ELSEVIER

Journal of Computational and Applied Mathematics 120 (2000) 1–25

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

Numerical methods for optimum experimental design in DAE systems ☆

Irene Bauer, Hans Georg Bock, Stefan Körkel, Johannes P. Schlöder *

Interdisciplinary Center for Scientific Computing (IWR), Im Neuenheimer Feld 368, 69120 Heidelberg, Germany

Received 5 March 1999; received in revised form 6 December 1999

Abstract

Subject of this paper is the design of optimal experiments for chemical processes described by nonlinear DAE models. The optimization aims at maximizing the statistical quality of a parameter estimate from experimental data. This leads to optimal control problems with an unusual and intricate objective function which depends implicitly on first derivatives of the solution of the underlying DAE. We treat these problems by the direct approach and solve them using a structured SQP method. The required first and second derivatives of the solution of the DAE are computed very efficiently by a special coupling of the techniques of internal numerical differentiation and automatic differentiation. The performance of our approach is demonstrated for an application to chemical reaction kinetics. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Optimum experimental design; Parameter estimation; Nonlinear DAE models; Direct approach; Internal numerical differentiation; Chemical reaction kinetics

1. Introduction

Computational simulation and optimization of chemical processes based on mathematical models is essential in today's chemical industry. From the planning of a production plant via the optimization of the production process, regarding, e.g., yield maximization or minimization of the formation of pollutants, up to the training of employees and the simulation of critical scenarios, mathematical tools can be applied to support the decisions of chemical engineers. To reproduce the real-world behaviour properly the processes must be described by mathematical models which are valid over a wide range. In this paper we consider the so-called rigorous models which are based on physical and chemical laws. Especially, we concentrate on processes in reaction kinetics and chemical engineering

☆ Supported by the German Ministry for Education and Research (BMBF), FKZ: 03 D 0043.

* Corresponding author.

PII: S0377-0427(00)00300-9

which can be modeled by systems of differential algebraic equations (DAE). Typically, these DAE systems are nonlinear and stiff.

Often the models depend on parameters the values of which have to be determined from measurements. Typically, these parameters describe specific properties of the reactions and of the involved species, for example reaction orders, reaction rates, velocity constants, equilibrium constants, or reaction enthalpies.

The parameters are estimated by fitting measurement data given from experiments. The conduction of these experiments is often extremely expensive and time consuming. Depending on the design of the experiments the confidence regions of the estimated parameters may be small or large. The aim of optimum experimental design is to determine experiments that yield a parameter estimation with maximum statistical quality. This leads to intricate infinite-dimensional nonlinear optimization problems. This paper deals with the formulation of these problems, the solution procedures, and their capabilities.

Additionally, several constraints have to be considered in the experimental design problem, such as cost limitations, restrictions on the number of measurements, safety and environmental protection rules, or the validity of the model describing the process. The mathematical formulation of these restrictions yields numerous nonlinear and particularly time-dependent equality and inequality constraints.

Optimum experimental design applied to linear models has been investigated already for a long time, see, e.g., the textbooks of Atkinson and Donev [1] or Pukelsheim [22]. Nowadays, these methods are frequently used in industry.

In contrast to this, there are only a few papers on optimum experimental design for nonlinear models and especially for models describing dynamic processes. Rudolph and Herrendörfer [23] applied experimental design to nonlinear regression models used in long-term selection. For a system with three unknown model parameters, the optimum experimental design is calculated analytically. Baltes et al. [2] studied a nonlinear dynamical system in unstructured growth models. To design optimal experiments they set up an unconstrained nonlinear optimization problem which they solved by a simplex algorithm, see [20]. Nonlinear optimum experimental design for chemical reaction systems was treated by Lohmann et al. [18,19]. He considered the optimal selection of the measurements but the variables describing the behavior of the process were kept fixed. Consequently, his optimization problems have a simpler structure than the ones presented in this paper. Hilf [16] developed and implemented methods for nonlinear optimum experimental design for mechanical problems described by DAE systems. In this case, he only had to consider optimum experimental design problems in which the underlying parameter estimation problems are unconstrained. Moreover, the differential equations describing the mechanical models are not stiff.

Our paper is structured as follows: in Section 2, the experimental design optimization problem is formulated, resulting in a nonlinear state-constrained optimal control problem in DAE systems where the objective function implicitly depends on first derivatives of the solution of the underlying constrained parameter estimation problem. In Section 3, we present a direct approach to solve these problems. In Section 4, we introduce the SQP method which we use for the numerical solution and we describe the special requirements in the case of optimum experimental design. How we compute the required derivatives of the objective function and the constraints using techniques of internal numerical differentiation combined with automatic differentiation and semi-analytic formulas is explained in Section 5. In Section 6, we present a sequential procedure for optimum

experimental design and parameter estimation and briefly sketch a method based on multiple shooting to solve the occurring constrained parameter estimation problems numerically. Finally, in Section 7, we give applications of our methods to an example from chemical reaction kinetics, the reaction of urethane.

2. Formulation of the optimum experimental design problem

The aim of optimum experimental design is to plan experiments in order to maximize the statistical reliability of some variables estimated from the experimental data. These variables are typically the unknown parameters in the process model – here we consider DAE systems – but may also be additional characteristic quantities such as, e.g., initial values of the state variables or other trajectory values that may be important.

2.1. Parameter estimation

We first describe the underlying constrained parameter estimation problem and consider an augmented formulation which already contains the quantities required for optimum experimental design. Typically, not all unknown variables can be estimated from a single experiment with sufficient reliability. Therefore, we consider parameter estimation problems for multiple experiments containing N_{ex} single experiments:

$$\min_{x,p} \sum_{i=1}^{N_{\mathrm{ex}}} \sum_{j=1}^{n_{m_i}} w_{ij} \cdot \frac{(\eta_{ij} - b_{ij}(t_{ij}, x_i(t_{ij}), p, q_i))^2}{\sigma_{ij}^2} \tag{1}$$

subject to for $(i = 1, \ldots, N_{\mathrm{ex}})$:

$x_i = (y_i, z_i)$ fulfills the DAE system

$$\dot{y}_i = f_i(t, x_i, p, q_i, u_i), \tag{2}$$

$$0 = g_i(t, x_i, p, q_i, u_i) \tag{3}$$

$+$ additional interior point constraints

$$d_i(x_i(t_{i,0}), \ldots, x_i(t_{i,f_i}), p, q_i) = 0. \tag{4}$$

A least-squares functional is minimized subject to the parameters $p \in \mathbb{R}^{n_p}$ and the differential and algebraic state variables $y_i : [t_{i,0}; t_{i,f_i}] \to \mathbb{R}^{n_{y_i}}$ and $z_i : [t_{i,0}; t_{i,f_i}] \to \mathbb{R}^{n_{z_i}}$ in order to approximate the experimental data best. We consider experimental data

$$\eta_{ij} = b_{ij}(t_{ij}, x_i(t_{ij}), p, q_i) + \varepsilon_{ij}, \quad i = 1, \ldots, N_{\mathrm{ex}}, \ j = 1, \ldots, n_{m_i} \tag{5}$$

with measurement values $\eta_{ij} \in \mathbb{R}$ and the corresponding values of the model responses $b_{ij}(t_{ij}, x_i(t_{ij}), p, q_i)$. The measurement errors $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_{N_{\mathrm{ex}}})$, $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{in_{m_i}})$, $i = 1, \ldots, N_{\mathrm{ex}}$, are assumed to be independent and normally distributed with expected value zero and known variance–covariance matrix $\Sigma^2 = \mathrm{diag}(\sigma_{ij}^2)$:

$$\varepsilon \sim \mathcal{N}(0, \Sigma^2).$$

The weights $w_i \in \{0,1\}^{n_{m_i}}$ describe a selection of measurement points out of the possible ones. They are fixed in the parameter estimation problem as well as the control variables $q_i \in \mathbb{R}^{n_{q_i}}$ and the control profiles $u_i : [t_{i,0}; t_{i,f_i}] \to \mathbb{R}^{n_{u_i}}$. The measurement values η_{ij} and the standard deviations $\sigma_i \in \mathbb{R}^{n_{m_i}}$ are given from the experimental data.

Note that all variables besides the parameters p are only locally defined in each experiment. The model equations (2) and (3) of the DAE system may differ from one experiment to another as well as the constraints and the dimensions of the variables y_i, z_i, q_i, u_i, w_i, and σ_i.

2.2. Parametrization of the solution of the DAE system

The constrained parameter estimation problem (1)–(4) depends on the solution functions $x_i : [t_{i,0}; t_{i,f_i}] \to \mathbb{R}^{n_{y_i}+n_{z_i}}$, $i = 1, \ldots, N_{\mathrm{ex}}$, of the DAE system (2) and (3). We use a direct approach to transfer the infinite-dimensional optimization problem to a finite-dimensional one. Therefore, we introduce additional variables s_i, $i = 1, \ldots, N_{\mathrm{ex}}$, which, e.g., can be the initial values of the DAE system at time $t = t_{i,0}$. We add the consistency conditions

$$g_i(t_{i,0}, s_i, p, q_i, u_i) = 0, \quad i = 1, \ldots, N_{\mathrm{ex}}, \tag{6}$$

to the vector of constraints d_i, $i = 1, \ldots, N_{\mathrm{ex}}$, of the parameter estimation problem. The solution of the DAE system can now be considered as a function of the finite-dimensional variables s_i (of course also depending on the parameters p and the fixed controls q and u), which is computed by an integration method solving the DAE system with initial values $x(t_{i,0}) = s_i$. Possible additional constraints for the variables s_i are also added to the constraints d_i.

