IGPE – A Novel Method for Efficient Global Parameter Estimation in Dynamical Systems

C. Michalik[1], B. Chachuat[2], W. Marquardt[1]
[1]AVT - Process Systems Engineering, RWTH Aachen University, Germany
[2]Laboratoire d'Automatique, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland,
Current affiliation: Department of Chemical Engineering, McMaster University, Canada

Corresponding author: W. Marquardt, AVT - Process Systems Engineering, RWTH Aachen University, Turmstraße 46, D-52064 Aachen, Germany, Wolfgang.Marquardt@avt.rwth-aachend.de

Abstract. In model discrimination a model's quality is usually judged based on its ability to reproduce the available measurement data, after optimal values for its parameters have been determined. Globally optimal rather than locally optimal parameter estimates are clearly warranted to avoid false conclusions. However, obtaining globally optimal parameter estimates is computationally very expensive for models comprised of differential equations. In this paper, we present a novel pseudo-deterministic global parameter estimation methodology, which is capable of significantly reducing the computational load for global parameter estimation in dynamical systems. This method, named IGPE, builds upon the *incremental identification* approach developed by Marquardt and co-workers for the identification of reaction-diffusion systems. Unlike most existing deterministic methods, IGPE can handle both ODE and DAE models, and its application relies on readily available software packages. This paper presents the IGPE methodology and illustrates its pros and cons through a case study in the field of chemical reaction kinetics.

1 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. If uncorrelated, Gaussian noise is assumed for the measurement errors, this leads to the well-known class of least squares problems, where the objective 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]. 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. [3]; this is also the approach implemented in state-of-the-art process simulation software [4, 5]. 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 [6, 7].

The aforementioned approaches rely on local search methods and can, at best, achieve convergence to a local minimum. This deficiency to converge to a global minimum may have dramatic consequences. In the context of reaction kinetics, for instance, getting a bad fit (due to only locally optimal parameter values) may lead to the erroneous conclusion that a proposed kinetic model yields an incorrect description of the chemistry [8]; in other words, only a global minimum allows one to conclude that a proposed model structure is invalid.

Both *stochastic* and *deterministic global optimization* have been developed to increase the likelihood of finding a global minimizer. Stochastic search methods rely on probabilistic approaches [9, 10]. 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., Guus et al. [11] 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 [12]. These sound theoretical convergence properties have stimulated the development of deterministic global optimization for problems with embedded differential equations [13, 14, 15, 8, 16, 17]. Briefly, a spatial 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. 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 indeed

exhibits 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 [18]. Recently, the so-called *incremental* approach for model identification has been introduced by Marquardt and coworkers [19, 20, 21]. 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 this approach is related to the well-known *differential* method of parameter estimation in kinetic models [22]. Although incremental identification approaches do not share the same theoretical properties as simultaneous methods, namely they are neither unbiased nor consistent [18], their major advantage lies in the computational tractability. Moreover, comparisons have indicated that good results can be obtained provided that sufficient care is taken during the estimation [20].

This paper proposes a novel methodology for parameter estimation in dynamical systems. Building upon the incremental identification method, the original problem is split into 5 steps, which are detailed in Section 3. 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 [23], 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 downside, of course, is that – in the presence of measurement noise – the global solution in the incremental approach will generally be different from the global solution in the simultaneous approach. 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 finds a global optimizer to the original problem.

2 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 $[F_{\dot{\mathbf{x}}}\, F_{\mathbf{y}}]$ has full rank; that is the index of the DAEs is less than or equal to one. Initial values $\mathbf{x}(t_0) = \mathbf{x}_0$ are given for the differential state variables and corresponding consistent initial conditions $\dot{\mathbf{x}}(t_0)$, $\mathbf{y}(t_0)$ are assumed to be available. The inputs $\mathbf{u}(t)$ are also assumed to be given.

Let $I_m \subseteq \{1,\ldots,n_x\}$ denote the indices of the measured differential state variables[1] and let $t_1,\ldots,t_\ell$ such that $t_0 \le t_1 < \ldots < t_\ell = t_f$ 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, $i = 1,\ldots,\ell$, and let the corresponding vector of state predictions be denoted by $\mathbf{x}_{m,i}$. The measurements and corresponding model predictions are arranged in the column vectors

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

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

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

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.

