Aachen Institute for Advanced Study in Computational Engineering Science

Preprint: AICES-2009-7

11/March/2009

Incremental Global Parameter Estimation in Dynamical Systems

C. Michalik,B. Chachuat and W. Marquardt

RWTH AACHEN UNIVERSITY

AICES

Financial support from the Deutsche Forschungsgemeinschaft (German Research Association) through grant GSC 111 is gratefully acknowledged.

List of AICES technical reports: http://www.aices.rwth-aachen.de/preprints

Incremental Global Parameter Estimation in Dynamical Systems

Claas Michalik,[†] Benoit Chachuat,[‡,§] and Wolfgang Marquardt[*,¶]

AVT - Process Systems Engineering, RWTH Aachen University, Turmstraße 46, D-52064 Aachen, Germany, Laboratoire d'Automatique, École Polytechnique Fédérale de Lausanne (EPFL), Station 9, CH-1015 Lausanne, Switzerland., and AVT - Process Systems Engineering, RWTH Aachen University, Turmstraße 46, D-52064 Aachen, Germany

E-mail: Wolfgang.Marquardt@avt.rwth-aachen.de

Abstract

Estimating the parameters of a dynamical system based on measurements is an important task in industrial and scientific practice. Since a model's quality is directly linked to its parameter values, obtaining globally rather than locally optimal values is especially important in this context. In practice, however, local methods are used almost exclusively. This is mainly due to the high computational cost of global dynamic parameter estimation, which limits its application to relatively small problems comprising no more than a few equations and parameters. In addition, there is still a lack of software packages that allow global parameter estimation in dynamical systems without expert knowledge. Therefore, we propose an efficient computational method for obtaining globally optimal parameter estimates of dynamical systems using

[†]AVT - Process Systems Engineering, RWTH Aachen University, Turmstraße 46, D-52064 Aachen, Germany

[‡]Laboratoire d'Automatique, École Polytechnique Fédérale de Lausanne (EPFL), Station 9, CH-1015 Lausanne, Switzerland.

[¶]AVT - Process Systems Engineering, RWTH Aachen University, Turmstraße 46, D-52064 Aachen, Germany

[§]Current address: Department of Chemical Engineering, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4L7, Canada.

well-established, user-friendly software packages. The method is based on the so-called incremental identification procedure, in combination with deterministic global optimization tools for nonlinear programs.

Introduction

Mathematical models giving accurate predictions of physical phenomena are essential tools in engineering and scientific fields. In chemical engineering, such models form the basis for the design, optimization and control of process systems.[1] However, these models often contain adjustable parameters (semi-empirical models), the values of which are to be determined from available experimental data to yield accurate predictions. In many cases, the task of parameter estimation is rendered more complex by using models that are comprised of mixed sets of nonlinear differential and algebraic equations (DAEs).

Traditionally, parameter estimation is performed following a maximum likelihood approach, where the objective typically is to minimize the weighted squared error between a set of measured data and corresponding model predictions. Early attempts to solve such problems in a systematic, computationally tractable manner using local search methods can be traced back to the 1960s.[2,3] Both, the *sequential* and *simultaneous* approaches of dynamic optimization have been widely studied in this context. In the sequential approach, an integration routine is used to determine the state values for a given set of model parameters, which in turn allows for the evaluation of the objective function and its derivatives in a master nonlinear program (NLP). One such implementation has been reported by Kim et al.[4] and this is also the approach implemented in process simulation software such as gPROMS[5] and JACOBIAN.[6] Multiple shooting methods are a robust alternative within the sequential approach and have successfully been applied by various authors (see for instance the works by Bock,[7] Lohman et al.[8] or Peifer and Timmer[9]). In the simultaneous approach, on the other hand, the dynamic system is converted into a set of algebraic equations, which are solved along with the model parameter values in a large-scale NLP. This approach has for instance

been employed by Van Den Bosch and Hellinckx[10] and Tjoa and Biegler.[11]

A common limitation of the aforementioned approaches is that they can, at best, achieve convergence to a local minimum since they rely on local search methods. This deficiency to converge to a global minimum may have dramatic consequences. In the context of reaction kinetics, for instance, getting a bad fit may lead to the erroneous conclusion that a proposed kinetic model yields an incorrect description of the chemistry based only on the knowledge of a local minimum; [12] in other words, only a global minimum allows one to conclusively determine whether or not a proposed model is inconsistent with a set of experimental data.

To increase the likelihood of finding a global minimizer, both *stochastic* and *deterministic global optimization* have been developed. Stochastic search methods rely on probabilistic approaches.[13,14] They are usually quite simple to implement and their efficiency has been demonstrated on many applications. Yet, they cannot guarantee locating a global solution in a finite number of iterations; see, e.g.,[15] for a recent comparison of various global optimization methods in parameter estimation of biochemical pathways. Deterministic methods, on the other hand, can provide a level of assurance that the global optimum will be located in finite time.[16] These sound theoretical convergence properties have stimulated the development of deterministic global optimization for problems with embedded differential equations.[12,17–21] Briefly, a standard branch-and-bound (B&B) algorithm is employed to converge to the global solution by systematically eliminating portions of the feasible space; this is achieved by solving a sequence of upper- and lower-bounding problems on refined sub-partitions; the proposed approaches mostly differ in the derivation and solution of the lower-bounding problems. Although tremendous progress has been achieved in recent years, global dynamic optimization methods are currently limited to problems containing no more than a few decision variables; from a computational viewpoint, B&B methods indeed exhibit worst-case complexity that is exponential in the number of decision variables. Moreover, still no general method has been proposed to rigorously address problems with DAEs embedded.

All of the foregoing parameter estimation approaches (be they local or global) fall into the

scope of *simultaneous* identification (also called *integral* method), in the sense that *all* the adjustable model parameters are estimated simultaneously. Note that these approaches give statistically optimal parameter estimates in a maximum likelihood framework.[22] Recently, the so-called *incremental* approach for model identification has been introduced by Marquardt and coworkers.[23,24] The key idea therein is to follow the steps that are usually taken in the development of model structures, thus yielding a sequence of algebraic parameter estimation problems, which are simpler than the original differential problem; note that there is a relation to the well-known *differential* method of parameter estimation in kinetic models.[25] Although incremental identification approaches do not share the same theoretical properties as simultaneous methods, namely they are neither unbiased nor consistent,[22] their major advantage lies in the computational ease. Moreover, comparisons have indicated that good results can be obtained provided that sufficient care is taken during the estimation.[23]

This paper proposes a novel methodology for parameter estimation in dynamical systems. Building upon the incremental identification method, the original problem is split into separate steps, each of which yields an algebraic estimation problem. The contribution of this paper is twofold. First, the incremental approach is extended to encompass general parameter identification problems in DAE systems, i.e., not necessarily reaction kinetic models. Second, deterministic global optimization is used in a systematic way for solving the (potentially nonconvex) algebraic estimation problems, hence the name *Incremental Global Parameter Estimation* (IGPE). Note that efficient global optimization software, such as BARON,[26] has indeed become available during the last decade for the solution of algebraic problems. The main advantage of IGPE, as compared to simultaneous global optimization, lies in a significant decrease of the overall computational time. The major drawbacks on the other hand are, that (i) the method relies on the availability of sufficiently rich measurement data and that (ii) the global solution in the incremental approach will generally be different from the global solution in the simultaneous approach in the presence of measurement noise. However, the incremental solution can be used as a starting point in a local simultaneous estimation problem. In particular, our experience is that this heuristic procedure

typically yields the global solution to the original problem.