Thus we can write the parameter estimation problem (1)–(4) in an equivalent form as a finite-dimensional constrained least-squares problem in the variables s and p:

$$\begin{aligned} &\min_{p,s} \sum_{i=1}^{N_{\mathrm{ex}}} \sum_{j=1}^{n_{m_i}} w_{ij} \cdot \frac{(\eta_{ij} - b_{ij}(t_{ij}, x_i(t_{ij}, p, s_i, q_i, u_i), p, q_i))^2}{\sigma_{ij}^2} \\ &\quad =: \sum_{i=1}^{N_{\mathrm{ex}}} \|r_i(p, s_i, q_i, u_i, w_i)\|_2^2 =: \|r(p, s, q, u, w)\|_2^2 \end{aligned} \tag{7}$$

subject to

$$d_i(p, s_i, q_i, u_i) = 0, \quad i = 1, \ldots, N_{\mathrm{ex}}. \tag{8}$$

Remark. Like the parameters p, also the introduced variables s_i are random variables. If the user is not only interested in the reliability of the model parameters and initial values, but also in the reliability of the trajectories of some important components, additional variables $s_{i,j}$ at some interior

points $t_{i,j} \in [t_{i,0}; t_{i,f_i}]$ may be introduced. The underlying parameter estimation problem is then of the form (29)–(32) as described in Section 6.2.

2.3. Variance–covariance matrix

We linearize the constrained parameter estimation problem (7)–(8) at a given point $(\hat{p}, \hat{s})$ and obtain a linear constrained least-squares problem in the unknowns $(\Delta p, \Delta s)$:

$$\min_{\Delta p, \Delta s} \left\| J_1(\hat{p}, \hat{s}) \cdot \begin{pmatrix} \Delta p \\ \Delta s \end{pmatrix} + r(\hat{p}, \hat{s}) \right\|_2^2$$

subject to

$$J_2(\hat{p}, \hat{s}) \cdot \begin{pmatrix} \Delta p \\ \Delta s \end{pmatrix} + d(\hat{p}, \hat{s}) = 0,$$

where the matrix

$$J = \begin{pmatrix} J_1 \\ J_2 \end{pmatrix} = \begin{pmatrix} \dfrac{\mathrm{d}r}{\mathrm{d}(p,s)}(\hat{p}, \hat{s}) \\ \dfrac{\mathrm{d}d}{\mathrm{d}(p,s)}(\hat{p}, \hat{s}) \end{pmatrix}$$

denotes the Jacobian of the constrained parameter estimation problem (7)–(8). In vector $d = (d_1^{\mathrm{T}}, \ldots, d_{N_{\mathrm{ex}}}^{\mathrm{T}})^{\mathrm{T}}$ we summarize the equality constraints.

A necessary condition for $(\hat{p}, \hat{s})$ to be a local minimum of (7)–(8) is that

$$J^{+}(\hat{p}, \hat{s}) \cdot \begin{pmatrix} r(\hat{p}, \hat{s}) \\ d(\hat{p}, \hat{s}) \end{pmatrix} = 0,$$

where J^{+} denotes the generalized inverse

$$J^{+} = \begin{pmatrix} I & 0 \end{pmatrix} \begin{pmatrix} J_1^{\mathrm{T}} J_1 & J_2^{\mathrm{T}} \\ J_2 & 0 \end{pmatrix}^{-1} \begin{pmatrix} J_1^{\mathrm{T}} & 0 \\ 0 & I \end{pmatrix} \tag{9}$$

of the Jacobian J of the constrained parameter estimation problem.

As the experimental data is random the estimated solution $(\hat{p}, \hat{s})$ of the parameter estimation problem is also a random variable which is in a first-order approximation normally distributed

$$(\hat{p}, \hat{s}) \sim \mathcal{N}((p^*, s^*), C)$$

with the (unknown) true value (p^*, s^*) as expected value and variance–covariance matrix

$$C = C(\hat{p}, \hat{s}, q, u, w) = J^{+} \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} J^{+^{\mathrm{T}}}. \tag{10}$$

The variance–covariance matrix describes the confidence ellipsoid which is an approximation of a nonlinear confidence region of the estimated variables. The $100\alpha\%$ confidence ellipsoid $(0 \leqslant \alpha \leqslant 1)$

can be described by

$$G_L(\alpha;\hat{p},\hat{s},q,u,w)=\{(p,s):(p-\hat{p},s-\hat{s})^{\mathrm{T}}C(\hat{p},\hat{s},q,u,w)(p-\hat{p},s-\hat{s})\leqslant\gamma^2(\alpha)\}.$$

The probability factor $\gamma(\alpha)$ is given by

$$\gamma^2(\alpha)=\chi^2_{l_1-l_2}(1-\alpha),$$

where l_1 is the dimension of (p,s), l_2 is the dimension of the constraints d, and $\chi^2_{l_1-l_2}(1-\alpha)$ is the quantile of the χ^2-distribution.

2.4. Experimental design optimization problem

Optimum experimental design aims at maximizing the reliability of the variables to be estimated by the underlying constrained parameter estimation problem. Therefore we minimize a design criterion on the variance–covariance matrix C. As described above, the infinite-dimensional solution of the DAE system is parametrized by finite-dimensional variables s.

The optimum experimental design approach then leads to an optimal control problem of the following form:

$$\min_{q,u,w,s}\ \phi(C(p,s,q,u,w)) \tag{11}$$

subject to (for every experiment i, $i=1,\ldots,N_{\mathrm{ex}}$):

control and path constraints

$$c_i(t,x_i(t,p,s_i,q_i,u_i),p,q_i,u_i(t),w_i)\geqslant 0,\quad t\in[t_{i,0};t_{i,f_i}], \tag{12}$$

constraints of the underlying parameter estimation problem

$$d_i(p,s_i,q_i,u_i)=0 \tag{13}$$

and the integrality constraints

$$w_i\in\{0,1\}^{n_{m_i}}. \tag{14}$$

Note that for every evaluation of the objective function and of the state constraints the solutions of the underlying DAE systems are required. For the objective function additionally derivatives of the solution of the DAE with respect to the parameters have to be computed.

Remark. Formally, we have introduced all initial values as variables s to obtain a finite-dimensional formulation of the optimization problem. If parts of the initial values are known in advance, they can be eliminated a priori in order to save computational effort.

Free variables of the optimization problem are the control profiles $u_i(t)$, the time-independent control variables q_i, and the weights w_i on possible measurements. The elements σ_{ij} of the variance–covariance matrix Σ^2 of the measurement errors are now no longer considered as fixed values. This takes into account that the standard deviations of the measurement methods may depend on the experimental settings q_i or on the states x_i: $\sigma_{ij}=\varsigma_{ij}(x_i,q_i)$ with $\varsigma_{ij}:\mathbb{R}^{n_{y_i}+n_{z_i}}\times\mathbb{R}^{n_{q_i}}\to\mathbb{R}^{n_{m_{ij}}}$. Thus we are able to formulate also measurements with a relative or more complex measurement error.

Possible design criteria to minimize the confidence ellipsoid are:

- *A-criterion:* Minimizing the average half-axis length of the confidence ellipsoid resp. the average variance of the parameters

$$\phi_{\mathrm{A}}(C) = \frac{1}{l_1}\mathrm{trace}(C), \quad l_1 = n_p + n_s.$$

- *E-criterion:* Minimizing the largest expansion of the confidence ellipsoid

$$\phi_{\mathrm{E}}(C) = \lambda_{\max}(C),$$

where $\lambda_{\max}$ denotes the largest eigenvalue of C.

- *D-criterion:* Minimizing the volume of the confidence ellipsoid.
In case of constrained parameter estimation problems we restrict on a regular submatrix of the variance–covariance matrix C corresponding to a linear combination $K^{\mathrm{T}} \cdot (\hat{p}, \hat{s})$ of the variables of the parameter estimation problem. Matrix $K \in R^{l_1 \times l_K}$ is of full rank. We obtain a restricted variance–covariance matrix $C_K = \mathrm{Cov}(K^{\mathrm{T}}(\hat{p}, \hat{s})) = K^{\mathrm{T}} \cdot \mathrm{Cov}(\hat{p}, \hat{s}) \cdot K$. For regular $C_K \in \mathbb{R}^{l_K \times l_K}$ we apply the D-criterion to the matrix C_K

$$\phi_D(C_K) = \det(C_K)^{1/l_K}.$$

In principle, also other functions on the variance–covariance matrix can be treated, e.g., minimizing the maximal edge of a box enclosing the confidence ellipsoid, i.e., minimizing the maximal (scaled) variance of the parameters, or criteria defined on the off-diagonals of the variance–covariance matrix or on the correlation matrix, minimizing correlations of the parameters.

3. Direct approach

To solve the state-constrained optimal control problem (11)–(14) numerically we use a direct approach. We parametrize the control functions and discretize the solution of the DAE system and the state constraints to obtain a finite-dimensional constrained nonlinear optimization problem.

3.1. Parametrizing the control functions

The control functions are parametrized on an appropriate user-defined grid

$$t_{i,0} = \tau_{i,0} < \cdots < \tau_{i,n_{\tau_i}} = t_{i,f_i} \tag{15}$$

and replaced by piecewise continuous functions with a finite-dimensional number of degrees of freedom. Typically, these functions are piecewise constant or piecewise linear, but also other parametrizations are possible. If continuity at grid points $\tau_{i,j}$ is required, additional constraints have to be introduced.

In the following we denote by the vector q not only the time-independent control variables but also the the additional control variables introduced by the parametrization of the control functions.

3.2. Discretization of state constraints and state variables

The state constraints in (12) are discretized on an appropriate grid defined by the user. They are replaced by interior point constraints on this grid. The grid does not need to be the same as for the parametrization of the control functions.