[1] For simplicity of presentation we assume that only differential state variables are measured.
[2] We use the identity matrix as the weighting matrix throughout this paper.

3 The IGPE Method

The IGPE methodology is based on the incremental identification procedure developed by Marquardt and co-workers for reaction, diffusion and coupled reaction-diffusion systems [19, 20, 21]. The novelty in IGPE is twofold: (i) generic models comprised of DAEs can be handled, and (ii) global optimization is systematically employed to solve the parameter estimation subproblems.

IGPE proceeds by breaking a dynamic parameter estimation problem into five steps. These steps are described in subsections 3.1 to 3.5 and illustrated by the case study of a tubular gas-phase reactor, where the reaction $\mathsf{A} \rightarrow m\,\mathsf{B}$ takes place. The following model is due to Ingham [24] and assumes steady-state conditions in the tubular reactor:

$$n_{\mathsf{A}} = \frac{y_{\mathsf{A},G} P}{RT}, \tag{4}$$

$$n_{\mathsf{B}} = \frac{y_{\mathsf{B},G} P}{RT}, \tag{5}$$

$$G = \frac{(n_{\mathsf{A}} + n_{\mathsf{B}} + n_{\mathsf{inert}})RT}{P}, \tag{6}$$

$$y_{\mathsf{A}} = \frac{y_{\mathsf{A},G}}{G}, \tag{7}$$

$$y_{\mathsf{B}} = \frac{y_{\mathsf{B},G}}{G}, \tag{8}$$

$$\frac{dy_{\mathsf{A},G}}{dz} = -k y_{\mathsf{A}} \Omega, \tag{9}$$

$$\frac{dy_{\mathsf{B},G}}{dz} = mk y_{\mathsf{A}} \Omega. \tag{10}$$

In Eqs. (4)-(10), G stands for the total gas volumetric flow rate; n_i, y_i and $y_{i,G}$ denote the molar flow rate, the mole fraction and the volumetric flow rate of species $i \in \{\mathsf{A}, \mathsf{B}, \mathsf{inert}\}$, respectively; Ω, the cross-sectional area of the reactor; z, the axial position; P, the total pressure; R, the gas constant; T, the temperature; m, the stoichiometric coefficient of the reaction; and k, the reaction constant.

The measurement data used in this case study correspond to $y_{\mathsf{A},G}$ and $y_{\mathsf{B},G}$ along the axial position of the reactor under steady state operation and for different temperatures $T = 200$ K, 300 K and 400 K. These data are due to Schittkowski [25] and have been generated by numerical simulation and addition of Gaussian noise with a standard deviation of 5%.

The identification task is to estimate the values of P, m and k. Each of the 3 data sets consists of 13 measurements along the axial position of the reactor, ranging from $z = 0.1$ to $z = 20$. The measurement data, the model files and some MatLab™ routines can be retrieved from `http://www.avt.rwth-aachen.de/AVT/index.php?id=689`. The parameter values used for data generation are given in Table 1, along with the lower and upper bounds used for parameter estimation.

parameter	P	k	m
lower bound	10^{-4}	10^{-4}	0
exact	10^{4}	15	2
upper bound	10^{9}	10^{2}	10^{2}

Table 1: Parameter values used to create the measurement data and lower and upper bounds for the parameter estimation

3.1 Step 1 - Estimation of Derivatives

The first step in the IGPE method is to estimate the state derivatives based on the available measurements. Estimating derivatives is an ill-posed problem in the sense of Hadamard [26] and thus requires appropriate regularization strategies. A variety of methods, such as filter [27] or spline-based [28] methods, have been proposed for this task. In this paper, we use cubic smoothing splines, as described by Tikhonov [27], and subsequent differentiation of the smooth spline function to estimate the derivatives. The smoothing splines algorithm automatically filters the data such that low to high resolution and low to high noise level data can be used in exact the same manner.

