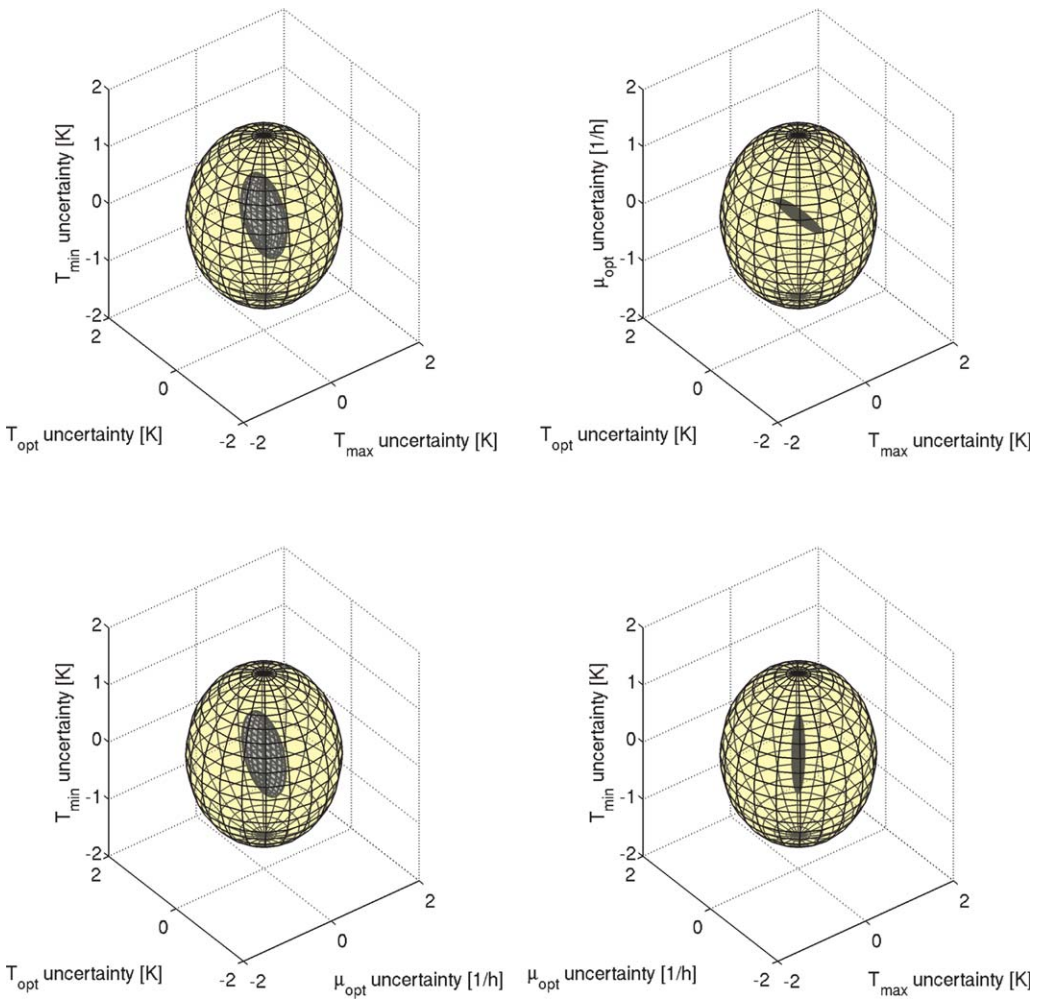

Figure 6. Projections of the 4-D joint confidence region in four 3-D joint regions for the optimal value (dark ellipsoids) and F_{bound} (light ellipsoids) for the maximization of the minimal eigenvalue.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

confidence region area described by the designed experiments is half the value of the joint confidence region described by the initial Q_{bound}. In Figures 3 and 4, the controls and the expected joint confidence region obtained in each iteration of the algorithm are depicted. The first 11 iterations are shown and subsequently only every 10 iterations are illustrated as the changes between the controls and joint confidence regions decrease over the number of iterations.

In the algorithm three different computational blocks can be distinguished: (1) the integration of the dynamic system and the generation of the derivatives, (2) the solution of the SDP by YALMIP, and (3) the line search in which possibly additional forward integrations of the system need to be performed. The computation times reported are for the minimization of the trace of the variance–covariance matrix, the A-criterion (i.e., Formulation 1) but similar values for the E-criterion (i.e., Formulation 2) are observed. An overview is provided in Table 3. The typical time for the integration and derivatives generation part is 1.5 s for the finite differences approach and 0.037 s for the implicit Runge–Kutta scheme. For the solution of the SDP problem this is 0.26 s for the finite differences and 0.28 s for the implicit Runge–Kutta implementation. The line search has a duration of 0.26 and 0.054 s, respectively. In the line search, several additional integrations of the system are needed which explains the amount of time spent in this part of the code. This clearly illustrates that the majority of the time is spent in the integration of the system dynamics, when using the finite difference approach. However, when the fast integrators are used, solving the SDP becomes the slowest part. The implementation with the ACADO integrators is significantly faster. This is mainly due to the substantial amount of time spent in the integration and linearization of the finite differences approach.