The solution of the DAE system is discretized applying methods of variable step size and order based on backward differentiation formulae. The algebraic equations (3) of the DAE system depend on the control variables q which vary in every iteration of the optimization process. Thus the algebraic equations will normally not be consistent. Therefore, we use a relaxed formulation of the algebraic equations (see, e.g., [12])

$$0 = g_i(t, x_i, p, q_i) - \vartheta_i(t) g_i(t_{i,0}, x_i(t_{i,0}), p, q_i) \tag{16}$$

instead of (3) which allows us to solve the initial value problem for the modified DAE system (2) and (16) with consistent initial values $x_i(t_{i,0})$. The damping factor $\vartheta_i(t)$, with $\vartheta_i(t_{i,0}) = 1$ and $0 \leqslant \vartheta_i(t) \leqslant 1$ for $t \in [t_{i,0}, t_{i,f_i}]$, should be non-increasing and may be chosen, e.g., as $\vartheta_i(t) \equiv 1$ or as $\vartheta_i(t) = \mathrm{e}^{-\theta(t-t_{i,0})}$ with $\theta > 0$.

Consistent initial values for the original DAE system are eventually ensured by the consistency conditions

$$g_i(t_{i,0}, x_i(t_{i,0}), p, q_i) = 0 \tag{17}$$

which we add to the constraints of the optimization problem.

3.3. Relaxation of the integrality constraints

The weights w on the measurements are 0–1-variables:

$$w_i \in \{0,1\}^{n_{m_i}}, \quad i = 1, \ldots, N_{\mathrm{ex}}. \tag{18}$$

For the numerical solution of the optimization problem, we use a relaxed formulation of these integrality constraints:

$$w_i \in [0,1]^{n_{m_i}}, \quad i = 1, \ldots, N_{\mathrm{ex}}. \tag{19}$$

A measurement with a weight 0.5, e.g., may be interpreted as a measurement with twice the variance as a measurement with weight 1. In an analogous way, also measurements with weights > 1 can be treated if desired. To obtain an integer solution we apply rounding heuristics to the solution of the relaxed problem if necessary. For example if we only have to fulfill restrictions on the number of measurements for each measurement method, we can round up the largest weights belonging to a measurement method and round off the smallest ones keeping the sum of all weights below or equal to the allowed maximum number. Another possibility is to accumulate the weights for each measurement method for increasing measurement times. If the accumulated sum is smaller than 1, round the current weight off to 0. If the sum reaches 1, round the current weight up to 1, decrease the sum by 1 and continue accumulating. Our computational results have shown that the solutions of the relaxed problems tend to be almost integer. Rounding in the described ways affects the value of the optimality criterion only slightly.

4. Numerical solution of the nonlinear optimization problem

We obtain a nonlinear finite-dimensional equality and inequality constrained optimization problem which we write in abbreviated form as

$$\begin{aligned} \min_{v} \quad & F(v) \\ \text{s.t.} \quad & G(v) = 0, \\ & H(v) \geqslant 0. \end{aligned} \tag{20}$$

In the vector v we summarize the parametrized control variables q, the variables s describing the parametrization of the solution of the DAE system, and the weights w on the measurements. The consistency conditions (17) of the DAE system are contained in the equality constraints $G(v) = 0$.

We choose the method of sequential quadratic programming (SQP) for the solution of the optimization problem (20). The general approach of this method is as follows.

We start with an initial guess v^0 and calculate in a loop the subsequent iterate

$$v^{k+1} = v^k + r^k \Delta v^k, \quad k = 0, 1, \ldots . \tag{21}$$

The step length $r^k \in (0, 1]$ is determined by a line-search. The search direction Δv^k solves the quadratic subproblem (QP)

$$\begin{aligned} \min_{\Delta v} \quad & \tfrac{1}{2} \Delta v^{\mathrm{T}} A^k \Delta v + \nabla F(v^k)^{\mathrm{T}} \Delta v \\ \text{s.t.} \quad & G(v^k) + \nabla G(v^k)^{\mathrm{T}} \Delta v = 0, \\ & H(v^k) + \nabla H(v^k)^{\mathrm{T}} \Delta v \geqslant 0. \end{aligned} \tag{22}$$

Here the matrix A^k denotes an approximation of the Hessian of the Lagrangian function

$$L(v, \lambda, \mu) = F(v) - \lambda^{\mathrm{T}} G(v) - \mu^{\mathrm{T}} H(v),$$

where λ and μ are the Lagrangian multipliers.

5. Derivative generation

The solution of the optimization problem with an SQP-method requires the evaluation of the objective function F and of the equality and inequality constraints G and H. Furthermore, the gradient of the objective function ∇F and the derivatives of the constraints ∇G and ∇H with respect to the optimization variables v are needed. The approximation of the Hessian A^k is updated from step to step. The modifications are based on derivatives of the Lagrangian function.

We obtain the derivatives of the constraints c, d, and g with respect to the optimization variables $v = (s, q, w)$ by applying the chain rule

$$\frac{\mathrm{d}(c, d, g)}{\mathrm{d}v} = \frac{\partial(c, d, g)}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial(c, d, g)}{\partial v}. \tag{23}$$

The derivatives of the constraint functions with respect to x and v are calculated by automatic differentiation, see [9]. The calculation of the derivatives of the solution trajectory $x(t)$ with respect to the variables s and the controls q is explained in Section 5.4. Since the variable $x(t)$ does not depend on w, the derivative of x with respect to the weights w is trivially equal to zero.

It remains to derive formulas needed for the computation of the derivatives of the objective function with respect to the optimization variables $v = (s, q, w)$. Since in most applications only directional derivatives are required, only $(\mathrm{d}\phi/\mathrm{d}v)\Delta v$ with a suitable direction Δv has to be computed.

Using the chain rule we obtain

$$\Delta\phi := \frac{\mathrm{d}\phi}{\mathrm{d}v}\Delta v = \frac{\mathrm{d}\phi}{\mathrm{d}C}\frac{\mathrm{d}C}{\mathrm{d}J}\frac{\mathrm{d}J}{\mathrm{d}v}\Delta v. \tag{24}$$

In the following sections, we describe in detail how the required terms

$$\Delta J := \frac{\mathrm{d}J}{\mathrm{d}v}\,\Delta v, \quad \Delta C := \frac{\mathrm{d}C}{\mathrm{d}J}\,\Delta J, \quad \Delta\phi = \frac{\mathrm{d}\phi}{\mathrm{d}C}\,\Delta C$$

are calculated.

5.1. Derivative of the design criteria with respect to the variance–covariance matrix

For the derivative of the design criteria ϕ with respect to the variance–covariance matrix C explicit formulas are available (see, e.g., [21]):

- *A-criterion*:

$$\frac{\mathrm{d}\phi_{\mathrm{A}}}{\mathrm{d}C}\,\Delta C = \frac{1}{l_1}\cdot \mathrm{trace}(\Delta C).$$

- *E-criterion*: Let $\phi_{\mathrm{E}}(C) = \lambda_{\max}(C)$ be a single eigenvalue of C and let z be the corresponding normalized eigenvector which is unique up to the sign. Then

$$\frac{\mathrm{d}\phi_{\mathrm{E}}}{\mathrm{d}C}\Delta C = z^{\mathrm{T}}\Delta C z.$$

- *D-criterion*:

$$\frac{\mathrm{d}\phi_{\mathrm{D}}}{\mathrm{d}C}\,\Delta C = \frac{1}{l_K}\,(\det C_K)^{1/l_K}\sum_{i=1}^{l_K}\sum_{j=1}^{l_K}(C_K^{-1})_{ij}\cdot(K^{\mathrm{T}}\cdot\Delta C\cdot K)_{ij}.$$

5.2. Derivative of the variance–covariance matrix with respect to the Jacobian

In this subsection, the calculation of the derivative of the variance–covariance matrix C with respect to the Jacobian J of the underlying parameter estimation problem and the corresponding numerical implementations are discussed. The variance–covariance matrix

$$C = J^{+}\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} J^{+^{\mathrm{T}}}$$

may be written as

$$C = (I \quad 0)\begin{pmatrix} J_1^{\mathrm{T}}J_1 & J_2^{\mathrm{T}} \\ J_2 & 0 \end{pmatrix}^{-1}\begin{pmatrix} J_1^{\mathrm{T}}J_1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} J_1^{\mathrm{T}}J_1 & J_2^{\mathrm{T}} \\ J_2 & 0 \end{pmatrix}^{-1}\begin{pmatrix} I \\ 0 \end{pmatrix}$$

$$=: I_1 \cdot M^{-1} \cdot S \cdot M^{-1} \cdot I_1^{\mathrm{T}}.$$

Some technical manipulations yield

$$C = (I_1 M^{-1})((I_1 M^{-1})S)^{\mathrm{T}}$$

for the variance–covariance matrix and

$$\frac{\mathrm{d}C}{\mathrm{d}J}\Delta J = (I_1 M^{-1})((I_1 M^{-1})(-(SI_1^{\mathrm{T}}(I_1 M^{-1})\Delta M)^{\mathrm{T}} + \Delta S - (SI_1^{\mathrm{T}}(I_1 M^{-1})\Delta M)^{\mathrm{T}})^{\mathrm{T}}$$

for its derivative, with

$$\Delta M = \begin{pmatrix} \Delta(J_1^{\mathrm{T}}J_1) & \Delta J_2^{\mathrm{T}} \\ \Delta J_2 & 0 \end{pmatrix} = \begin{pmatrix} \Delta J_1^{\mathrm{T}}J_1 + J_1^{\mathrm{T}}\Delta J_1 & \Delta J_2^{\mathrm{T}} \\ \Delta J_2 & 0 \end{pmatrix}$$

and

$$\Delta S = \begin{pmatrix} \Delta(J_1^{\mathrm{T}}J_1) & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \Delta J_1^{\mathrm{T}}J_1 + J_1^{\mathrm{T}}\Delta J_1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Both for the calculation of the variance–covariance matrix C and for its derivative, formally $(I_1 M^{-1})$ has to be multiplied with different matrices. To realize this operation, we use a generalized QR decomposition of the Jacobian J. The matrices J_1 and J_2 have to be decomposed only once.