This step is illustrated in Figure 1 for $y_{\mathsf{B},G}$ at $T = 400$ K. The measurements are shown on the left plot along with the exact values; the estimated derivative (as obtained from differentiation of the smooth spline function) and the exact derivative (calculated using the model and the exact parameter values given in Table 1) are given in the right plot. It is seen that the derivatives are approximated well, even though they are slightly biased close to $z = 0$.

3.2 Step 2: Problem Transformation and Analysis

The objective of problem transformation is to determine a set of algebraic equations that can be used in a global algebraic parameter estimation problem. This is achieved by first removing those equations containing derivatives

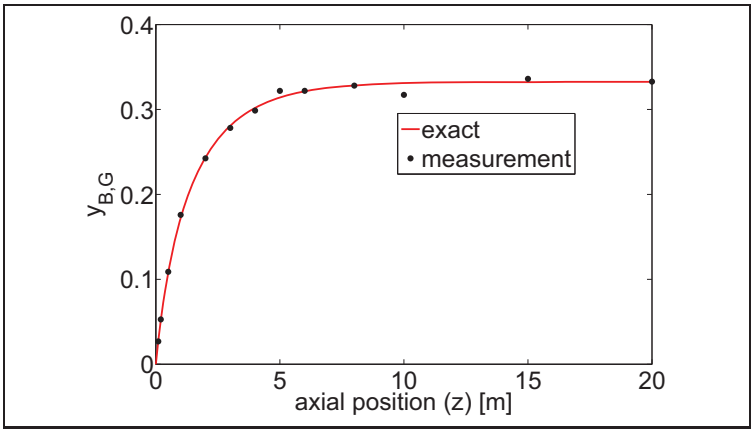

Figure 1: Measurements and exact trajectory.

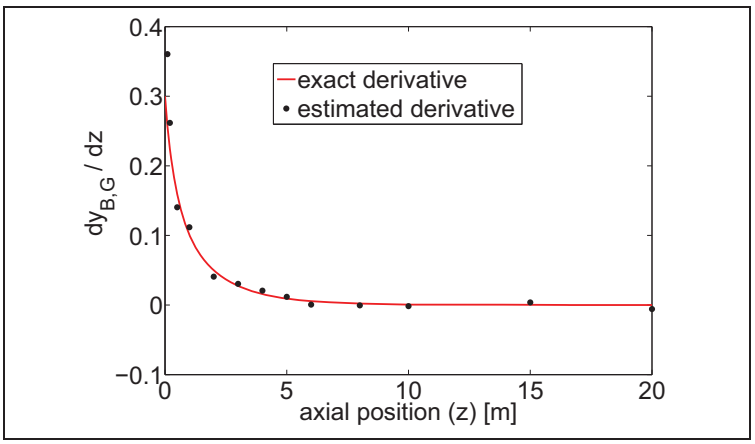

Figure 2: Estimated and exact derivative.

of unmeasured state variables from the dynamic model (1), and then discretizing the remaining equations at the measurement instances. Next, subsets of equations appropriate for the algebraic parameter estimation subproblem (see subsection 3.3) have to be identified. In these algebraic subproblems, the task is to determine those parameter values that minimize the difference between the estimated state derivatives (see subsection 3.1) and those predicted by the model. In particular, a subset of equations is said to form a suitable subsystem if it can be solved for the derivatives of the participating differential state variables ($\frac{dy_{A,G}}{dz}$ and $\frac{dy_{B,G}}{dz}$, here) at each measurement instance.

If the set of equations obtained by discretizing the model and removing those equations containing derivatives of unmeasured state variables does not form a suitable subsystem, a suitable subsystem can often be obtained by removing more equations from the model. In general, different suitable subsystems can be obtained depending on which and how many equations are removed from the original model; those subsystems may involve different subsets of parameters. Note however that no algorithmic procedure for the automatic generation of suitable subsystems is currently available. Therefore, this step has to be performed manually by a modeling expert.