The remainder of the paper is organized as follows. The general problem formulation is given in Section 2. The proposed IGPE approach is presented in Section 3. A case study illustrating the various steps of the IGPE approach is then detailed in Section 4. The application of IGPE to test problems from the literature is reported in Section 5. Finally, conclusions and future work are discussed in Section 6.

Problem Definition

In the following, we consider a class of dynamical systems described by means of DAEs,

$$\mathbf{F}(\dot{\mathbf{x}}(t),\mathbf{x}(t),\mathbf{y}(t),\mathbf{u}(t),\mathbf{p}) = \mathbf{0}, \tag{1}$$

where $\mathbf{x}(t) \in \mathbb{R}^{n_x}$ is the vector of differential state variables at time $t \in [t_0, t_f]$, $\mathbf{y}(t) \in \mathbb{R}^{n_y}$ is the vector of algebraic state variables, $\mathbf{u}(t) \in \mathbb{R}^{n_u}$ is the vector of inputs, and $\mathbf{p} \in P$ is the vector of time-invariant parameters to be estimated, with $P \subset \mathbb{R}^{n_p}$ a compact set.

Throughout this paper it is assumed that the matrix $[\mathbf{F}_{\dot{\mathbf{x}}}\, \mathbf{F}_{\mathbf{y}}]$ has full rank; that is the index of the DAEs is less than or equal to one. Consistent initial conditions $\mathbf{x}(t_0) = \mathbf{x}_0$, $\mathbf{y}(t_0) = \mathbf{y}_0$ are assumed to be available. It is also assumed that the inputs $\mathbf{u}(t)$ are given.

Let $I_m \subseteq \{1,\ldots,n_x\}$ denote the indices of the measured differential state variables[1] and let $t_1,\ldots,t_l$ such that $t_0 \leq t_1 < \ldots < t_\ell = t_f$ denote the measurement times. Let $\tilde{\mathbf{x}}_{m,i} \in \mathbb{R}^{n_m}$, $i = 1,\ldots,\ell$, be the vector of measurements at times t_i, $i = 1,\ldots,\ell$. The measurements and corresponding model predictions are arranged in the column vectors

$$\tilde{\mathbf{X}}_m := \begin{pmatrix} \tilde{\mathbf{x}}_{m,1} \\ \vdots \\ \tilde{\mathbf{x}}_{m,\ell} \end{pmatrix} \in \mathbb{R}^{\ell \cdot n_m}, \quad \text{and} \quad \mathbf{X}_m(\mathbf{p}) := \begin{pmatrix} \mathbf{x}_m(t_1) \\ \vdots \\ \mathbf{x}_m(t_\ell) \end{pmatrix} \in \mathbb{R}^{\ell \cdot n_m}. \tag{2}$$

[1] For simplicity of presentation we assume that only differential state variables are measured.

We consider global weighted least-squares parameter estimation problems of the form

$$\begin{aligned} \min_{\mathbf{p}\in P} \quad & (\tilde{\mathbf{X}}_m - \mathbf{X}_m(\mathbf{p}))^T \mathbf{W} (\tilde{\mathbf{X}}_m - \mathbf{X}_m(\mathbf{p})) \\ \text{s.t.} \quad & \text{Eq. (1)} \\ & \mathbf{p}_l \leq \mathbf{p} \leq \mathbf{p}_u, \end{aligned} \tag{3}$$

where $\mathbf{W}$ is a properly chosen weighting matrix[2] and $\mathbf{p}_l$ and $\mathbf{p}_u$ are the vectors of lower and upper parameter bounds, respectively.

The IGPE Method

The IGPE method is based on ideas presented by Brendel et al.[24] for the identification of chemical reaction systems. Brendel et al.[24] split the identification task into three steps. In the first step, the available measurements are used to obtain estimates of the state derivatives, which in turn provide the so-called reaction fluxes. In the second step, the reaction rates are estimated based on both the structure of the reaction network and the estimated reaction fluxes. In the last step, the model parameters are determined for each reaction taken separately. In the framework of global optimization this approach has the following appealing features:

1. The dynamic estimation problem is reduced to a number of algebraic estimation problems by using the time derivatives of the states estimated from the measurements.

2. Because the resulting algebraic problems tend to be decoupled from one another, smaller parameter subsets are to be estimated in each subproblem as compared to the simultaneous estimation of all model parameters.

The first feature is thrilling since algebraic parameter estimation problems can be solved to global optimality at a significantly lower computational cost than their dynamic counterparts.[27] In particular, well-established, robust and fast global optimizers for algebraic models are now available.

[2]We use the identity matrix as the weighting matrix throughout this paper.

[26,28,29] The second feature is also crucial as deterministic global optimization methods typically exhibit exponential complexity in the number of decision variables. Hence, solving a sequence of estimation tasks, each of which comprises a smaller number of parameters, can be expected to significantly decrease the overall computational requirement. Last but not least, although Brendel et al.[24] developed the incremental approach for identifying reaction systems, extending these ideas to general DAE systems poses no conceptual difficulty.

A somewhat similar approach has been presented by Ljung and Glad[30] in the context of global identifiability analysis. They show that a parameter estimation problem for a general ODE model comprising n parameters can be broken down (provided it is globally structurally identifiable) into n algebraic parameter estimation problems, one for each individual parameter. However, derivatives of the inputs and measured outputs up to an arbitrary order have to be calculated in this approach. In practice, it is almost impossible to obtain reasonable estimates for derivatives of order three or higher.

Observe that the proposed IGPE method lies in-between the classical, simultaneous approach, where no differentiation of the measurements is necessary and all model parameters are estimated simultaneously in one dynamic parameter estimation problem, and the approach by Ljung and Glad,[30] where each parameter can be estimated in a separate algebraic parameter estimation problem that contains high-order derivatives of the measurements. In the IGPE approach, only first-order derivatives of the measurements must be estimated and the dynamic estimation problem is typically broken down into a small number of algebraic problems, each of which comprises a reduced set of parameters to be estimated.

The steps of the proposed IGPE method for solving problem (3) are depicted in Figure 1 and are detailed subsequently.