5.3. Derivative of the Jacobian with respect to the optimization variables

It remains to establish formulas for the computation of the derivatives of the matrices J_1 and J_2 with respect to the optimization variables v.

We write the Jacobian J_1 of the least-squares problem (7) in abbreviated form as

$$J_1 = \frac{\mathrm{d}r}{\mathrm{d}(p,s)} = -W\Sigma^{-1}\left(\frac{\partial b}{\partial x}\frac{\partial x}{\partial(p,s)} + \frac{\partial b}{\partial(p,s)}\right)$$

with $W = \mathrm{diag}(\sqrt{w_{ij}})$, $\Sigma = \mathrm{diag}(\varsigma_{ij}(x_i, q_i))$, and the model response function $b = (b_{ij}(t_{ij}, x_i(t_{ij}), p, q_i))$ as introduced in (5).

For the derivative of row ij of the Jacobian J_1 with respect to the control variables q we obtain

$$\begin{aligned}\frac{\mathrm{d}J_{1ij}}{\mathrm{d}q}\Delta q = -W\ \Bigg(&\frac{\partial}{\partial x}\left(\frac{1}{\varsigma_{ij}}\frac{\partial b_{ij}}{\partial x}\right)\frac{\partial x}{\partial(p,s)}\frac{\partial x}{\partial q}\Delta q + \frac{\partial}{\partial q}\left(\frac{1}{\varsigma_{ij}}\frac{\partial b_{ij}}{\partial x}\right)\frac{\partial x}{\partial(p,s)}\Delta q \\ &+ \frac{\partial}{\partial x}\left(\frac{1}{\varsigma_{ij}}\frac{\partial b_{ij}}{\partial(p,s)}\right)\frac{\partial x}{\partial q}\Delta q + \frac{\partial}{\partial q}\left(\frac{1}{\varsigma_{ij}}\frac{\partial b_{ij}}{\partial(p,s)}\right)\Delta q \\ &+ \frac{1}{\varsigma_{ij}}\frac{\partial b_{ij}}{\partial x}\frac{\partial^2 x}{\partial q\partial(p,s)}\Delta q\Bigg).\end{aligned}$$

To compute the terms $(\partial/\partial x)((1/\varsigma_{ij})(\partial b_{ij}/\partial x))$, etc. we generate subroutines evaluating $((1/\varsigma_{ij})(\partial b_{ij}/\partial x))$, etc. which we then differentiate by automatic differentiation.

The derivative w.r.t. the variables s is established analogously.

Finally, we have to provide the derivative of $J_1^{\mathrm{T}}J_1$ with respect to the weights w. Let $\Delta w_r := e_r$. Then

$$\frac{\mathrm{d}}{\mathrm{d}w}(J_1^{\mathrm{T}}J_1)\Delta w_r = \left(e_r^{\mathrm{T}}\Sigma^{-1}\left(\frac{\partial b}{\partial x}\frac{\partial x}{\partial(p,s)} + \frac{\partial b}{\partial(p,s)}\right)\right)^{\mathrm{T}}\left(e_r^{\mathrm{T}}\Sigma^{-1}\left(\frac{\partial b}{\partial x}\frac{\partial x}{\partial(p,s)} + \frac{\partial b}{\partial(p,s)}\right)\right),$$

where e_r is the rth unit vector.

The derivatives of the constraints

$$d = (d_1^{\mathrm{T}}, \ldots, d_{N_{\mathrm{ex}}}^{\mathrm{T}})^{\mathrm{T}}, \quad d_i = d_i(x_i(t_{i,0}), \ldots, x_i(t_{i,f_i}), p, s_i, q_i))$$

of the underlying parameter estimation problem with respect to the control variables q (and analogously the variables s) can be computed as

$$\begin{aligned}\frac{\mathrm{d}J_2}{\mathrm{d}q} &= \frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{\mathrm{d}d}{\mathrm{d}(p,s)}\right) = \frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{\partial d}{\partial x}\frac{\partial x}{\partial(p,s)} + \frac{\partial d}{\partial(p,s)}\right)\\ &= \frac{\partial^2 d}{\partial x^2}\frac{\partial x}{\partial(p,s)}\frac{\partial x}{\partial q} + \frac{\partial^2 d}{\partial q \partial x}\frac{\partial x}{\partial(p,s)} + \frac{\partial^2 d}{\partial x \partial(p,s)}\frac{\partial x}{\partial q} + \frac{\partial^2 d}{\partial q \partial(p,s)} + \frac{\partial d}{\partial x}\frac{\partial^2 x}{\partial q \partial(p,s)}.\end{aligned}$$

The Jacobian J_2 does not depend on the weights w.

5.4. Derivative of the solution of the initial value problem of the DAE system

It remains to derive techniques to calculate the following expressions efficiently:

$$x_i(t), \quad \frac{\partial}{\partial(s_i, p)}x_i(t), \quad \frac{\partial}{\partial q_i}x_i(t), \quad \text{and} \quad \frac{\partial^2}{\partial q_i \partial(s_i, p)}x_i(t).$$

This requires the numerical solution of an initial value problem for the DAE system and the first and second derivatives of the solution trajectory with respect to initial values, parameters, and controls. This is the main computational effort in the evaluation of the quadratic subproblem (22) and therefore has to be done very efficiently. In our approach the variables s_i, introduced for the parametrization of the solution of the DAE system, are the initial values. For an extended approach, where the variables s_i consist of initial values and interior points, the calculation of the derivatives with respect to s_i reduces to the calculation of the derivatives with respect to initial values starting at the points $t_{i,j}$ and can therefore be implemented in nearly the same way as in the approach presented here.

In case of multiple experiments the solution of the DAE system $x_i(t)$ is computed separately for each experiment. In the following, we will omit the index i specifying the particular experiment for the sake of readability.

The state variables $x(t)$ are computed by the integrator DAESOL, a multistep-method with variable step size and order based on Backward Differentiation Formulae (BDF) which has proven to be very successful for stiff systems. Error control and step size selection take into account the true nonuniform grid (for the special techniques implemented see for example Eich [14], Bauer [3], Bauer et al. [5,6]). For derivative generation see also Bauer et al. [4].

The basic idea in the BDF-discretization is to approximate the unknown $\dot{y}_{n+1}$ in step $n+1$ by a polynomial $\delta y_{n+1} = -(1/h)\sum_{i=0}^{k} \alpha_i y_{n+1-i}$ of degree k through the last (already computed) values $y_n, \ldots, y_{n+1-k}$

$$\begin{aligned} &\sum_{i=0}^{k} \alpha_i y_{n+1-i} + h f(t_{n+1}, x_{n+1}, p, q) = 0, \\ &g(t_{n+1}, x_{n+1}, p, q) - \vartheta(t) g(t_0, x(t_0), p, q) = 0. \end{aligned} \tag{25}$$

The unknowns $x_{n+1} = (y_{n+1}, z_{n+1})$ are iteratively determined by a modified Newton method.

In order to compute the derivatives of the state variables x with respect to initial values, parameters, and controls, we apply the techniques of Internal Numerical Differentiation (IND) which have been introduced by Bock 1981 [10]. The idea is to calculate the exact derivative of the approximation of the solution of the initial value problem. Here we approximate the solution of the initial value problem by the BDF-discretization scheme (25). We consider the discretization as a mapping of the initial values, parameters, and controls to the discretized solution trajectory and differentiate this mapping by applying the chain rule. This mapping is continuous and differentiable with respect to the initial values, parameters, and controls, if we freeze the adaptive grid and all other adaptive decisions made by the integrator.

In one step (here step $n+1$) the derivative of the discretized DAE system (25) with respect to the parameters p is of the form

$$\begin{aligned} &\sum_{i=0}^{k} \alpha_i \frac{\partial y_{n+1-i}}{\partial p} + h \cdot \left(\frac{\partial f}{\partial x}(t_{n+1}, x_{n+1}, p, q) \cdot \frac{\partial x_{n+1}}{\partial p} + \frac{\partial f}{\partial p}(t_{n+1}, x_{n+1}, p, q) \right) = 0, \\ &\frac{\partial g}{\partial x}(t_{n+1}, x_{n+1}, p, q) \cdot \frac{\partial x_{n+1}}{\partial p} + \frac{\partial g}{\partial p}(t_{n+1}, x_{n+1}, p, q) - \vartheta(t) \frac{\partial g}{\partial p}(t_0, x(t_0), p, q) = 0. \end{aligned} \tag{26}$$

On the other hand if we discretize the variational DAE

$$\begin{aligned} &\dot{W}_p^y = f_x W_p + f_p, \\ &0 = g_x W_p + g_p - \vartheta(t) g_p(t_0) \end{aligned} \tag{27}$$

for the Wronskian matrix $W_p = (W_p^{y\mathrm{T}}, W_p^{z\mathrm{T}})^{\mathrm{T}} = ((\partial y/\partial p)^{\mathrm{T}}, (\partial z/\partial p)^{\mathrm{T}})^{\mathrm{T}}$ we obtain the same system of equations as (26). This means that the solution of the variational DAE (27) is the derivative of the solution of the DAE system calculated by the BDF discretization – presupposed that we use the same discretization scheme for both.

In an analogous way we set up variational DAEs for the Wronskian matrices $W_{x_0} = (\partial x/\partial x_0)$ and $W_q = (\partial x/\partial q)$ for the calculation of the derivatives with respect to initial values and controls.