Table 4. Overview of the Typical Computational Times for the Implicit Runge–Kutta Approach for the Maximization of the Minimal Eigenvalue of the Fisher Information Matrix in the Microbial Growth Model

Computational Block/Approach	Implicit RK
Integration and sensitivity generation	0.0070 s
Solving semidefinite programming	0.22 s
Line search and additional integration	0.0062 s
Average backtracking steps	0.52

Predictive microbial growth model

Optimization of the E- and A-Criterion. In previous work by[36] some insight has been obtained regarding

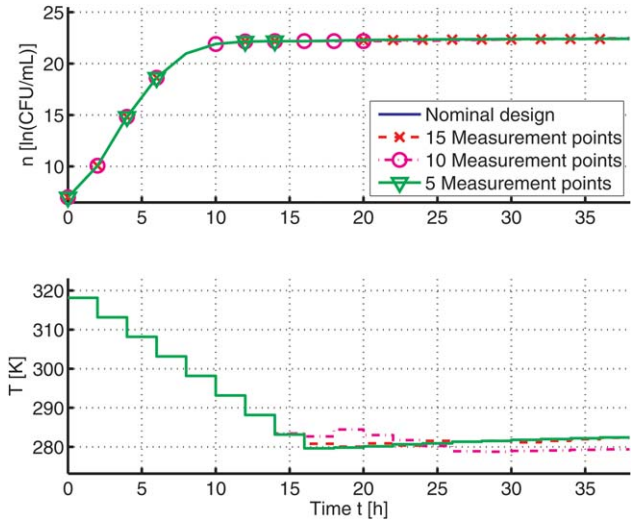

Figure 7. Resulting state and control inputs for the maximization of the minimal eigenvalue while reducing the number of measurements per experiment.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

interesting profiles for the initialization of the OED procedure. In this section, the minimal eigenvalue of the 4 by 4 Fisher information matrix is maximized (i.e., Formulation 2 using the Fisher information matrix approach). Furthermore, the trace of the inverse of the Fisher information matrix is minimized by the approach of Formulation 3 outlined in the third section. The obtained state and temperature profiles are depicted in Figure 5. The temperature profile starts at 318.15 K and subsequently decreases every 2 h with 5 K until the temperature reaches 283.15 K. For the E-criterion there is decrease to 279.65 K, for the A-criterion there is one more intermediate step and a slightly higher minimal temperature of 279.95 K. After this minimal temperature value, there is for both cases a steady increase to almost 283.15 K toward the end of the experiment. The E-criterion experiment has a minimal eigenvalue of 2.08 and the trace of the inverse of the Fisher information matrix is 0.933, whereas the A-criterion results in a minimal eigenvalue of 2.05 and a trace value of 0.927. These observations illustrate how closely related the designed experiments are.

To assess the information content of the E-criterion experiment, the four three-dimensional (3-D) projections of the joint confidence region are displayed in Figure 6. From this figure, it is clear that the joint confidence region of the designed experiment is completely contained by the joint confidence region of the lower bound of F_{bound}. Furthermore, one can infer that for $T_{\min}$ and T_{opt} the information content increase in the optimal experiment is lower than for $T_{\max}$ and μ_{opt}. An overview of the computational time is given in Table 4. Similar values are obtained for the two objective functions, so only the E-criterion is described. As the implicit Runge–Kutta approach is faster, the simulations are only performed with this approach. The time needed for integration is less than in the fed-batch reactor case study. The solution of the SDP is also slightly faster. The line search takes less time than in the previous case study and needs less backtracking steps.

Determination of the Optimal sampling scheme. In this section, both the temperature profile and the sampling scheme are assumed to be the decision variables in the dynamic optimization procedure. Only the maximization of the minimal eigenvalue is considered (Formulation 2). The decision when to sample can be taken into account in the computation of the Fisher information matrix, according to (7). Three different schemes are considered: 5, 10, and 15 measurement points. The obtained temperature profiles, sampling schemes, and state profiles are depicted in Figure 7.