In the case study treated here, discretization at the measurement times yields the following algebraic system:

$$n_{\mathrm{A},i} = \frac{y_{\mathrm{A},G,i}P}{RT_i}, \tag{11}$$

$$n_{\mathrm{B},i} = \frac{y_{\mathrm{B},G,i}P}{RT_i}, \tag{12}$$

$$G_i = \frac{(n_{\mathrm{A},i} + n_{\mathrm{B},i} + n_{\mathrm{inert},i})RT_i}{P}, \tag{13}$$

$$y_{\mathrm{A},i} = \frac{y_{\mathrm{A},G,i}}{G}, \tag{14}$$

$$y_{\mathrm{B},i} = \frac{y_{\mathrm{B},G,i}}{G}, \tag{15}$$

$$\frac{dy_{\mathrm{A},G,i}}{dz} = -ky_{\mathrm{A},i}\Omega, \tag{16}$$

$$\frac{dy_{\mathrm{B},G,i}}{dz} = mky_{\mathrm{A},i}\Omega, \tag{17}$$

with $i = 1,\ldots 39$ representing the measurement instances. Besides the given quantities R, Ω and n_{inert}, this algebraic model contains 7×39 equations and the same number of unknown variables ($n_{\mathrm{A},i}, n_{\mathrm{B},i}, G_i, y_{\mathrm{A},i}, y_{\mathrm{B},i}, \frac{dy_{\mathrm{A},G,i}}{dz}, \frac{dy_{\mathrm{B},G,i}}{dz}$). For given values of the parameters k, m and P, this system can therefore be solved. In other words, Eqs. (11)-(17) form a suitable subsystem and can be used in the subsequent steps without further modifications. Moreover, two lower-dimensional suitable subsystems can be obtained by dropping either Eq. (16) or Eq. (17). The various suitable subsystems available along with the parameters present in each subsystem are reported in Table 2.

Subsystem	Eqs.	Parameters
1	(11)-(17)	P,k,m
2	(11)-(16)	P,k
3	(11)-(15), (17)	P,k,m

Table 2: Available suitable subsystems and parameters therein

3.3 Step 3: Global Algebraic Parameter Estimation

Any suitable subsystem singled out in step 2 of the IGPE procedure gives rise to an algebraic parameter estimation problem that involves part or all of the unknown model parameters. Such NLP problems can be solved to global optimality using state-of-the-art software packages such as BARON [29]; (see also, Arnold Neumaier's global optimization page at http://www.mat.univie.ac.at/~neum/glopt/software_g.html for an up-to-date list of global optimization solvers).

In the eventuality that several suitable subsystems are available, one often has the option to solve a sequence of lower-dimensional estimation problems rather than one single large problem. The advantage of solving a sequence of lower-dimensional problems is mostly of computational nature, as it can decrease the overall effort tremendously for problems having many parameters. On the other hand, a lower-dimensional problem uses part of the available structural information and measurement data only, which may have an adverse effect on the accuracy of the resulting parameter estimates. As mentioned above, an algorithmic procedure automatically determining the most appropriate suitable subsystems is still lacking. However, as a general guideline, it is advisable to consider the largest possible estimation problem that can be solved in an acceptable amount of time.

Back to the case-study problem, three suitable subsystems have been identified in step 2 of the IGPE procedure. One possibility is to estimate parameters P and k based on subsystem 2 and then estimate m based on either subsystem 1 or 3. Another possibility is to estimate all three parameters at once by using either subsystem 1 or 3. Since the algebraic estimation problems are small and can be solved to global optimality fairly quickly, the latter approach is advisable here. In addition, subsystem 1 is preferred to subsystem 3 since it uses more of the available structural information.

It takes about 14 CPU-sec to get a global solution for the resulting estimation problem using BARON [29][3]. The resulting parameter estimates are reported in the first row of Table 3; the globally optimal parameter values for the simultaneous estimation problem are also given in the second row for comparison purpose (they are different from the values used to generate the data due to the presence of measurement noise). It can be seen that the parameter estimates are close to the true global solution, yet slightly biased. This is to be expected, since a bias is introduced in step 1 through the estimation of derivatives (see Figure 2).