For the calculation of second derivatives we again differentiate the discretized system (26) with respect to the controls q. If we use the same discretization scheme as for the nominal solution (25) and the variational DAEs of first order (27) then discretizing the variational DAE of second order

$$\begin{aligned} &\dot{W}_{p,q}^y = (f_{xx} W_q + f_{xq}) W_p + f_x W_{p,q} + f_{px} W_q + f_{pq}, \\ &0 = (g_{xx} W_q + g_{xq}) W_p + g_x W_{p,q} + g_{px} W_q + g_{pq} - \vartheta(t) g_{pq}(t_0) \end{aligned} \tag{28}$$

for the expression $W_{p,q} = \partial^2 x/(\partial q\, \partial(s, p))$ is the second derivative of the discretized nominal solution with respect to parameters and controls. Similar schemes hold for the calculation of the second derivatives $W_{x_0,q}$ with respect to initial values and controls.

The required first and second derivatives of the model functions f and g of the DAE system are calculated by automatic differentiation (for more details see the ADIFOR User's Guide of Bischof et al. [9]) which allows us to evaluate the derivatives with high accuracy. Moreover the user does not need to specify any derivatives of the model functions. We evaluate the complete expressions, e.g., $f_{xx}W_qW_p$, by directional derivatives.

In every step of the BDF discretization, we have to solve a nonlinear system of Eqs. (25) for the unknown x_{n+1} and linear systems of equations for the derivatives W_{n+1}. In our implementation we only need one evaluation and decomposition of the Jacobian of (25) to compute all first and second derivatives W_{n+1}. For the calculation of x_{n+2} by a modified Newton method we reuse the Jacobian again in the next step.

Remark. Note that the BDF polynomials represent a global error controlled approximation of the solution of the initial value problem (see [13]). Therefore, the grid of the BDF-discretization scheme may be decoupled from the grid for the evaluation of the model response and interior point constraints – and their derivatives. The computation of these quantities requires only the cheap evaluation of the BDF polynomials and no additional function calls of the (often costly) right-hand side of the DAE.

6. Sequential proceeding for optimum experimental design

As the experimental design optimization problem (11)–(14) depends on the values of the parameters which are fixed within the optimization we have to apply a strategy which both determines the parameters and designs the experiments.

In Körkel et al. [17] we suggest a sequential approach which consists of the alternate solution of experimental design problems (with fixed parameters p) and parameter estimation problems (where the experimental settings (q, u, w) are fixed). Experimental data are evaluated by carrying out the optimized experiments in laboratory according to the ones suggested by the previous optimum experimental design. Further, we need an initial guess for the parameters which can be taken, e.g., from prior experiments, similar processes, or literature. We stop the sequential procedure if the quality for the parameter estimates satisfies the user's requirements.

Algorithm. Sequential parameter estimation and experimental design

(1) Choose an initial guess $p^{(1)}$ for the parameters. Set $k := 1$.
(2) Determine $N_{\mathrm{ex}k}$ (new) experiments by solving the experimental design problem (11)–(14) with $p = p^{(k)}$. If $k > 1$ use the information from the old $N_{\mathrm{ex}1} + \cdots + N_{\mathrm{ex}k-1}$ experiments.
(3) Carry out the optimal experiments to obtain experimental data.
(4) Compute a new parameter estimate $p^{(k+1)}$ by solving a constrained least-squares parameter estimation problem using the experimental data from all $N_{\mathrm{ex}1} + \cdots + N_{\mathrm{ex}k}$ experiments. Calculate the variance–covariance matrix $C(p^{(k+1)}, s, q, u, w)$.
(5) If a user specific criterion based on the variance–covariance matrix C is fulfilled then stop. Otherwise set $k := k + 1$ and go to step 2.

6.1. Use of a priori information in experimental design

When we optimize the experimental design in step 2 of the algorithm above we want to obtain new experiments which complement the information already present in the old experiments.

The rows of the Jacobian of the underlying parameter estimation problem consist of two parts – one for the old experiments and one for the new experiments. The variance–covariance matrix ensuing from this Jacobian describes the estimation error of a parameter estimate out of the experimental data from both the old experiments and the new ones. But of course, only the control variables and weights on the measurements belonging to the new experiments are variables for the experimental design optimization, whereas the old experiments are kept fixed.

6.2. Solution of the parameter estimation problem

For the optimum experimental design approach the underlying parameter estimation problem is parametrized by only a few variables as described in Section 2.2. In the following, we present a slightly different approach for the numerical solution of the infinite-dimensional constrained least-squares problem (1)–(4). To reduce the infinite-dimensional optimization problem to a finite-dimensional one we again use a direct approach and apply the method of multiple shooting (introduced for parameter estimation by [10], see also [11] or [24]). We define an appropriate multiple shooting grid

$$t_{i,0} = \tau_{i,0} < \cdots < \tau_{i,n_{\tau_i}} = t_{i,f_i},$$

which needs not to be the same as the grid for the parametrization of the control functions for the optimum experimental design problem. At the grid points $\tau_{i,j}$, $j = 1,\ldots,n_{\tau_i}$, we introduce additional variables $v_{i,j}$ to parametrize the solution of the DAE system. On each interval $[\tau_{i,j}, \tau_{i,j+1}]$, we solve an initial value problem for the DAE system with relaxed formulation of the algebraic equations (cf. Section 3.2)

$$\dot{y}_i = f_i(t, x_i, p, q_i, u_i),$$

$$0 = g_i(t, x_i, p, q_i, u_i) - \vartheta_i(t)\, g_i(\tau_{i,j}, v_{i,j}, p, q_i, u_i)$$

with initial values $x(\tau_{i,j}) = v_{i,j} = ((v^y_{i,j})^{\mathrm{T}}, (v^z_{i,j})^{\mathrm{T}})^{\mathrm{T}}$.

Now, we can formulate a nonlinear finite-dimensional least-squares problem in the unknowns (p, v)

$$\min_{p,v} \quad \|W\,\Sigma^{-1}\,(\eta - b(p,v))\|_2^2 \tag{29}$$

$$\text{s.t.} \qquad d(p,v) = 0. \tag{30}$$

We add the consistency conditions

$$g_{ij}(v_{i,j}, p) := g_i(\tau_{i,j}, v_{i,j}, p, q_i, u_i) = 0, \quad i = 1,\ldots,N_{\mathrm{ex}},\ j = 0,\ldots,n_{\tau_i}, \tag{31}$$

to the constraints of the least-squares problem to ensure consistent intial values at the solution point. Additional continuity conditions for the differential state variables

$$h_{ij}(v_{i,j}, v_{i,j+1}, p) := y(\tau_{i,j+1}; v_{i,j}, p) - v^y_{i,j+1} = 0, \quad i = 1,\ldots,N_{\mathrm{ex}},\ j = 0,\ldots,n_{\tau_i} - 1, \tag{32}$$

guarantee continuity of the final solution also at the multiple shooting grid points. Here the control functions u_i may be represented by additional control variables q_i like in the optimum experimental design problem or by some other functions. In any case they are fixed within the parameter estimation problem.

Advantages of this approach are the free choice of the multiple shooting grid and of the initial guesses $v_{i,j}$ of the state variables $x_i(\tau_j)$ which enables the user to make use of information on the values of the states that, e.g., is given by the measurements. Moreover, the influence of bad initial guesses for the parameter estimates is reduced.

We solve the constrained parameter estimation problem (7)–(8) by a modified Gauss–Newton method: we linearize the problem (29)–(32) at a given point (v, p) and obtain a linear constrained least-squares problem in the iterates $(\Delta v, \Delta p)$. If we order the parts of the objective and the constraints experiment-wise we yield a linearized least-squares problem with a Jacobian of the following structure:

$$J_{\mathrm{LS}} = \begin{pmatrix} M_1^v & & 0 & M_1^p \\ & \ddots & & \vdots \\ 0 & & M_{N_{\mathrm{ex}}}^v & M_{N_{\mathrm{ex}}}^p \end{pmatrix}$$

with

$$M_i^v = \begin{pmatrix} R_i^{v_{i,0}} & \cdots & R_i^{v_{i,n_{\tau_i}}} \\ D_i^{v_{i,0}} & \cdots & D_i^{v_{i,n_{\tau_i}}} \\ G_{i,0}^{v_{i,0}} & & 0 \\ & \ddots & \\ 0 & & G_{i,n_{\tau_i}}^{v_{i,n_{\tau_i}}} \\ H_{i,0}^{v_{i,0}} & H_{i,0}^{v_{i,1}} & 0 \\ & \ddots \quad \ddots & \\ 0 & H_{i,n_{\tau_i}-1}^{v_{i,n_{\tau_i}-1}} & H_{i,n_{\tau_i}-1}^{v_{i,n_{\tau_i}}} \end{pmatrix}, \quad M_i^p = \begin{pmatrix} R_i^p \\ D_i^p \\ G_{i,0}^p \\ \vdots \\ G_{i,n_{\tau_i}}^p \\ H_{i,0}^p \\ \vdots \\ H_{i,n_{\tau_i}-1}^p \end{pmatrix}$$

and right-hand side

$$\mathrm{rhs} = \begin{pmatrix} \mathrm{rhs}_1 \\ \vdots \\ \mathrm{rhs}_{N_{\mathrm{ex}}} \end{pmatrix}, \quad \mathrm{rhs}_i = - \begin{pmatrix} r_i \\ d_i \\ g_{i,0} \\ \vdots \\ g_{i,n_{\tau_i}} \\ h_{i,0} \\ \vdots \\ h_{i,n_{\tau_i}-1} \end{pmatrix}.$$

Here we use the following notations:

$$R_i^\kappa = \frac{\partial r_i}{\partial \kappa}, \quad D_i^\kappa = \frac{\partial d_i}{\partial \kappa}, \quad G_{i,j}^\kappa = \frac{\partial g_{i,j}}{\partial \kappa}, \quad H_{i,j}^\kappa = \frac{\partial h_{i,j}}{\partial \kappa}, \quad \kappa \in \{v_{i,j}, p\},$$

$$i = 1, \ldots, N_{\mathrm{ex}}, \ \ j = 0, \ldots, n_{\tau_i} \text{ resp. } j = 0, \ldots, n_{\tau_i} - 1,$$

and

$$H_{i,j}^{v_{i,j+1}} = (-I_{n_{y_i}}, 0_{n_{z_i}}), \quad i = 1, \ldots, N_{\mathrm{ex}}, \ j = 0, \ldots, n_{\tau_i} - 1.$$

For the calculation of $H_{i,j}^\kappa$ we solve initial value problems for the DAE system and the first derivatives of the solution trajectory with respect to $\kappa \in \{v_{i,j}, p\}$ on each interval $[\tau_{i,j}, \tau_{i,j+1}]$. These are computed experiment-wise by DAESOL using the techniques of internal numerical differentiation as described in Section 5.4.

