

Guaranteed Parameter Estimation of Nonlinear Dynamic Systems using High-Order Bounding Techniques with Domain and CPU-Time Reduction Strategies

RADOSLAV PAULEN

Process Dynamics and Operations Group, Department of Biochemical and Chemical Engineering, Technische Universität Dortmund, Emil-Figge-Str. 70, 44221 Dortmund, Germany.

MARIO E. VILLANUEVA, BENOÎT CHACHUAT

Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, UK.

[Received on December 16, 2014]

This paper is concerned with guaranteed parameter estimation of nonlinear dynamic systems in a context of bounded measurement error. The problem consists of finding—or approximating as closely as possible—the set of all possible parameter values such that the predicted values of certain outputs match their corresponding measurements within prescribed error bounds. A set-inversion algorithm is applied, whereby the parameter set is successively partitioned into smaller boxes and exclusion tests are performed to eliminate some of these boxes, until a given threshold on the approximation level is met. Such exclusion tests rely on the ability to bound the solution set of the dynamic system for a finite parameter subset, and the tightness of these bounds is therefore paramount; equally important in practice is the time required to compute the bounds, thereby defining a trade-off. In this paper, we investigate such a trade-off by comparing various bounding techniques based on Taylor models with either interval or ellipsoidal bounds as their remainder terms. We also investigate the use of optimization-based domain reduction techniques in order to enhance the convergence speed of the set-inversion algorithm, and we implement simple strategies that avoid recomputing Taylor models or reduce their expansion orders wherever possible. Case studies of various complexities are presented, which show that these improvements using Taylor-based bounding techniques can significantly reduce the computational burden, both in terms of iteration count and CPU time.

Keywords: parameter estimation, dynamic systems, bounded-error estimation, measurement noise, Taylor models, polyhedral relaxations, domain reduction

1. Introduction

Mathematical modelling has become an integral part of modern process design methodologies as well as in control system design and operations optimization. A typical model development procedure is divided into two main phases, namely specification of the model structure and estimation of the unknown/uncertain model parameters. The latter phase, often referred to as model fitting, normally proceeds by determining parameter values for which the model predictions closely match the available process measurements. Failure to find an acceptable agreement calls for a revision of the model structure, before repeating the parameter estimation.

Most commonly, the parameter estimation problem is posed as an optimization problem that determines the parameter values minimizing the gap between the measurements and the model predictions,

for instance in the least-square sense. Nonetheless, several factors can impair a successful and reliable estimation procedure. First of all, structural model mismatch is inherent to the modeling exercise, and it is illusive to look for the ‘true’ parameter values in this context. Even in the absence of model mismatch, fitting a set of experimental data exactly is generally not possible due to various sources of uncertainty. A measurement’s accuracy is always tied to the resolution of the corresponding apparatus. Moreover, measured data are typically corrupted with noise, for instance Gaussian white noise or more generally colored noise.

Among the available approaches to account for uncertainty in parameter estimation, the focus in this paper is on *guaranteed* parameter estimation (Walter, 1990), namely the determination of *all* parameter values—referred to as the solution set subsequently—that are consistent with the measurements under given uncertainty scenarios. Specifically, we consider the case that the uncertainty enters the estimation problem in the form of bounded measurement errors. An inherent advantage of this approach over more traditional parameter estimation is that no (consistent) solution to the problem will be lost, and this can help detect problems arising due to lack of identifiability. Moreover, the estimation process does not rely on a particular statistical description of the uncertainty, as is typically the case when applying maximum likelihood or Bayesian techniques. On the downside nonetheless, performing guaranteed parameter estimation turns out to be a very challenging and demanding task from a computational standpoint.

In nonlinear algebraic models, the problem of approximating the solution set by a box partition, at an arbitrary precision, has been shown to be tractable using exhaustive search and interval analysis (Moore, 1992), for instance based on the set-inversion algorithm SIVIA (Jaulin and Walter, 1993). This approach has been later extended to dynamic systems using ODE bounding techniques (e.g., Jaulin, 2002; Raissi et al., 2004). In a recent paper, Kieffer and Walter (2011) have identified the main computational bottlenecks of set-inversion algorithms for guaranteed parameter estimation in dynamic systems to be: (i) the need for tight bounds on the solutions of the dynamic system; and, (ii) the need for efficient domain-reduction strategies as part of the exclusion tests. It is the objective of this paper to investigate strategies that can enhance the convergence speed of these algorithms, with special emphasis on higher-order ODE bounding techniques and domain-reduction techniques.