[3] All computation times have been obtained using an Intel Core Duo 2.13 GHz PC with 2 GB of RAM, running Windows XP.

Parameter	P	k	m
Optimal estimate after step 3	1.0e5	1.357e1	2.694
Globally optimal value	1.006e5	1.496e1	2.007

Table 3: Parameter estimates after step 3 of the IGPE method

3.4 Step 4 - Dynamic Complement

In the eventuality that not all parameters of the model could be estimated in step 3 of the IGPE procedure, i.e. if no suitable subsystem could be isolated for one or more parameters, the remaining parameters must be determined simultaneously by solving a global dynamic optimization problem. Although this step may reduce the efficiency of the IGPE method significantly, observe that the dynamic complement problem contains fewer parameters than the original dynamic estimation problem, so the IGPE method remains advantageous in terms of computational effort.

Since all the parameters of the case study problem could be estimated in step 3, this step can be skipped here.

3.5 Step 5 - Simultaneous Correction

In this last step, all the model parameters are estimated simultaneously by solving the original dynamic estimation problem locally, starting from the incremental parameter estimates calculated in step 4 as the initial guess. Unlike the incremental parameter estimates that are inevitably biased due to error propagation from the estimated state derivatives, the simultaneous parameter estimates determined in this step are statistically optimal. Unfortunately, no guarantee can be given that the resulting simultaneous parameter estimates will correspond to a global solution of the problem (3) since only a local optimization is performed. However, our experience is that the IGPE procedure typically finds a global solution to the original problem when reasonably informative measurement data are available.

For the case-study problem, the simultaneous correction step is performed by using gEST, the parameter estimation routine provided by gPROMS. The computation takes approximately 4 CPU-sec and yields the globally optimal parameter values. Overall, it takes about 18 CPU-sec to find the global optimum to the case-study problem.

3.6 Discussion

Several remarks on the IGPE algorithm are in order:

- 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.

  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 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 .

- It is important to mention that situations exist, where simultaneous approaches can be applied while IGPE cannot. In particular, this may occur when few states are measured compared to the number of equations in the model. In principle, a simultaneous dynamic parameter estimation problem can be formulated and solved with just one data point available (even though the model parameters would likely be unidentifiable with this minimal information); on the other hand, obtaining a suitable subsystem in the IGPE approach usually requires that a few states at least be measured. Hence, the IGPE approach typically requires more measurement data than simultaneous approaches, but at the same time this extra information helps reveal situations where the parameters are unidentifiable.

4 Future Work

The presented case study shows that IGPE offers a high potential to efficiently identify the globally optimal parameter values of a DAE model. However, the IGPE method also has a number of limitations. First and foremost, no guarantee that a global optimum is been found can be given. Particularly, the use of low quality and scarce measurement data may lead to the global optimum being missed.

In future work the IGPE method will thus further be 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, almost equally well solutions would also be interesting from a model development point of view. 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 might not reveal it. In this case either more data or a less complex model structure should be considered.

Furthermore 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 every, or better the most appropriate, suitable subsystems is desirable and will be the focus of future research.

In addition, estimates of the errors on the estimated derivatives could be obtained at relatively low computational cost, using for instance bootstrapping methods [30, 31]. These error estimates could then be used to decide on the number of local optima that are to be retained in step 5. 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.

5 References

[1] L.T. Biegler, I.E. Grossmann, and A.W. Westerberg. *Systematic Methods of Chemical Process Design*. Prentice-Hall, Englewood Cliffs, NJ, 1997.

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

[3] I.-W. Kim, M. J. Liebman, and T. F. Edgar. A sequential error-in-variables method for nonlinear dynamic systems. *Comput. Chem. Eng.*, 15(9):663–670, 1991.

[4] Process Systems Enterprise Ltd. gPROMS advanced process modelling environment. http://www.psenterprise.com/gproms/, 2008.

[5] Numerica Technology LLC. JACOBIAN dynamic modeling and optimization software. http://www.numericatech.com/, 2008.

[6] B. Van Den Bosch and L. Hellinckx. A new method for the estimation of parameters in differential equations. *AIChE J*, 20(2):250–255, 1974.