Solving the structured linearized constrained least-squares problem with Jacobian J_{LS} and right-hand side rhs we exploit both the structure of the multiple experiments and of the multiple shooting approach in each experiment (see [24,25]). First, we eliminate the local variables $v_{i,1}, \ldots, v_{i,n_{\tau_i}}$ in each experiment and solve the structured condensed least-squares problem in the reduced set of variables $(v_{i,0}, p)$ with reduced Jacobian

$$J_{\mathrm{RLS}} = \begin{pmatrix} E_1^v & & 0 & E_1^p \\ & \ddots & & \vdots \\ 0 & & E_{N_{\mathrm{ex}}}^v & E_{N_{\mathrm{ex}}}^p \end{pmatrix}$$

with

$$E_i^v = \begin{pmatrix} \tilde{R}_i^{v_{i,0}} \\ \tilde{D}_i^{v_{i,0}} \\ \tilde{G}_{i,0}^{v_{i,0}} \end{pmatrix}, \quad E_i^p = \begin{pmatrix} \tilde{R}_i^p \\ \tilde{D}_i^p \\ \tilde{G}_{i,0}^p \end{pmatrix}, \quad i = 1, \ldots, N_{\mathrm{ex}}$$

and new right-hand side

$$\mathrm{rhs}_R = \begin{pmatrix} \widetilde{\mathrm{rhs}}_1 \\ \vdots \\ \widetilde{\mathrm{rhs}}_{N_{\mathrm{ex}}} \end{pmatrix}, \quad \widetilde{\mathrm{rhs}}_i = - \begin{pmatrix} \tilde{r}_i \\ \tilde{d}_i \\ \tilde{g}_{i,0} \end{pmatrix}, \quad i = 1, \ldots, N_{\mathrm{ex}}.$$

In each experiment we use a LU decomposition for the matrices corresponding to the local variables $v_{i,0}$. The submatrices corresponding to the parameters p are also transformed as well as the right-hand side. The remaining system is summarized and ordered by equality and least-squares conditions and solved by a generalized QR decomposition. We obtain a new iterate for the parameters p and then calculate successively the iterates for the local variables $v_{i,0}, v_{i,1}, \ldots, v_{i,n_{\tau_i}}$ in a backward recursion.

7. Application of the methods to the reaction of urethane

We applied our methods and software tools to the reaction of urethane, provided by Kud from the BASF AG. Even though this example involves only few reactions, it already shows typical properties

and difficulties of nonlinear experimental design, but it also demonstrates the enormous potential of this method.

Further applications are presented in [8].

7.1. The reaction of urethane

The reaction of urethane is a simultaneous and consecutive reaction including a reversible reaction:

$$A + B \to C,$$

$$A + C \rightleftharpoons D,$$

$$3A \to E$$

with A we denote phenylisocyanate, B is butanol, C is urethane, D is allophanate, and E is isocyanurate. Solvent is dimethylsulfoxide.

The experiments for this reaction are carried out in a semi-batch reactor with two feed vessels, one for phenylisocyanate and one for butanol. In the beginning, the reactor contains phenylisocyanate and butanol. The temperature inside the reactor can be controlled.

We describe this process by a nonlinear DAE model:

$$\dot{n}_3 = V \cdot (r_1 - r_2 + r_3),$$

$$\dot{n}_4 = V \cdot (r_2 - r_3),$$

$$\dot{n}_5 = V \cdot r_4,$$

$$0 = n_1 + n_3 + 2n_4 + 3n_5 - n_{a1} - n_{1ea},$$

$$0 = n_2 + n_3 + n_4 - n_{a2} - n_{2eb},$$

$$0 = n_6 - n_{a6} - n_{6ea} - n_{6eb},$$

$$n_3(t_0) = n_4(t_0) = n_5(t_0) = 0 \text{ mol}, \quad t_0 = 0\,\text{h}, \quad t_\text{f} = 80\,\text{h}$$

with

$$V = \sum_{i=1}^{6} n_i \frac{M_i}{\rho_i}, \qquad k_i = k_i = k_{\text{ref}\,i} \exp\left(-\frac{E_{ai}}{R}\left(\frac{1}{T} - \frac{1}{T_{\text{ref}\,i}}\right)\right),$$

$$r_1 = k_1 \cdot \frac{n_1}{V} \cdot \frac{n_2}{V}, \quad i = 1, 2, 4,$$

$$r_2 = k_2 \cdot \frac{n_1}{V} \cdot \frac{n_3}{V}, \quad k_3 = \frac{k_2}{k_c},$$

$$r_3 = k_3 \cdot \frac{n_4}{V}, \qquad k_c = k_{c2} \exp\left(-\frac{dh_2}{R} \cdot \left(\frac{1}{T} - \frac{1}{T_{g2}}\right)\right)$$

$$r_4 = k_4 \cdot \left(\frac{n_1}{V}\right)^2.$$

The molar numbers $n_1,\ldots,n_5$ of the species A to E and n_6 of the solvent are the state variables of the DAE system.

In this model eight parameters are unknown: the steric factors $k_{\mathrm{ref}\,i}$, $i=1,2,4$, the activation energies E_{ai}, $i=1,2,4$, as well as the equilibrium constant k_{c2} (for the reference temperature T_{g2}) and the reaction enthalpy dh_2 of the reversible reaction.

The two feeds are modelled by two monotonously increasing control functions

$$\mathrm{feed}_a,\ \mathrm{feed}_b : [t_0; t_\mathrm{f}] \to [0; 1]$$

describing the profiles of the accumulated feeds. Multiplied with the initial molar numbers within the feed vessels, we get the feed molar numbers:

$$n_{1ea} = n_{a1ea} \cdot \mathrm{feed}_a, \quad n_{2eb} = n_{a2eb} \cdot \mathrm{feed}_b,$$

$$n_{6ea} = n_{a6ea} \cdot \mathrm{feed}_a, \quad n_{6eb} = n_{a6eb} \cdot \mathrm{feed}_b.$$

The third control function is the temperature profile

$$T : [t_0; t_\mathrm{f}] \to [293.16\,\mathrm{K}; 473.16\,\mathrm{K}].$$

Each experiment lasts 80 h. The beginning is Monday, 8 pm, the end Thursday, 4 pm. During the nights, the feed rates and the heating/cooling rate have to be zero due to safety rules.

Further control variables for experimental design are the mole ratios $\mathrm{MV}_1 \in [0.1; 10]$, $\mathrm{MV}_2 \in [0; 1000]$, and $\mathrm{MV}_3 \in [0; 10]$, the parts of active ingredients $g_a \in [0; 0.8]$, $g_{aea} \in [0; 0.9]$, and $g_{aeb} \in [0; 1]$, and the initial volume $V_a \in [0\,\mathrm{m}^3; 0.00075\,\mathrm{m}^3]$ of the species in the reactor.

These quantities are connected to the initial molar numbers as follows:

$$\mathrm{MV}_1 = \frac{n_{a2} + n_{a2eb}}{n_{a1} + n_{a1ea}}, \quad g_a = \frac{n_{a1} \cdot M_1 + n_{a2} \cdot M_2}{n_{a1} \cdot M_1 + n_{a2} \cdot M_2 + n_{a6} \cdot M_6},$$

$$\mathrm{MV}_2 = \frac{n_{a1ea}}{n_{a1}}, \quad g_{aea} = \frac{n_{a1ea} \cdot M_1}{n_{a1ea} \cdot M_1 + n_{a6ea} \cdot M_6},$$

$$\mathrm{MV}_3 = \frac{n_{a2eb}}{n_{a1}}, \quad g_{aeb} = \frac{n_{a2eb} \cdot M_2}{n_{a2eb} \cdot M_2 + n_{a6eb} \cdot M_6},$$

$$V_\mathrm{a} = \frac{n_{a1}}{\rho_1} \cdot M_1 + \frac{n_{a2}}{\rho_2} \cdot M_2 + \frac{n_{a6}}{\rho_6} \cdot M_6.$$

All other quantities in the model are constants: the molar masses M_i, the densities ρ_i, the reference temperatures $T_{\mathrm{ref}\,i}$ and T_{g2}, and the gas constant R, see Table 1.

Three measurement methods are available:

- titration, measuring mass percent of phenylisocyanate with a standard deviation of the measurement error of 0.5,
- HPLC1, giving mass percent of urethane and allophanate with standard deviations 0.5 resp. 0.005, and
- HPLC2, for mass percent of isocyanurate with standard deviation 0.0005.