The computation of exact bounds on the solution set of nonlinear parametric ODEs belongs to the class of computationally intensive problems (nonconvex optimization). In response to this, approximate methods that overestimate the solution set of parametric ODEs, yet provide sufficiently tight bounds, have been developed over the years. These methods can be classified as discrete or continuous methods according to the way the enclosures are propagated through time. *Discrete-time methods* proceed by discretizing the integration horizon into a finite number of steps, whereby each step consists of two phases. Phase I is concerned with the computation of a coarse enclosure and a step-size for which existence and uniqueness of the solutions can be established; then, the enclosure is refined at the end of each time-step during Phase II. Recently, Houska et al. (2013) also proposed an algorithm reverting the order of the two phases. The types of enclosures that can be propagated with discrete-time methods include intervals (Lohner, 1992; Nedialkov et al., 1999; Rauh et al., 2006), Taylor models with interval remainders (Berz and Makino, 1998; Neher et al., 2007; Lin and Stadtherr, 2007b; Sahlodin and Chachuat, 2011) and Taylor models with ellipsoidal remainders (Houska et al., 2013). *Continuous-time methods*, on the other hand, involve formulating a set of (parameter-independent) auxiliary ODEs, whose solutions enclose those of the original dynamic model. Similar to discrete-time approaches, the types of enclosures that can be propagated in continuous time include intervals based on the classical theory of differential inequalities (Walter, 1970) as well as ellipsoids using ellipsoidal calculus (Kurzanski and Varaiya, 2002; Houska et al., 2012). Extensions of these approaches have recently been proposed to enable the propagation of Taylor models with either interval remainders (Chachuat and Villanueva,

2012) or ellipsoidal remainders (Villanueva et al., 2013). See also (Villanueva et al., 2014) for a unified framework and convergence analysis of continuous-time bounding techniques for nonlinear parametric ODEs.

In the context of guaranteed parameter estimation, the use of ODE bounding techniques based on Taylor models has been investigated by Lin and Stadtherr (2007a) and Kletting et al. (2011) using discrete-time bounding techniques and, more recently, by Paulen et al. (2013) using a continuous-time approach. These authors have reported significant improvements in the convergence speed of the set-inversion algorithm compared to classical approaches based on interval enclosures. In principle, the higher the Taylor expansion order of the ODE solutions with respect to the uncertain parameters, the smaller the number of iterations required by the set-inversion algorithm to converge. Nonetheless, a higher-order expansion can incur a significant computational overhead, thereby defining a trade-off in terms of the overall computational burden with regards to the expansion order. This trade-off is investigated further in this paper through the comparison of continuous-time methods propagating Taylor models (with either interval or ellipsoidal remainders) against simple interval box propagation.

Another approach to enhancing the convergence involves applying contractors to the parameter boxes in order to reduce their width. Contractors based on optimality tests were derived by Jaulin et al. (2001) using interval analysis, and later applied to dynamic system, e.g., by Kieffer and Walter (2011). Besides enabling higher-order convergence, the use of Taylor models to enclose the ODE solutions provides an explicit representation of parameter dependencies via the multivariate polynomial part. Lin and Stadtherr (2007a) and Kletting et al. (2011) took advantage of this representation and used a constraint-propagation strategy in order to contract the parameter boxes. Inspired by developments in the field of global optimization (Zamora and Grossmann, 1999; Neumaier, 2004; Tawarmalani and Sahinidis, 2004; Belotti et al., 2009) and their recent extension to global dynamic optimization (Sahlodin, 2012), this paper investigates a domain-reduction technique that solves linear programs (LPs) constructed from the polyhedral relaxation of Taylor models of the predicted outputs as a means to exclude those parameter subsets whose corresponding response does not intersect with the measurement bounds. In order to further reduce the computational burden, we also investigate new strategies that avoid recomputing Taylor models of the predicted outputs or reduce their order as soon as the corresponding overestimation is within a given threshold.