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

[8] A. B. Singer, J. W. Taylor, P. I. Barton, and W. H. Green Jr. Global dynamic optimization for parameter estimation in chemical kinetics. *J Phys Chem A*, 110(3):971–976, 2006.

[9] C. Guus, E. Boender, and H. E. Romeijn. Stochastic methods. In R. Horst and P. M. Pardalos, editors, *Handbook of global optimization*, pages 829–869, Dordrecht, The Netherlands, 1995. Kluwer Academic Publishers.

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

[11] C. G. Moles, P. Mendes, and J. R. Banga. Parameter estimation in biochemical pathways: A comparison of global optimization methods. *Genome Res.*, 13:2467–2474, 2003.

[12] R. Horst and H. Tuy. *Global Optimization: Deterministic Approaches*. Springer-Verlag, Berlin, Germany, 3rd edition edition, 1996.

[13] William R. Esposito and Christodoulos A. Floudas. Global optimization for the parameter estimation of differential-algebraic systems. *Ind. Eng. Chem. Res.*, 39:1291–1310, 2000.

[14] I. Papamichail and C. S. Adjiman. Global optimization of dynamic systems. *Computers Chem. Eng.*, 28:403–415, 2004.

[15] B. Chachuat and M. A. Latifi. A new approach in deterministic global optimization of problems with ordinary differential equations. In C. A. Floudas and P. M. Pardalos, editors, *Frontiers in Global Optimization*, volume 74 of *Nonconvex Optimization and Its Applications*, pages 83–108. Kluwer Academic Publishers, 2003.

[16] B. Chachuat, A. B. Singer, and P. I. Barton. Global solution of dynamic optimization and mixed-integer dynamic optimization. *Ind. Eng. Chem. Res.*, 45(25):8373–8392, 2006.

[17] Y. Lin and M. A. Stadtherr. Deterministic global optimization of nonlinear dynamic systems. *AIChE J*, 53:866–875, 2007.

[18] L. H. Hosten. A comparative study of short-cut procedures for parameter estimation in differential equations. *Comput. Chem. Eng.*, 3:117–126, 1979.
[19] A. Bardow and W. Marquardt. Identification of diffusive transport buy means of an incremental approach. *Computers Chem. Eng.*, 28(5):585–595, 2004.
[20] A. Bardow and W. Marquardt. Incremental and simultaneous identification of reaction kinetics: Methods and comparison. *Chem. Eng. Sci.*, 59(13):2673–2684, 2004.
[21] M. Brendel, D. Bonvin, and W. Marquardt. Incremental identification of complex reaction kinetics in homogeneous systems. *Chemical Engineering Science*, 61:5404–5420, 2006.
[22] G. F. Froment and K. B. Bischoff. *Chemical Reactor Analysis and Design*. Wiley, New York, NY, 1990.
[23] M. Tawarmalani and N. Sahinidis. Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming: Theory, Algorithms, Software and Applications. 2005.
[24] John Ingham. *Chemical Engineering Dynamics*. VCH, 1997.
[25] Schittkowski. *Numerical Data Fitting in Dynamical Systems*. Kluwer, Academic Publishers, Dordrecht, The Netherlands, 2002.
[26] J. Hadamard. Lectures on cauchy problem in linear partial differential equations. Oxford University Press, London, 1923.
[27] A. N. Tikhonov and V. A. Arsenin. *Solution of Ill-posed Problems*. Winston and Sons, Washington, 1977.
[28] Christian H. Reinsch. Smoothing by spline functions. *Numer. Math.*, 10:177–183, 1967.
[29] N. V. Sahinidis. BARON: A general purpose global optimization software package. *Journal of Global Optimization*, 8(2):201–205, 1996.
[30] Carl Duchesne and John F. MacGregor. Jackknife and bootstrap methods in the identification of dynamic models. *Journal of Process Control*, 11:553–564, 2001.
[31] Ron Wehrens, Hein Putter, and Lutgarde M. C. Buydens. The bootstrap: a tutorial. *Chemometrics and Intelligent Laboratory Systems*, 54:35–52, 2000.