In each experiment, 16 measurements can be selected out of 30 possible ones. We parametrize the time depending control functions using piecewise linear and continuous polynomials. Altogether we have 90 experimental design variables for each experiment: 7 control variables, 7 initial molar

Table 1
Constants in the model for the reaction of urethane

Molar masses	Densities	Reference temperatures
$M_1 = 0.11911$ kg/mol	$\rho_1 = 1095.0$ kg/m^3	$T_{\mathrm{ref}\,1} = 363.16$ K
$M_2 = 0.07412$ kg/mol	$\rho_2 = 809.0$ kg/m^3	$T_{\mathrm{ref}\,2} = 363.16$ K
$M_3 = 0.19323$ kg/mol	$\rho_3 = 1415.0$ kg/m^3	$T_{\mathrm{ref}\,4} = 363.16$ K
$M_4 = 0.31234$ kg/mol	$\rho_4 = 1528.0$ kg/m^3	$T_{g2} = 363.16$ K
$M_5 = 0.35733$ kg/mol	$\rho_5 = 1451.0$ kg/m^3	Gas constant
$M_6 = 0.07806$ kg/mol	$\rho_6 = 1101.0$ kg/m^3	$R = 8.314$ J/(K mol)

Table 2
Sequential parameter estimation for the reaction of urethane. The table shows the initial guesses for the parameters, the estimates after one and two optimized experiments, and the parameter estimate from the intuitive experimental design after 15 experiments. The estimated parameters are shown with the corresponding standard deviations. The parameters are scaled such that all initial guesses are equal to 1

Parameter	Initial guess	First estimate	Second estimate	Intuitive result
$k_{\mathrm{ref}\,1}$	1	$2.51 \pm 3 \cdot 10^{-2}$	$2.504 \pm 6 \cdot 10^{-3}$	$2.50 \pm 2 \cdot 10^{-2}$
E_{a1}	1	$0.835 \pm 4 \cdot 10^{-3}$	$0.836 \pm 1 \cdot 10^{-3}$	$0.835 \pm 7 \cdot 10^{-3}$
$k_{\mathrm{ref}\,2}$	1	$91.2 \pm 1 \cdot 10^{-1}$	$91.25 \pm 2 \cdot 10^{-2}$	$91.3 \pm 7 \cdot 10^{-1}$
E_{a2}	1	$0.8355 \pm 2 \cdot 10^{-4}$	$0.83537 \pm 8 \cdot 10^{-5}$	$0.834 \pm 2 \cdot 10^{-3}$
$k_{\mathrm{ref}\,4}$	1	$58.0 \pm 1 \cdot 10^{-1}$	$58.004 \pm 6 \cdot 10^{-3}$	$57.990 \pm 9 \cdot 10^{-3}$
E_{a4}	1	$0.658 \pm 1 \cdot 10^{-3}$	$0.6574 \pm 2 \cdot 10^{-4}$	$0.65725 \pm 7 \cdot 10^{-5}$
dh_2	1	$1.09 \pm 2 \cdot 10^{-2}$	$1.082 \pm 6 \cdot 10^{-3}$	$0.9 \pm 3 \cdot 10^{-1}$
k_{c2}	1	$1.30 \pm 3 \cdot 10^{-2}$	$1.29 \pm 1 \cdot 10^{-2}$	$1.1 \pm 3 \cdot 10^{-1}$
A-criterion		$3.8 \cdot 10^{-3}$	$5.5 \cdot 10^{-5}$	$7.2 \cdot 10^{-2}$

numbers, 30 weights on the measurements, and 46 variables due to the parametrization of the control functions.

7.2. Computational results

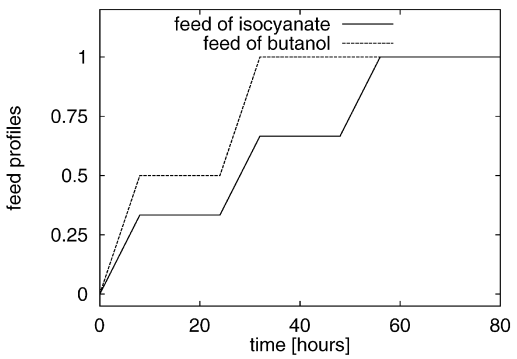
As initial guesses for the parameters values from literature were chosen. In the first iteration, one experiment was computed by minimizing the A-criterion. The parameter estimation from the corresponding experimental data resulted in parameters with standard deviations of at most 2.5% (see Table 2). We decided to design a second A-optimal experiment taking into account the a priori information from the first experiment to improve this estimation. After the second parameter estimation, the standard deviations were smaller than 0.8%.

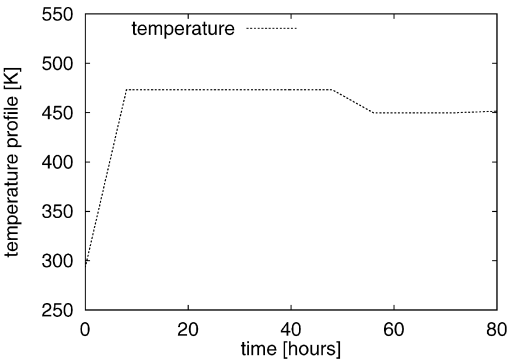
The experimental data were generated by simulations. A set of parameters was assigned to be the true one. The simulation results of the measurements were perturbed by normally distributed errors with the variances of the measurement methods.

7.3. Initial values for the optimization

We started the optimization with the following initial values for the experimental design variables: $MV_1 = 1.0$, $MV_2 = 0.3$, $MV_3 = 0.3$, $g_a = 0.75$, $g_{aea} = 0.5$, $g_{aeb} = 0.4$, $V_a = 2.75 \cdot 10^{-5}\,m^3$.

For the control functions we used the following initial profiles:





7.4. Experimental set-up of the two computed experiments

First experiment

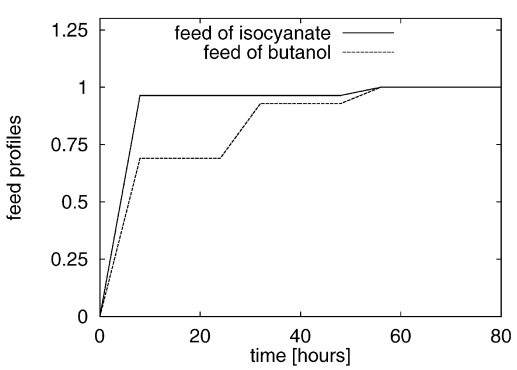
Control variables:

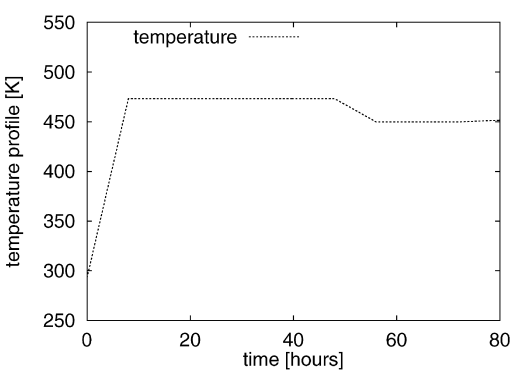
MV_1	MV_2	MV_3	g_a	g_{aea}	g_{aeb}	V_a
0.637635	15.9264	10	0.8	0.9	1	$2.09045 \cdot 10^{-5}\,m^3$

Measurement points for the different methods:

- titration: 0.5, 1,
- HPLC 1: 0.5, 1, 4, 8, 24, 32, 48, 56, 72, 80,
- HPLC 2: 4, 8, 32, 48.

Control functions:





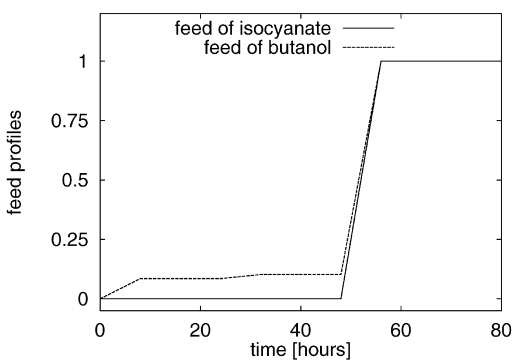
Second experiment:
Control variables:

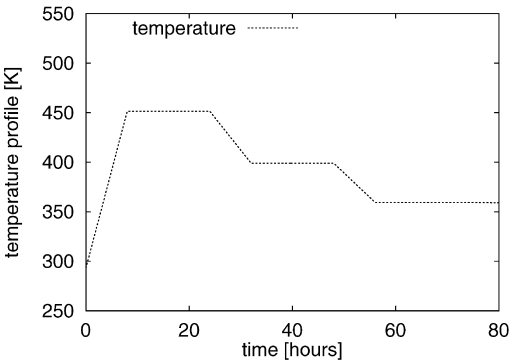
MV_1	MV_2	MV_3	g_a	g_{aea}	g_{aeb}	V_a
0.370994	25.9546	10	0.8	0.9	1	$9.00762 \cdot 10^{-6}\,\mathrm{m}^3$,

Measurement points for the different methods:

- titration: none,
- HPLC 1: 4, 8, 24, 32, 48, 56, 72, 80,
- HPLC 2: 4, 8, 24, 32, 48, 56, 72, 80.