Although, several initializations have been tried, none outperformed the temperature profile obtained in the previous section. So, the obtained temperature and state profiles do not differ a lot. Only after 16 h in the experiment, there is a difference between the different experimental conditions. Note that there is a small difference between the expected state evolutions. Except for the nominal design case and the case where five measurement points are considered, these designs coincide. A second interesting aspect is the decision when to sample. If the input profile is fixed, the resulting optimization problem is convex. As the temperature profile does not differ a lot, the optimization routine focuses mainly on the decision when to sample and can almost be considered to be a convex optimization problem.

In Figure 7, the different points when to sample are also illustrated. The difference between 10 and 15 points is remarkable. Where the 10 points result focuses more on the initial part of the experiment, the 15 measurement points result omits four points in first part of the experiment. The three designs however do share five common points. As the design is not different in the first 16 h, these points can be considered as the five most informative measurement points of the experiment. Note that none of the designs take a measurement at 8 h in the experiment.

Conclusions

In this article, a novel optimization algorithm for OED of nonlinear dynamic processes by sequential SDP is presented. In the presented algorithm, the variance–covariance matrix/Fisher information matrix is linearized and constrained by a linear matrix inequality way, which leads to a SDP. A first advantage of the proposed methodology is that linear matrix inequalities on the variance–covariance matrix can be taken into account. This means that it can be ensured that the expected variance–covariance matrix is better than an initial predetermined value. A second advantage is that the sequential SDP formulation allows an easy formulation of the minimization of the maximum eigenvalue or maximization of the minimal eigenvalue in a direct dynamic optimization formulation. Possible problems with the computation of derivatives of the maximum eigenvalue function are, thus, avoided. A third advantage is that problems involving the minimization of the trace of the inverse of the Fisher information matrix can be reformulated using a linear matrix inequality. This approach avoids the need of computing the derivatives through an inverse of a matrix. The presented methodology is successfully applied to two different (bio)chemical case studies.

Acknowledgments

Dries Telen has a Ph.D. grant of the Agency for Innovation through Science and Technology in Flanders (IWT). Rien Quirynen holds a Ph.D scholarship by the FWO. Jan Van Impe holds the chair Safety Engineering sponsored by

the Belgian Chemistry and Life Sciences Federation essenscia. The research was supported by the KU Leuven Research Fund: PFV/10/002 (OPTEC), OT/10/035, GOA/10/09 (MaNet), GOA/10/11 the KU Leuven Industrial Research Fund: IOF, KP, SCORES4CHEM, FWO: PhD/postdoc grants and projects: FWO KAN2013 1.5.189.13, G.0930.13, G.0320.08, G.0377.09, IWT: PhD grants and project: LeCoPro, the Belgian Federal Science Policy Office: IAP VII/19 (DYSCO), EU: FP7- EMBOCON, ERC ST HIGHWIND.