Step 1 - Estimation of Derivatives

The first step in the IGPE method is to estimate the state time derivatives based on the available measurements. Estimating time derivatives is an ill-posed problem in the sense of Hadamard[31]

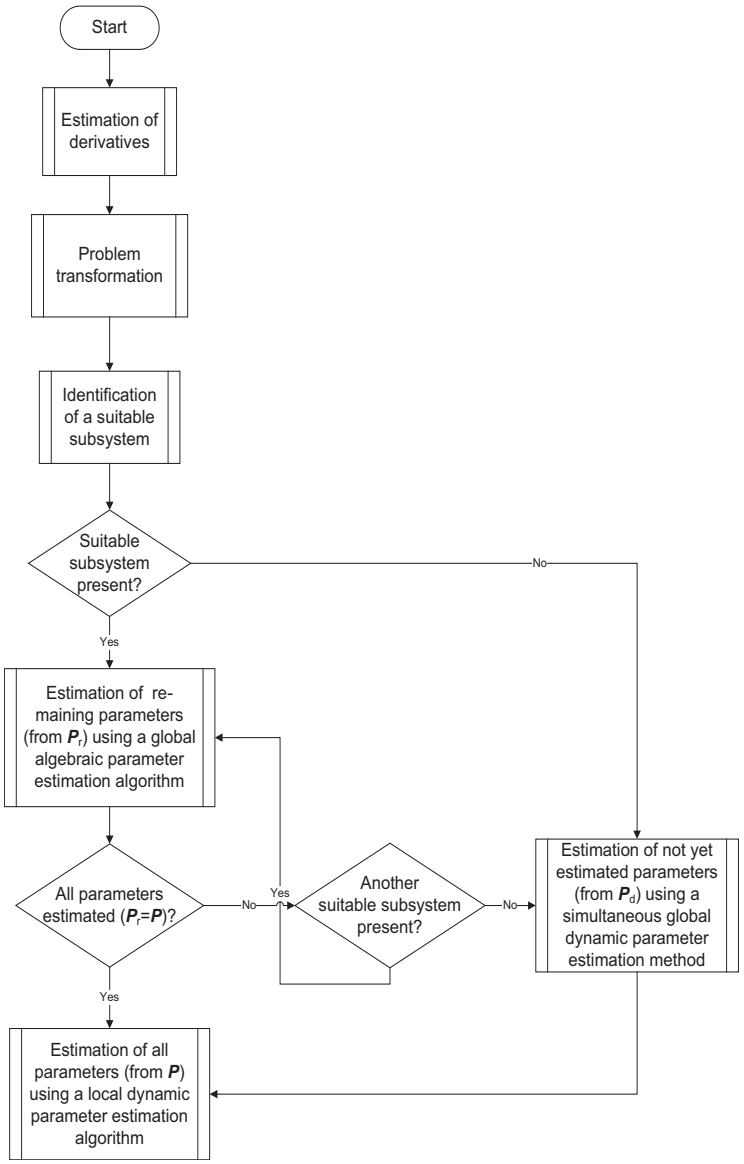

Figure 1: Flowchart of the IGPE method.

and thus requires appropriate regularization strategies. A variety of methods, such as filter[32,33] or spline-based[34–36] methods, have been proposed for this task. Throughout this paper, we use boundary corrected cubic smoothing splines, as described by Huang,[37] and subsequent differentiation of the smooth spline function to estimate the derivatives. The regularization parameter is determined using generalized cross validation. A comprehensive introduction to the solution of ill-posed problems is provided by Hansen[38] who also has provided a freely available MATLAB toolbox.[39] It is important to mention that the estimated derivatives are generally erroneous, no matter which method is used. The quality of the estimates mainly depends on the quality of the

available measurement data in terms of their frequency and noise content. However, the examples treated in this work show that, even when noisy and relatively scarce data are used, good estimates of the derivatives can be obtained with an adequate regularization strategy.

Subsequently, $\hat{\dot{\mathbf{x}}}_{m,i}$ denotes the vector of estimated time derivatives for the measured differential variables at the measurement times $t_i,\ i = 1,\ldots,\ell$.

Step 2 - Problem Transformation and Analysis

The objective of the problem transformation is to determine a set of algebraic equations that can be used in an algebraic parameter estimation problem. This is achieved by first removing those equations containing time derivatives of non-measured state variables from the model, then discretizing the remaining equations at the measurement times. There results a set of algebraic equations in the form:

$$\mathbf{F}_a(\dot{\mathbf{x}}_{m,i}, \mathbf{x}_{r,i}, \mathbf{y}_{r,i}, \mathbf{u}_{r,i}, \mathbf{p}_r) = \mathbf{0}, \quad i = 1,\ldots,\ell, \tag{4}$$

where $\mathbf{F}_a$ is a subset of algebraic equations from the original model $\mathbf{F}$; $\mathbf{p}_r \in P_r \subseteqq P$ stands for the subset of parameters present in $\mathbf{F}_a$; $\dot{\mathbf{x}}_{m,i} \in \mathbb{R}^{n_m}$ is the vector of time derivatives of the measured states at time t_i; and $\mathbf{x}_{r,i} \in \mathbb{R}^{n_{x,r}}$, $\mathbf{y}_{r,i} \in \mathbb{R}^{n_{y,r}}$ and $\mathbf{u}_{r,i} \in \mathbb{R}^{n_{u,r}}$ are the vectors of the remaining differential variables, algebraic variables and inputs at time t_i, respectively.[3] Let us also introduce the set $P_d \subseteqq P$ for those parameters not present in the reduced set of equations (4).

System analysis starts with the identification of a subset of equations that is suitable for use in the subsequent algebraic parameter estimation step (see Subsection 3.3). In this respect, the reduced system (4) must be solvable for the state time derivatives $\dot{\mathbf{x}}_{m,i},\ i = 1\ldots,\ell$. A prerequisite to this computation is that the number of unknown variables in (4) must be equal to the number of equations; in other words, the reduced subsystem (4) must be *well-constrained* in the sense of Dulmage and Mendelsohn.[40] In the case of an ODE system, for instance, well-constrainedness imposes that measurements for *all* the state variables $\mathbf{x}_{r,i}$ present in (4) must be available.

[3] Since potentially some of the original equations have been removed from the system, the set of differential variables, algebraic variables, inputs and parameters still present in the system might also be reduced.

Note that well-constrainedness is a necessary, yet not sufficient, condition for solvability of a set of algebraic equations. Nevertheless, testing a nonlinear system of equations for solvability (and verifying the uniqueness of its solution) is a very complex task that is beyond the scope of this work; see, e.g. Dulmage and Mendelsohn[40] or Borisevich et al.[41] for more details on this topic. In the remainder of the paper, a reduced subsystem $\mathbf{F}_a$ is called *suitable* provided that it is well-constrained.

If the naturally arising system $\mathbf{F}_a$, that is, the set of algebraic equations obtained by discretizing the original model according to the measurement grid and removing the equations containing derivatives of unmeasured states does not form a suitable subsystem, a suitable subsystem can generally be obtained by further removing equations from $\mathbf{F}_a$. To this end, no algorithmic procedure for automatically obtaining a suitable subsystem is available. Therefore, this step has to be performed manually by a modelling expert. In addition, different suitable subsystems might be obtained depending on which and how many equations are removed from $\mathbf{F}_a$. Hence, various choices of suitable subsystems (potentially containing different sets of parameters) will generally be possible. Again, no algorithmic procedure for the optimal choice of a suitable subsystem can be given here. The situation is comparable to modelling itself, where the same physical system can modelled in various ways and only vague guidelines are available on how to build a *good* model. Some general guidelines for the selection of an appropriate suitable subsystem are discussed in Section 3.6.