Control functions:

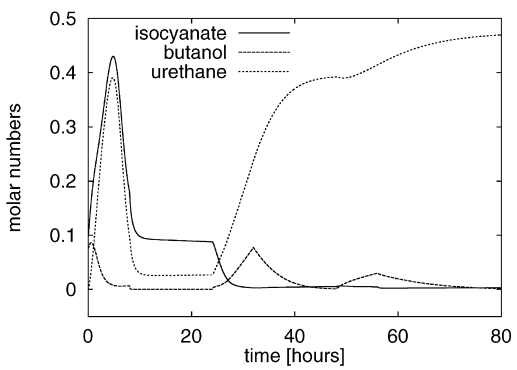


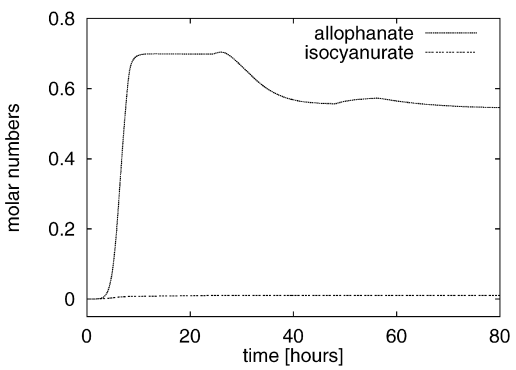


7.5. Trajectories

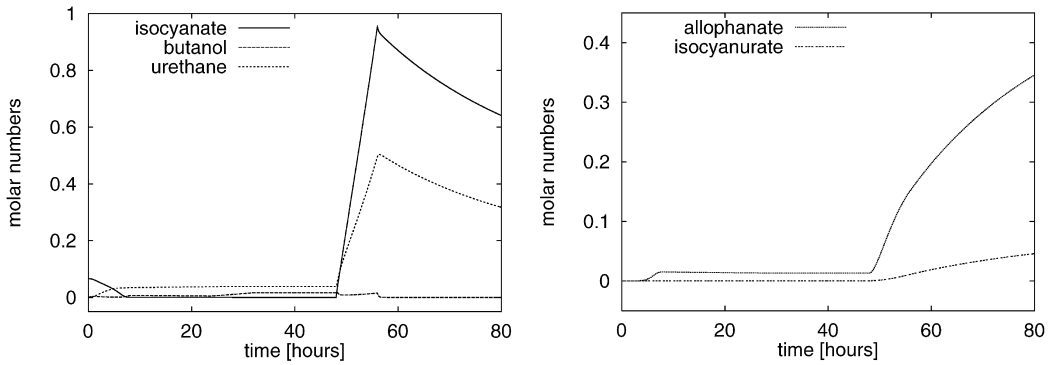
The following plots show the trajectories of the five species, evaluated for the second and final estimate of the parameters:

First experiment:





Second experiment:



We used an SGI O2 workstation with one mips R5000 processor and the IRIX 6.3 operation system. The code was compiled by the SGI C++ and FORTRAN compiler using the -O2 optimization level. The SQP library SNOPT of Gill et al. [15] was applied. The optimization runs took up to 250 SQP iterations and lasted up to 1 h and 15 min.

7.6. For comparison: intuitive experimental design by an experienced experimentor

Wörz, a chemist from BASF, was asked to design experiments and to estimate the parameters under the same conditions as for optimum experimental design, see [7]. Using his experience from chemical laboratory's everyday life he planned his experiments according to the following principles:

- separation of the several reactions,
- isothermal experiments,
- batch reactor.

He designed

- four experiments for the formation of urethane at three different temperatures using the titration measurement method,
- six experiments for the formation of allophanate at five different temperatures using the HPLC1 measurement method,
- and five experiments for the formation of isocyanurate at three different temperatures using the HPLC2 measurement method.

Altogether he employed 15 experiments with 90 measurements. After 15 experiments he obtained the parameter estimate shown in Table 2. The standard deviations still are up to 30% of the values of the parameters.

The comparison between intuitive and optimum experimental design shows the enormous potential arising from the application of numerical methods regarding the savings of experimental costs and experimental effort.

Acknowledgements

This work was financially supported by the German Ministry for Research and Technology (BMBF) within the project "Optimum Experimental Design for Nonlinear Processes" (FKZ 03 D 0043). The authors want to thank the project partners Aventis, BASF and FH Frankfurt am Main for the fruitful cooperation with the IWR. Special thanks are dedicated to Dr. A. Kud and Dr. O. Wörz from BASF for their contribution concerning the reaction of urethane.

References

[1] A.C. Atkinson, A.N. Donev, Optimum Experimental Designs. Oxford Statistical Science Series, Oxford University Press, Oxford, 1992.
[2] M. Baltes, R. Schneider, C. Sturm, M. Reuss, Optimal experimental design for parameter estimation in unstructured growth models, Biotechnol. Prog. 10 (1994) 480–488.
[3] I. Bauer, Numerische Behandlung differentiell-algebraischer Gleichungen mit Anwendungen in der Chemie, Diploma Thesis, Universität Augsburg, 1994.
[4] I. Bauer, H.G. Bock, S. Körkel, J.P. Schlöder, Numerical methods for initial value problems and derivative generation for DAE models with application to optimum experimental design of chemical processes, in: F. Keil, W. Mackens, H. Voss, J. Werther (Eds.), Scientific Computing in Chemical Engineering II, Vol. 2: Simulation, Image Processing, Optimization, and Control, Springer, Berlin, 1999, pp. 282–289.
[5] I. Bauer, H.G. Bock, J.P. Schlöder, DAESOL – a BDF-code for the numerical solution of differential algebraic equations, Technical Report, IWR, SFB 359, Universität Heidelberg, 1999.
[6] I. Bauer, F. Finocchi, W.J. Duschl, H.-P. Gail, J.P. Schlöder, Simulation of chemical reactions and dust destruction in protoplanetary accretion disks, Astronomy and Astrophys. 317 (1997) 273–289.
[7] I. Bauer, M. Heilig, S.Körkel, A. Kud, A. Mayer, O. Wörz, Versuchsplanung am Beispiel einer Phosphin- und Urethanreaktion, in: Optimale Versuchsplanung für nichtlineare Prozesse, DECHEMA e.V., Frankfurt, 1998.
[8] I. Bauer, S. Körkel, H.G. Bock, J.P. Schlöder, Optimale Versuchsplanung für dynamische Systeme aus der chemischen Reaktionskinetik, in: Optimale Versuchsplanung für nichtlineare Prozesse, DECHEMA e.V., Frankfurt, 1998.
[9] Ch.H. Bischof, A. Carle, P.M. Khademi, A. Mauer, P. Hovland, ADIFOR 2.0 User's Guide, Technical memorandum No. 192, Mathematics and Computer Science Division, 1998.
[10] H.G. Bock, Numerical treatment of inverse problems in chemical reaction kinetics, in: K.H. Ebert, P. Deuflhard, W. Jäger (Eds.), Modelling of Chemical Reaction Systems, Springer Series in Chemical Physics, Vol. 18, Springer, Heidelberg, 1981, pp. 102–125.
[11] H.G. Bock, Randwertproblemmethoden zur Parameteridentifizierung in Systemen nichtlinearer Differentialgleichungen, Bonner Math. Schriften 183 (1987).
[12] H.G. Bock, E. Eich, J.P. Schlöder, Numerical solution of constrained least squares boundary value problems in differential-algebraic equations, in: K. Strehmel (Ed.), Numerical Treatment of Differential and Integral Equations, BG Teubner, Leipzig, 1988.
[13] H.G. Bock, J.P. Schlöder, Numerical solution of retarded differential equations with state dependent time lags, Z. Angew. Math. Mech. 61 (1981) T269–T271.
[14] E. Eich, Numerische Behandlung semi-expliziter differentiell-algebraischer Gleichungssysteme vom Index 1 mit BDF-verfahren, Diploma Thesis, Universität Bonn, 1987.
[15] Ph.E. Gill, W. Murray, M.A. Saunders, User's guide for SNOPT 5.3: a fortran package for large-scale nonlinear programming, Technical Report, NA 97-5, Department of Mathematics, University of California, San Diego, 1998.
[16] K.-D. Hilf, Optimale Versuchsplanung zur Dynamischen Roboterkalbibrierung, Fortschrittberichte, Vol. 8, VDI, Düsseldorf, 1996.
[17] S. Körkel, I. Bauer, H.G. Bock, J.P. Schlöder, A sequential approach for nonlinear optimum eperimental design in DAE systems, in: F. Keil, W. Mackens, H. Voss, J. Werther (Eds.), Scientific Computing in Chemical Engineering II, Simulation, Image Processing, Optimization, and Control, Vol. 2, Springer, Berlin, 1999, pp. 338–345.

[18] Th.W. Lohmann, Ein numerisches Verfahren zur Berechnung optimaler Versuchspläne für beschränkte Parameteridentifizierungsprobleme, Reihe Informatik, Verlag Shaker, Aachen, 1993.
[19] Th.W. Lohmann, H.G. Bock, J.P. Schlöder, Numerical methods for parameter estimation and optimal experiment design in chemical reaction systems, Ind. Eng. Chem. Res. 31 (1992) 54–57.
[20] J.A. Nelder, R.A. Mead, A simplex method for function minimisation, Comput. J. 7 (1964) 308–313.
[21] A. Pázman, Foundations of Optimum Experimental Design, Reidel, Dordrecht, 1986.
[22] F. Pukelsheim, Optimal Design of Experiments, Wiley Series in Probability and Mathematical Statistics, Wiley, New York, 1993.
[23] P.E. Rudolph, G. Herrendörfer, Optimal experimental design and accuracy of parameter estimation for nonlinear regression models used in long-term selection, Biometrical J. 37 (2) (1995) 183–190.
[24] J.P. Schlöder, Numerische Methoden zur Behandlung hochdimensionaler Aufgaben der Parameteridentifizierung, Bonner Math. Schriften 187 (1988).
[25] M. von Schwerin, Numerische Methoden zur Schätzung von Reaktionsgeschwindigkeiten bei der katalytischen Methankonversion und Optimierung von Essigsäure- und Methanprozessen, Ph.D. Thesis, Universität Heidelberg, 1998.