Literature Cited

1. Espie D, Macchietto S. The optimal design of dynamic experiments. *AIChE J*. 1989;35:223–229.
2. Walter E, Pronzato L. *Identification of Parametric Models from Experimental Data*. Paris: Springer, 1997.
3. Bauer I, Bock G, Körkel S, Schlöder JP. Numerical methods for optimum experimental design in DAE systems. *J Comput Appl Math*. 2000;120:1–25.
4. Franceschini G, Macchietto S. Model-based design of experiments for parameter precision: state of the art. *Chem Eng Sci*. 2008;63:4846–4872.
5. Heidebrecht P, Sundmacher K, Biegler L. Optimal design of nonlinear temperature programmed reduction experiments. *AIChE J*. 2011;57:2888–2901.
6. Telen D, Houska B, Logist F, Vanderlinden E, Diehl M, Van Impe J. Optimal experiment design under process noise using Riccati differential equations. *J Process Control*. 2013;23:613–629.
7. Hoang MD, Barz T, Merchan A, Biegler LT, Arellano-Garcia H. Simultaneous solution approach to model-based experimental design. *AIChE J*. 2013;59:4169–4183.
8. Telen D, Logist F, Vanderlinden E, Van Impe J. Optimal experiment design for dynamic bioprocesses: a multi-objective approach. *Chem Eng Sci*. 2012;78:82–97.
9. Tran Dinh Q, Gumussoy S, Michiels W, Diehl M. Combining convex-concave decompositions and linearization approaches for solving BMIs, with application to static output feedback. *IEEE Trans Automat Control*. 2012;57:1377–1390.
10. Tran Dinh Q, Savorgnan C, Diehl M. Adjoint-based predictor-corrector sequential convex programming for parametric nonlinear optimization. *SIAM J Optim*. 2012;22:1258–1284.
11. Ljung L. *System Identification: Theory for the User*. Upper Saddle River, NJ: Prentice Hall, 1999.
12. Sager S. Sampling decisions in optimum experimental design in the light of Pontryagin's maximum principle. *SIAM J Control Optim*. 2013;51:3181–3207.
13. Pukelsheim F. *Optimal Design of Experiments*. New York: Wiley, 1993.
14. Franceschini G, Macchietto S. Novel anticorrelation criteria for model-based experiment design: theory and formulations. *AIChE J*. 2008;54:1009–1024.
15. Sargent RWH, Sullivan GR. The development of an efficient optimal control package. In: Stoer J, editor. *Proceedings of the 8th IFIP Conference on Optimization Techniques*. Heidelberg: Springer-Verlag, 1978:158–168.
16. Vassiliadis VS, Sargent RWH, Pantelides CC. Solution of a class of multistage dynamic optimization problems. 1. Problems without path constraints. *Ind Eng Chem Res*. 1994;33:2111–2122.
17. Vassiliadis VS, Sargent RWH, Pantelides CC. Solution of a class of multistage dynamic optimization problems. 2. Problems with path constraints. *Ind Eng Chem Res*. 1994;33:2123–2133.
18. Bock HG, Plitt KJ. A multiple shooting algorithm for direct solution of optimal control problems. In: *Proceedings of the 9th IFAC world congress*. Budapest: Pergamon Press, 1984:243–247.
19. Leineweber DB, Bauer I, Bock HG, Schlöder JP. An efficient multiple shooting based reduced SQP strategy for large-scale dynamic process optimization. Part I: theoretical aspects. *Comput Chem Eng*. 2003;27:157–166.
20. Biegler LT. Solution of dynamic optimization problems by successive quadratic programming and orthogonal collocation. *Comput Chem Eng*. 1984;8:243–248.
21. Biegler LT. An overview of simultaneous strategies for dynamic optimization. *Chem Eng Proces: Process Intensificat*. 2007;46:1043–1053.
22. Fares B, Noll D, Apkarian P. Robust control via sequential semidefinite programming. *SIAM J Control Optim*. 2002;40:1791–1820.
23. Vandenberghe L, Boyd S. Semidefinite programming. *Soc Ind Appl Math*. 1996;38:49–95.
24. Sturm JF. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optim Methods Software*. 1999;11–12:625–653. Version 1.05. Available at: http://fewcal.kub.nl/sturm. accessed on June 2013.
25. Toh KC, Todd MJ, Tütüncü RH. SDPT3—a MATLAB software package for semidefinite programming. *Optim Methods Software*. 1999;11:545–581.
26. Nocedal J, Wright SJ. *Numerical Optimization*, 2nd ed. New York: Springer, 2006.
27. Vandenberghe L, Boyd S, Wu, SP. Determinant maximization with linear matrix inequality constraints. *SIAM J Matrix Anal Appl*. 1998;19:499–533.
28. MATLAB. version 7.10.0 (R2010a). MA: The MathWorks Inc., Natick, 2010.
29. Lofberg J. Yalmip: a toolbox for modeling and optimization in matlab. In: *2004 IEEE International Symposium on Computer Aided Control Systems Design*. Taipei, 2004:284–289.
30. Bock HG. Recent advances in parameter identification techniques for ODE. In: Deuflhard P, Hairer E, editors. *Numerical Treatment of Inverse Problems in Differential and Integral Equations*. Boston: Birkhäuser Verlag, 1983:95–121.
31. Quirynen R, Vukov M, Diehl M. Auto generation of implicit integrators for embedded NMPC with microsecond sampling times. In: *Proceedings of the 4th IFAC nonlinear model predictive control conference,* Noordwijkhout, 2012:175–180.
32. Houska B, Ferreau HJ, Diehl M. ACADO Toolkit–an open-source framework for automatic control and dynamic optimization. *Optim Control Appl Methods*. 2011;32:298–312.
33. Versyck KJ, Van Impe JF. Feed rate optimization for bed-batch bioreactors: from optimal process performance to optimal parameter estimation. *Chem Eng Commun*. 1999;172:107–124.
34. Rosso L, Lobry JR, Flandrois, JP. An unexpected correlation between cardinal temperatures of microbial growth highlighted by a new model. *J Theor Biol*. 1993;162:447–463.
35. Baranyi J, Roberts TA. A dynamic approach to predicting bacterial growth in food. *Int J Food Microbiol*. 1994;23:277–294.
36. Van Derlinden E, Bernaerts K, Van Impe J. Simultaneous versus sequential optimal experiment design for the identification of multi-parameter microbial growth kinetics as a function of temperature. *J Theor Biol*. 2010;264:347–355.

Manuscript received Jun. 17, 2013, and revision received Jan. 9, 2014.