Step 3 - Global Algebraic Parameter Estimation

Having identified a suitable subsystem in the previous step it is possible to post a least-squares problem minimizing the difference between the predicted and estimated time derivatives. Because the independent variables of this parameter estimation problem also contain the original measurements, which are corrupted by noise, an error-in-variables approach could be applied. Nevertheless, the errors in the estimated state time derivatives are expected to be significantly larger than the errors in the measured states, so a least squares approach can also be applied, here.[24] This latter

approach leads to the parameter estimation problem

$$\begin{aligned} \min_{\mathbf{p}_r \in P_r} \quad & (\hat{\dot{\mathbf{X}}}_m - \dot{\mathbf{X}}_m(\mathbf{p}_r))^T \mathbf{W} (\hat{\dot{\mathbf{X}}}_m - \dot{\mathbf{X}}_m(\mathbf{p}_r)) \\ \text{s.t.} \quad & \text{Eq. (4)} \\ & \mathbf{p}_{r,l} \leq \mathbf{p}_r \leq \mathbf{p}_{r,u} \end{aligned} \tag{5}$$

where $\dot{\mathbf{X}}_m$ and $\hat{\dot{\mathbf{X}}}_m$ are column vectors containing the predicted and estimated state time derivatives, respectively, analogous to (2); $\mathbf{p}_{r,l}$ and $\mathbf{p}_{r,u}$ are the vectors of lower and upper bounds on $\mathbf{p}_r$, respectively.

If the estimation problem (5) is nonlinear in the parameters it can be solved using any code for the global optimization of algebraic models. If (5) is linear in the parameters, the problem can be solved non iteratively to global optimality. In this work, the code BARON[29] is used to solve the nonlinear algebraic global parameter estimation problems and SNOPT[42] is used within BARON as nonlinear optimization routine. Note that BARON is freely available via the NEOS server for optimization[4]. All computations presented in this paper have been performed on a Intel Core Duo computer with 2.13 GHz with 2 GB RAM, running Windows XP.

Step 4 - Dynamic Complement

This step consists in estimating the remaining parameters, i.e., those parameters which could not be estimated from any of the suitable subsystems. The solution of the so-called dynamic complement problem employs a global dynamic parameter estimation method, and is only necessary if $P_r \neq P$. However, the computational load is significantly reduced in comparison to the full simultaneous global parameter estimation, since the dynamic complement problem typically contains a reduced set of unknown parameters or can often be skipped completely.

[4] `http://www-neos.mcs.anl.gov`

Step 5 - Simultaneous Correction

This last step is the same as the simultaneous correction step in the incremental identification procedure by Brendel et al.[24] It consists in estimating all the parameters simultaneously by using a *local* dynamic optimization solver, wherein the parameter estimates obtained incrementally are used as the initial guess. Unlike incremental parameter estimates, which are inevitably biased due to error propagation from the estimated state derivatives, the simultaneous parameter estimates determined in this step are statistically optimal.

Discussion

Several remarks on the IGPE algorithm are in order:

- The algorithm as depicted in Figure 1 contains a loop. This loop refers to the option to solve a sequence of parameter estimation problems in step 3, i.e., when several suitable subsystems are present. In this case, each subproblem comprises fewer parameters than the original simultaneous estimation problem. The advantage of such a decomposition lies in the significant reduction in computational load which typically is obtained when the estimation is performed in a reduced search space. On the other hand, however, only part of the available structural information and measurement data is used in each subproblem, so the parameter estimates can end-up being less accurate. Overall, the decision about which subsystem to choose generally results from a tradeoff between accuracy and computational time. As a general guideline, it is advisable to choose the largest estimation problem that can be solved within an acceptable amount of time.

- Obviously, the quality of the estimated derivatives strongly depends on the quality and time resolution of the measurement data. If the data are too scarce or noisy, the estimated derivatives will be highly erroneous. The errors in the estimated derivatives will then propagate to the parameter estimates obtained in step 3. The question therefore arises whether (and

under which circumstances) the solution of the IGPE method is identical to the solution obtained using simultaneous global parameter estimation, which suffers from erroneous and low resolution measurement data in a different way.

Let $\mathbf{p}_{opt}$ denote the statistically sound parameter estimates as obtained from simultaneous global parameter estimation, and let $\mathbf{p}_{IGPE,4}$ and $\mathbf{p}_{IGPE,5}$ denote the parameter estimates obtained after steps 4 and 5 of the proposed incremental approach. Provided that the global estimates $\mathbf{p}_{opt}$ are unique and that sufficiently rich data are available to obtain good estimates of the state time derivatives, $\mathbf{p}_{IGPE,4}$ is generally in the neighborhood of $\mathbf{p}_{opt}$. In turn, this allows local simultaneous correction in step 5 of the IGPE approach to converge to the globally optimal parameter estimates, $\mathbf{p}_{IGPE,5} = \mathbf{p}_{opt}$. Nevertheless, conditions on the measurement data under which $\mathbf{p}_{IGPE,4}$ would end up sufficiently close to $\mathbf{p}_{opt}$ are presently unavailable.

Therefore, the IGPE algorithm can be seen as a *pseudo-deterministic global optimization algorithm*, in the sense that deterministic global optimization is used, but no guarantee can be given that the global optimum is obtained. Our experience with IGPE, however, is that the likelihood of finding the global optimum is excellent in case of high quality measurements and still reasonably good in case of relatively scarce and noisy data (see case studies in Sections 4 and 5).

- It is important to mention that situations may arise where Step 3 of the IGPE procedure cannot be applied, in which case the IGPE proccdure reduces to a simultaneous dynamic estimation problem (see Figure 1). This is typically the case when very few state variables are measured, which precludes the formulation of suitable subsystems. In contrast, a simultaneous dynamic parameter estimation problem can be formulated (and solved) as soon as a single data point is available.

To illustrate these important considerations, consider the simple reaction system:

$$x_1 + x_2 \rightarrow x_3$$

$$x_4 + x_5 \rightarrow x_6.$$

for which a model is given by:

$$\dot{x}_1 = -k_1 x_1 x_2 \tag{6}$$

$$\dot{x}_2 = -k_1 x_1 x_2 \tag{7}$$

$$\dot{x}_3 = k_1 x_1 x_2 \tag{8}$$

$$\dot{x}_4 = -k_2 x_4 x_5 \tag{9}$$

$$\dot{x}_5 = -k_2 x_4 x_5 \tag{10}$$

$$\dot{x}_6 = k_2 x_4 x_5, \tag{11}$$

and consider the problem of estimating both k_1 and k_2 based on the measurements of x_1 only. Clearly, this available measurements do not contain any information on the second reaction since the two reactions are independent, and it is therefore inadequate for estimating k_2. However, this deficiency cannot be detected when following a simultaneous estimation approach, which then results in an arbitrary value for k_2. On the other hand, no suitable subsystem can be isolated for estimating the value of k_2 when using IGPE, which is reasonable. Accordingly, the search for suitable subsystems in the IGPE approach can help reveal situations where part of the parameters are not structurally identifiable. Though one must be aware that situations also exist where suitable subsystems cannot be found, yet simultaneous estimation methods produce unique estimates for all the parameters. Hence, one cannot conclude that certain parameters are structurally unidentifiable based on the lack of a suitable subsystem for the estimation of these parameters. However, such a situation generally indicates that the available measurement data are inadequate for the considered identification

task.

Detailed Case Study

To demonstrate the IGPE approach, the case study of a DAE model identification is considered next. This test problem is new and, to the authors' best knowledge, has not yet been solved using global parameter estimation techniques. The DAE model describes a reaction system and is due to Bauer et al.[43] We reproduce the model in a slightly modified form:

$$
\begin{aligned}
\dot{n}_A &= V\,(-r_1 - r_2 - 3r_4) + f_A, \quad n_A(t=0) = n_{A_0},\\
\dot{n}_B &= -V\,r_1 + f_B, \quad n_B(t=0) = n_{B_0},\\
\dot{n}_C &= V\,(r_1 - r_2 + r_3), \quad n_C(t=0) = n_{C_0},\\
0 &= n_A + n_C + 2n_D + 3n_E - (n_{A_0} + n_{C_0} + 2n_{D_0} + 3n_{E_0} + f_A(t-t_0)),\\
0 &= n_B + n_C + n_D - (n_{B_0} + n_{C_0} + n_{D_0} + f_B(t-t_0)),\\
V &= n_A\frac{M_A}{\rho_A} + n_B\frac{M_B}{\rho_B} + n_C\frac{M_C}{\rho_C} + n_D\frac{M_D}{\rho_D} + n_E\frac{M_E}{\rho_E} + n_S\frac{M_S}{\rho_S},\\
k_1 &= k_{ref,1}\,e^{(-Ea_1/R)(1/T - 1/T_{ref})},\\
k_2 &= k_{ref,2}\,e^{(-Ea_2/R)(1/T - 1/T_{ref})},\\
k_3 &= k_2/k_c,\\
k_4 &= k_{ref,4}\,e^{(-Ea_4/R)(1/T - 1/T_{ref})},\\
k_c &= k_{c,2}\,e^{(-dh_2/R)(1/T - 1/T_{ref})},\\
r_1 &= k_1\frac{n_A\,n_B}{V^2},\\
r_2 &= k_2\frac{n_A\,n_C}{V^2},\\
r_3 &= k_3\frac{n_D}{V},\\
r_4 &= k_4\frac{n_A^2}{V^2}.
\end{aligned}
$$

Here, A, B, C, D and E are the chemical species present and S is the solvent, n_c, n_{c_0}, ρ_c and M_c, with $c \in \{A, B, C, D, E, S\}$, are the numbers of moles present in the system, the numbers of moles initially present, the density and the molecular weight of c, respectively. $r_i, i = 1, \ldots, 4$ is the reaction rate of the i-th reaction, $f_{cf}, cf \in \{A, B\}$ is the constant molar feed of cf, R is the gas constant, T is the time varying temperature, T_{ref} is the reference temperature of 363.16 K, and V is the reaction volume. The 8 parameters that need to be estimated are the steric factors $k_{ref,i}, i = 1, 2, 4$, the equilibrium constant $k_{c,2}$, the activation energies $Ea_i, i = 1, 2, 4$ and the reaction enthalpy dh_2. We refer to the original work by Bauer et al.[43] for more details on the reaction system and the model.

Measurements are generated by simulating four different, optimally planned experiments[5] and adding white noise with a standard deviation of 0.5%. The degrees of freedom for the experimental design are given in Table 1 and the parameter values used for the simulations are given in Table 2. We assume that for each experiment, 40 equally distributed measurements of the concentrations of A, B and C and of the volume (V) are available. Thus the molar amounts n_A, n_B and n_C can directly be calculated. For the global parameter estimation task, we use the parameter bounds as given in Table 2.

Table 1: Degrees of freedom for the optimal experimental design

variable	type
$n_{c_0}, c = \{A, B, C, D, E\}$	initial condition
$f_c, c = \{A, B\}$	constant input
T	time-variant input

Table 2: Bounds and true values for the parameter estimation task.

parameter	$k_{ref,1}$	$k_{ref,2}$	$k_{ref,4}$	$k_{c,2}$	Ea_1	Ea_2	Ea_4	dh_2
value	2.5	50	0.8	120	42000	45000	57000	32000
lower bound	0	0	0	20	10000	10000	10000	10000
upper bound	100	100	100	200	100000	100000	100000	100000

[5] The experimental conditions are determined based on a D-Optimal design[44]

Step 1 - Estimation of Derivatives

In the first step, we obtain a smooth approximation of the measurement data and the estimated derivatives as outlined above. Note that we provide the Matlab routines used to estimate the derivatives along with the measurement data online. The link can be found in the supporting information section.

Step 2 - Problem Transformation and Analysis

For this case study, the discretized model is given by:

$$\dot{n}_{A,i} = V_i\,(-r_{1,i} - r_{2,i} - 3r_{4,i}) + f_A, \tag{12}$$

$$\dot{n}_{B,i} = -V_i\,r_{1,i} + f_B, \tag{13}$$

$$\dot{n}_{C,i} = V_i\,(r_{1,i} - r_{2,i} + r_{3,i}), \tag{14}$$

$$0 = n_{A,i} + n_{C,i} + 2n_{D,i} + 3n_{E,i} - (n_{A_0} + n_{C_0} + 2n_{D_0} + 3n_{E_0} + f_A\,(t - t_0)), \tag{15}$$

$$0 = n_{B,i} + n_{C,i} + n_{D,i} - (n_{B_0} + n_{C_0} + n_{D_0} + f_B\,(t - t_0)), \tag{16}$$

$$V_i = n_{A,i}\frac{M_A}{\rho_A} + n_{B,i}\frac{M_B}{\rho_B} + n_{C,i}\frac{M_C}{\rho_C} + n_{D,i}\frac{M_D}{\rho_D} + n_{E,i}\frac{M_E}{\rho_E} + n_{S,i}\frac{M_S}{\rho_S}, \tag{17}$$

$$k_{1,i} = k_{ref,1}\,e^{(-Ea_1/R)(1/T_i - 1/T_{ref})}, \tag{18}$$

$$k_{2,i} = k_{ref,2}\,e^{(-Ea_2/R)(1/T_i - 1/T_{ref})}, \tag{19}$$

$$k_{3,i} = k_{2,i}/k_{c,i}, \tag{20}$$

$$k_{4,i} = k_{ref,4}\,e^{(-Ea_4/R)(1/T_i - 1/T_{ref})}, \tag{21}$$

$$k_{c,i} = k_{c,2}\,e^{(-dh_2/R)(1/T_i - 1/T_{ref})}, \tag{22}$$

$$r_{1,i} = k_{1,i}\frac{n_{A,i}\,n_{B,i}}{V_i^2}, \tag{23}$$

$$r_{2,i} = k_{2,i}\frac{n_{A,i}\,n_{C,i}}{V_i^2}, \tag{24}$$

$$r_{3,i} = k_{3,i}\frac{n_{D,i}}{V_i}, \tag{25}$$

$$r_{4,i} = k_{4,i}\frac{n_{A,i}^2}{V_i^2}, \tag{26}$$

with $i = 1, \ldots, 40$. Generally, the first step is to remove any equations containing the derivatives of non-measured species. In this case, no such equation exists. The model contains $15 \times 40 = 600$ equations, $15 \times 40 = 600$ unknown variables ($\dot{n}_{A,i}$, $\dot{n}_{B,i}$, $\dot{n}_{C,i}$, $n_{D,i}$, $n_{E,i}$, $n_{S,i}$, $k_{1,i}$, $k_{2,i}$, $k_{3,i}$, $k_{4,i}$,$k_{c,i}$, $r_{1,i}$, $r_{2,i}$, $r_{3,i}$, $r_{4,i}$), 14 known parameters (ρ_A, ρ_B, ρ_C, ρ_D, ρ_E, ρ_S, M_A, M_B, M_C, M_D, M_E, M_S, T_{ref}, R), 5 known initial conditions (n_{A_0}, n_{B_0}, n_{C_0}, n_{D_0}, n_{E_0}), 2 known time invariant inputs (f_A, f_B),

1 known time variant input (T) and 8 parameters which are to be estimated but assumed known (assigned) for the identifiability test. Hence this model forms a suitable subsystem and can be used in the IGPE approach.

Note that in this case, a decomposition of the problem would also be possible. The identification task can be split up into first solving for the parameters $k_{ref,1}$ and Ea_1 using the subsystem consisting of Eqs. (13),(18) and (23) and then solving for the remaining six parameters using the complete model as given in Eqs. (12) - (26) but considering $k_{ref,1}$ and Ea_1 known.

Step 3 - Global Parameter Estimation

The resulting algebraic parameter estimation task is to minimize the squared residuals between the estimated state derivatives ($\hat{\dot{n}}_{A,i}$, $\hat{\dot{n}}_{B,i}$, $\hat{\dot{n}}_{C,i}$) and those predicted by the algebraic model (12)-(26). This parameter estimation task could be solved using BARON in approximately 92 seconds on an Intel Core Duo 2,13 Ghz computer running windows XP; the results are reported in Table 3 as $p_{IGPE,3}$.

Table 3: Optimal parameter estimates after step 3 and step 5.

parameter	$k_{ref,1}$	$k_{ref,2}$	$k_{ref,4}$	$k_{c,2}$	Ea_1	Ea_2	Ea_4	dh_2
$\mathbf{p}_{IGPE,3}$	2.566	49.890	0.799	120.728	42383.595	44986.27	56930.189	31899.157
$\mathbf{p}_{IGPE,5}$	2.503	50.09	0.7995	120.3	42000	45010	56970	31970

The results obtained with a decomposition of the estimation problem into two suitable subsystems, as discussed in the previous subsection, are very similar to those given in Table 3. However, such a decomposition was not found to provide a significant reduction in CPU time here.

Step 4 - Dynamic Complement

Since all parameters could be estimated in step 3, step 4 can be skipped here.

Step 5 - Simultaneous Correction

Solving the original dynamic parameter estimation problem (3) with the local dynamic optimization method implemented in gPROMS, from the values obtained in step 3 as initial guess, takes approximately 19 second and yields the correct global solution; these values are reported as $p_{IGPE,5}$ in Table 3.

Conclusions

The detailed case study shows that high quality estimates of the globally optimal parameters can be obtained in step 3, if sufficient care is taken in the previous steps. Overall the global optimum could be obtained for this 8 parameter DAE example in only 111 seconds CPU-time. Since the performance of the method cannot be judged based on only one example, we consider six well known test cases from the literature in the following section.

Test Problems from the Literature

To allow fair comparisons between the IGPE method and simultaneous global dynamic parameter estimation, all six test problems from Floudas[45] have been solved, in addition to the DAE test problem treated previously (this latter problem is referred to as test problem 7 subsequently). Such test problems have been (completely or partly) studied by different authors — see for instance the works by Esposito and Floudas,[17] Lin and Stadtherr[46] or Papamichail and Adjiman.[18,47] Their formulations are summarized subsequently:

Test problem 1:

$$\dot{x}_1 = -p_1 x_1$$

$$\dot{x}_2 = p_1 x_1 - p_2 x_2$$

$$\mathbf{x}(0) = (1,0)^T$$

$$0 \le \mathbf{p} \le 10$$

Test problem 2:

$$\dot{x}_1 = -p_1 x_1 + p_2 x_2$$

$$\dot{x}_2 = p_1 x_1 - (p_2 + p_3) x_2 + p_4 x_3$$

$$\dot{x}_3 = -p_4 x_3 + p_3 x_2$$

$$\mathbf{x}(0) = (1,0,0)^T$$

$$(0,0,10,10)^T \le \mathbf{p} \le (10,10,50,50)^T,$$

Test problem 3:

$$\dot{x}_1 = -(p_1 + p_3) x_1^2$$

$$\dot{x}_2 = p_1 x_1^2 - p_2 x_2$$

$$\mathbf{x}(0) = (1,0)^T$$

$$0 \le \mathbf{p} \le 20$$

Test problem 4:

$$\dot{x} = p_1 (126.2 - x)(91.9 - x)^2 - p_2 x^2$$

$$x(0) = 0$$

$$0 \le \mathbf{p} \le 0.1$$

Test problem 5:

$$\dot{x}_1 = x_1 \Big(-2p_1 + \frac{p_1 x_2}{(p_2 + p_5) x_1 + x_2} - p_3 - p_4 \Big)$$

$$\dot{x}_2 = \frac{p_1 x_1 (p_2 x_1 - x_2)}{(p_2 + p_5) x_1 + x_2} + p_3 x_1$$

$$\dot{x}_3 = \frac{p_1 x_1 (x_2 + p_5 x_1)}{(p_2 + p_5) x_1 + x_2} + p_4 x_1$$

$$\mathbf{x}(0) = (1,0,0)^T$$

$$0 \le \mathbf{p} \le 100$$

Test problem 6:

$$\dot{x}_1 = p_1 x_1 (1 - x_2)$$

$$\dot{x}_2 = p_2 x_2 (x_1 - 1)$$

$$\mathbf{x}(0) = (1.2, 1.1)^T,$$

$$0 \le \mathbf{p} \le 10$$

The corresponding measurement data and model files are available online, the link can be found in the supporting information section.

It should be noted that most of these problems are not particularly well suited for application of the IGPE method, since the measurement data are rather scarce. The IGPE method clearly works better with high resolution measurement data, stemming for instance from spectroscopic measurement devices. Mention should also be made that the resulting algebraic parameter estimation problems are linear in the parameters in test problems 1, 2, 3, 4 and 6, while they are nonlinear in test problem 5 and 7.

Discussion

The results and computational times obtained using the novel IGPE method are summarized in Table 4; in the first row the test problem number is given. The second row shows the relative difference of the parameter estimates available after step 4 of the IGPE approach and the global optimum. The third row indicates whether or not the global optimum has been obtained in the last step, the simultaneous correction, while the last row gives the computational times. For test cases 5 and 7, BARON has been used to solve the global algebraic parameter estimation problem. The algebraic parameter estimations of the remaining test problems could be solved non-iteratively due to the linear nature of the resulting subproblems. For all test problems, gPROMS has been used for step 5, the simultaneous correction. It is seen that the proposed method works well for a set of standard test problems from the literature. The method could directly be applied to all six test problems from Floudas[45] and could also be used to solve one exemplary DAE test problem comprising eight parameters. Although the problems are not especially well suited for the method, because the available data are generally rather scarce and in part corrupted by noise (test problems 3-7), the estimates obtained after step 4 of the IGPE approach show on average a bias of only 8% and a maximum bias of 32.52%. The global optimum is obtained for 6 out of the 7 test cases.

The problem of not converging to a global minimizer in test problem 6 can be explained by

Table 4: Summary of the results and computational times for the 7 test problems considered using the IGPE method.

problem no.	1	2	3	4	5	6	7
No. of equations	2	3	2	1	3	2	15
No of parameters	2	3	2	1	5	2	8
relative mean difference between $\mathbf{p}_{IGPE,4}$ and the global optimum	0.38%	1.80%	15.40%	1.77%	4.04 %	32.52%	0.57%
$\mathbf{p}_{IGPE,5} = \mathbf{p}_{opt}$	yes	yes	yes	yes	yes	no	yes
CPU-time [sec.]	< 1	<1	<1	<1	≈ 350	< 1	≈ 110

means of Figure 2, which shows the trajectory calculated using the globally optimal parameter values, the provided measurement data and the smoothing splines approximation of the data. It is apparent that the data are too scarce to reflect the dynamics of the system. This results in low quality estimates for the time derivatives. In addition, Esposito and Floudas[17] showed that this problem comprises a large number of local minima relatively close to one another. Therefore, even slight deviations from the globally optimal solution can lead to convergence to a false, local optimum in the simultaneous correction step.

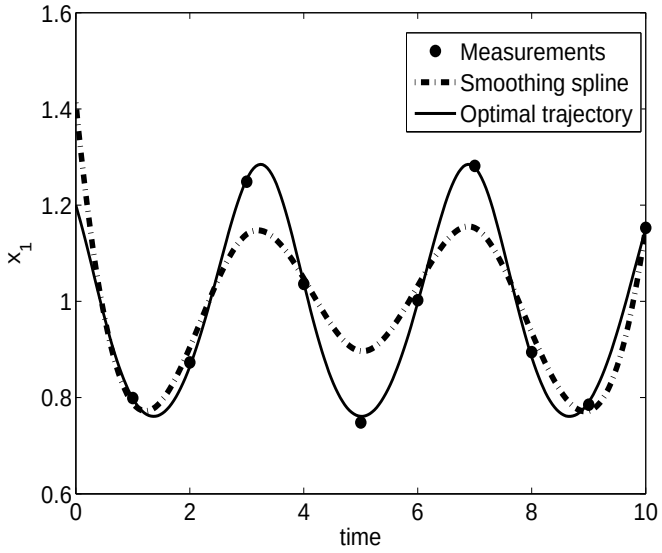

Figure 2: Measurements, optimal trajectory and smoothing splines approximation for test problem 6.

The computational times reported in Table 4 clearly validate that one of the major advantages of using IGPE lies in its computational efficiency. To the best knowledge of the authors, the IGPE

method solves the considered test problems at least one order of magnitude faster, and on average two to three orders of magnitude faster, than any other deterministic global optimization method. Hence, this method allows solution of deterministic global parameter estimation for reasonably-sized dynamic models comprising a large number of parameters. Thus far, such problems could only be tackled using stochastic optimization approaches.

The IGPE method also has a number of limitations. First and foremost, no guarantee that a global optimum is found can be given. As seen in test problem 6, the use of low quality and scarce measurement data to solve a problem having many local minima may result in the global optimum being missed. Therefore, the method is best suited for the identification of complex dynamic systems based on large sets of high resolution measurement data. Besides having high quality measurements, it is also advantageous to measure a large fraction of the state variables (all state variables ideally), since the number of suitable subsystems, and thus the potential to decompose the estimation task, becomes larger as more state derivatives can be estimated. This way, the computational efficiency also increases with the number of state measurements. Test problem 7 shows that high-resolution, low-noise data result in high-quality parameter estimates.

Overall, one can conclude that the benefit of using the proposed IGPE method strongly depends on the available measurement data. The treated case study demonstrate that the IGPE approach offers much potential to significantly decrease the computational effort. Nevertheless, situations may also arise where a classical simultaneous approach is more favorable or where IGPE cannot be applied in contrast to classical approaches.

In future work the IGPE method will be further refined so as to alleviate the issue of not finding a global optimum for the original simultaneous estimation problem. One idea would be to not only solve step 3 to global optimality, but also retain a number of local optima whose solution values are close enough to the global solution value. These solutions could then be used as various initial guesses in the simultaneous correction step. The occurrence of multiple local solutions that are close in solution value to one another is also interesting from a model development perspective. This information would indeed indicate that the confidence in the parameter values might be low,

even though a confidence analysis at the global optimal solution had not revealed it. In this case either more data or a lessRef. 1 is empty in the pdf file. Don't know why... complex model structure should be considered.

It would also be interested to use bootstrapping methods[48,49] for estimating the error on the estimated derivatives. In turn, these error estimates could be used to decide on the number of local optima that are to be retained in step 3. This way the probability of obtaining a global optimum could be estimated. Furthermore, knowledge about the errors in the estimated derivatives could be used to solve a weighted, as opposed to a non-weighted, least-squares problem in step 3 of IGPE.

Finally, in its current form, IGPE requires the user to identify suitable subsystems and decide on which one(s) to use. To this end, an algorithmic procedure that automatically determines all or a subset of the most promising suitable subsystems is highly desirable and will be in the focus of future research.

Acknowledgments

The financial support by the the *Deutsche Forschungsgemeinschaft* (DFG) via *collaborative research center 540* (SFB 540) is greatly appreciated. The authors would also like to thank Andreas Harwardt for providing support in the GAMS modeling language, which has been used to implement and solve the algebraic parameter estimation problems using BARON and Maka Karalashvili for many fruitful discussion on the properties of the IGPE method. Last but not least the authors would like to thank the staff of the Laboratoire d'Automatique, at the École Polytechnique Fédérale de Lausanne (EPFL) for their hospitality. Most of the work presented in this paper has been completed during a research visit of Claas Michalik at the Laboratoire d'Automatique.

Supporting Information

All model files, measurement data and Matlab routines used to solve the case studies treated in this work can be found online at: http://www.avt.rwth-aachen.de/AVT/index.php?id=686. This

information is available free of charge via the Internet at http://pubs.acs.org.

References

1. Biegler, L.; Grossmann, I.; Westerberg, A. *Systematic Methods of Chemical Process Design*; Prentice-Hall: Englewood Cliffs, NJ, 1997.

2. Rosenbrock, H. H.; Storey, C. *Computational Techniques for Chemical Engineers*; Pergamon Press, 1966.

3. Bellman, R.; Jacquez, J.; Kalaba, R.; Schwimmer, S. Quasiliniearization and the Estimation of Chemical Rate Constants from Raw Kinetic Data. *Mathematical Biosciences* **1967**, *1*, 71–76.

4. Kim, I.-W.; Liebman, M. J.; Edgar, T. F. A sequential error-in-variables method for nonlinear dynamic systems. *Comput. Chem. Eng.* **1991**, *15*, 663–670.

5. Process Systems Enterprise Ltd, `http://www.psenterprise.com/gproms/`.

6. Numerica Technology LLC, `http://www.numericatech.com/`.

7. Bock, H. G. Recent advances in parameter identification techniques for ODE. In *Numerical Treatment of Inverse Problems in Differential and Integral Equations*; Deuflhard, P., Hairer, E., Eds.; Birkhäuser: Boston, 1983; pp 95–121.

8. Lohmann, T.; Bock, H. G.; P., S. J. NUMERICAL-METHODS FOR PARAMETER-ESTIMATION AND OPTIMAL EXPERIMENT DESIGN IN CHEMICAL-REACTION SYSTEMS. *Industrial & Engineering Chemistry Research* **1992**, *31*, 54–57.

9. Peifer, M.; Timmer, J. Parameter estimation in ordinary differential equations for biochemical processes using the method of multiple shooting. *IET Systems Biology* **2007**, *1*, 78–88.

10. Van Den Bosch, B.; Hellinckx, L. A new method for the estimation of parameters in differential equations. *AIChE J* **1974**, *20*, 250–255.

11. Tjoa, T. B.; Biegler, L. T. Simultaneous solution and optimization strategies for parameter estimation of differential-algebraic equation systems. *Ind. Eng. Chem. Res.* **1991**, *30*, 376–385.

12. Singer, A. B.; Taylor, J. W.; Barton, P. I.; Green Jr., W. H. Global dynamic optimization for parameter estimation in chemical kinetics. *J Phys Chem A* **2006**, *110*, 971–976.

13. Guus, C.; Boender, E.; Romeijn, H. E. Stochastic methods. *Handbook of global optimization*, Dordrecht, The Netherlands, 1995; pp 829–869.

14. Törn, A.; Ali, M.; Viitanen, S. Stochastic global optimization: Problem classes and solution techniques. *J Global Optim* **1999**, *14*, 437–447.

15. Moles, C. G.; Mendes, P.; Banga, J. R. Parameter Estimation in Biochemical Pathways: A Comparison of Global Optimization Methods. *Genome Res.* **2003**, *13*, 2467–2474.

16. Horst, R.; Tuy, H. *Global Optimization: Deterministic Approaches*, 3rd ed.; Springer-Verlag: Berlin, Germany, 1996.

17. Esposito, W. R.; Floudas, C. A. Global Optimization for the Parameter Estimation of Differential-Algebraic Systems. *Ind. Eng. Chem. Res.* **2000**, *39*, 1291–1310.

18. Papamichail, I.; Adjiman, C. S. Global Optimization of dynamic systems. *Computers Chem. Eng.* **2004**, *28*, 403–415.

19. Chachuat, B.; Latifi, M. A. A new approach in deterministic global optimization of problems with ordinary differential equations. *Frontiers in Global Optimization*, 2003; pp 83–108.

20. Chachuat, B.; Singer, A. B.; Barton, P. I. Global Solution of Dynamic Optimization and Mixed-Integer Dynamic Optimization. *Ind. Eng. Chem. Res.* **2006**, *45*, 8373–8392.

21. Lin, Y.; Stadtherr, M. A. Deterministic Global Optimization of Nonlinear Dynamic Systems. *AIChE J* **2007**, *53*, 866–875.

22. Hosten, L. H. A comparative study of short-cut procedures for parameter estimation in differential equations. *Comput. Chem. Eng.* **1979**, *3*, 117–126.

23. Bardow, A.; Marquardt, W. Incremental and simultaneous identification of reaction kinetics: Methods and comparison. *Chem. Eng. Sci.* **2004**, *59*, 2673–2684.

24. Brendel, M.; Bonvin, D.; Marquardt, W. Incremental identification of complex reaction kinetics in homogeneous systems. *Chemical Engineering Science* **2006**, *61*, 5404–5420.

25. Froment, G. F.; Bischoff, K. B. *Chemical Reactor Analysis and Design*; Wiley: New York, NY, 1990.

26. Tawarmalani, M.; Sahinidis, N. Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming: Theory, Algorithms, Software and Applications **2005**.

27. Grossmann, I. E.; Biegler, L. T. Part II. Future perspective on optimization. *Computers & Chemical Engineering* **2004**, *28*, 1193–1218.

28. Floudas, C. A.; Akrotirianakis, I. G.; Caratzoulas, A.; Meyer, C. A.; Kallrath, J. Global optimization in the 21st century: Advances and challenges. *Computers & Chemical Engineering* **2005**, *29*, 1185–1202.

29. Sahinidis, N. V. BARON: A general purpose global optimization software package. *Journal of Global Optimization* **1996**, *8*, 201–205.

30. Ljung, L.; Glad, T. On Global Identifiability for arbitrary Model Parametrizations. *Automatica* **1994**, *30*, 265–276.

31. Hadamard, J. *Lectures on Cauchy problem in linear partial differential equations*, Oxford University Press, London, 1923.

32. Lázaro, M.; Santamaría., I.; Pantalén, C.; Ibénez, L., J. amd Vielva A regularized technique for the simultaneous reconstruction of a function and its derivatives with application to nonlinear transistor modeling. *Signal Processing* **2003**, *83*, 1859–1870.

33. Tikhonov, A. N.; Arsenin, V. A. *Solution of Ill-posed Problems*; Winston and Sons: Washington, 1977.

34. Nunhez, J. R.; Costa, C. B. B.; Guirardello, R. A new spline approach for data fitting. *Fluid Phase Equilibria* **2004**, *220*, 171–180.

35. Reinsch, C. H. Smoothing by Spline Functions. *Numer. Math.* **1967**, *10*, 177–183.

36. Ritter, K. Almost Optimal Differentiation Using Noisy Data. *Journal of Approximation Theory* **1996**, *86*, 293–309.

37. Huang, C. Boundary corrected cubic smoothing splines. *J. Statist. Comput. Simul.* **2001**, *70*, 107–121.

38. Hansen, P. C. *Rank-deficient and Discrete Ill-posed Problems*; SIAM, 1998.

39. Hansen, P. C. Regularization Tools Version 4.0 for Matlab 7.3. *Numerical Algorithms* **2007**, *46*, 189–194.

40. Dulmage, A. L.; Mendelsohn, N. S. Two algorithms for bipartite graphs. *SIAM Journal* **1963**, *11*, 183–194.

41. Borisevich, V. D.; Potemkin, V. G.; Stunkov, S. P.; Wood, H. G. Global Methods for Solving Systems of Nonlinear Equations. *Computers and Mathematics with Applications* **2000**, *40*, 1015–1025.

42. Gill, P. E.; Murray, W.; Saunders, M. A. An SQP algorithm for large-scale constrained optimization. *SIAM Journal on Optimization* **2002**, *12*, 979–1006.

43. Bauer, I.; Bock, H. G.; Koerkel, S.; Schloeder, J. P. Numerical methods for optimum experimental design in DAE systems. *Journal of Computational and Applied Mathematics* **2000**, *120*, 1–25.

44. Espie, D.; Macchietto, S. The Optimal Design of Dynamic Experiments. *AIChE Journal* **1989**, *35*, year.

45. Floudas, C. A. *Handbook of Test Problems in Local and Global Optimization*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1999.

46. Lin, Y.; Stadtherr, M. A. Deterministic Global Optimization for Parameter Estimation of Dynamic Systems. *Ind. Eng. Chem. Res.* **2006**, *45*, 8438–8448.

47. Papamichail, I.; Adjiman, C. S. A Rigorous Global Optimization Algorithm for Problems with Ordinary Differential Equations. *Journal of Global Optimization* **2002**, *24*, 1–33.

48. Duchesne, C.; MacGregor, J. F. Jackknife and bootstrap methods in the identification of dynamic models. *Journal of Process Control* **2001**, *11*, 553–564.

49. Wehrens, R.; Putter, H.; Buydens, L. M. C. The bootstrap: a tutorial. *Chemometrics and Intelligent Laboratory Systems* **2000**, *54*, 35–52.