The rest of the paper is organized as follows. In Sect. 2, the problem of guaranteed parameter estimation is defined mathematically and the set-inversion algorithm is concisely stated. Sect. 3 presents a new methodology for enhancing the convergence of set inversion in a guaranteed parameter estimation context, which relies on Taylor-model bounding in combination with domain-reduction and CPU-time-reduction strategies; a simple case study is carried out through this section to illustrate the developments. Then, Sect. 4 presents the case study of a more challenging model of anaerobic digestion with complex dynamics and multiple time-scales, demonstrating that the proposed improvements allow tackling guaranteed parameter estimation in up to seven parameters within reasonable computational times. Finally, Sect. 5 concludes the paper.

2. Guaranteed Parameter Estimation

2.1 Problem Statement

Consider a dynamic process described by parametric ODEs of the form

$$\dot{x}(t, p) = f(x(t, p), p) \quad \text{with} \quad x(0, p) = h(p), \quad (2.1a)$$

$$y(t, p) = g(x(t, p), p), \quad (2.1b)$$

where $x : [0, t_N] \times \mathbb{R}^{n_p} \rightarrow \mathbb{R}^{n_x}$ denotes the vector of process states, $p \in \mathbb{R}^{n_p}$ stands for the vector of (unknown) process parameters, and $y : [0, t_N] \times \mathbb{R}^{n_p} \rightarrow \mathbb{R}^{n_y}$ denotes the n_y -dimensional vector of model outputs (predictions). Notice the parametric dependencies in the right-hand side function f , the output function g , and the initial value function h . In particular, this latter dependency can be used to handle dynamic systems with uncertain initial conditions. Given a bounded subset $P \subseteq \mathbb{R}^{n_p}$ for the parameters, we introduce the point-wise-in-time reachable sets of (2.1a) and (2.1b) as

$$\forall t \in [0, t_N], \quad X(t, P) := \{x(t, p) \mid p \in P \subseteq \mathbb{R}^{n_p}\} \quad \text{and} \quad Y(t, P) := \{y(t, p) \mid p \in P \subseteq \mathbb{R}^{n_p}\}. \quad (2.2)$$

For a given set of output measurements $y_m(t_i)$ at N time points $t_1, \dots, t_i, \dots, t_N$, *classical* parameter estimation seeks for *one* particular instance p_e of the parameter values for which the (possibly weighted) normed difference between these measurements and the corresponding model outputs y is minimized. This optimization problem, for instance in the least-square sense, is given by:

$$p_e \in \arg \min_{p \in P_0} \sum_{i=1}^N \|y_m(t_i) - y(t_i, p)\|_2^2, \quad (2.3a)$$

$$\text{s.t. } \dot{x}(t, p) = f(x(t, p), p) \quad \text{with} \quad x(0, p) = h(p), \quad (2.3b)$$

$$y(t, p) = g(x(t, p), p), \quad (2.3c)$$

where the interval box $P_0 := [p_0^L, p_0^U]$ denotes the a priori set of admissible values for the parameters. The superscripts L and U representing the lower and upper bounds of an interval box are understood component-wise throughout.

In contrast, *guaranteed* (bounded-error) parameter estimation accounts for the fact that the actual process outputs, y_p , are only known within some bounded measurement error $e \in E := [e^L, e^U]$, so that

$$y_p(t_i) \in y_m(t_i) + [e^L, e^U] =: Y_p(t_i). \quad (2.4)$$

Then, the main objective is to estimate the set P_e of *all* possible parameter values p such that $y(t_i; p) \in Y_p(t_i)$ for every $i = 1, \dots, N$; that is,

$$P_e := \left\{ p \in P_0 \left| \begin{array}{l} \exists x \text{ such that:} \\ \dot{x}(t, p) = f(x(t, p), p) \quad \text{with} \quad x(0, p) = h(p), \\ g(x(t_i, p), p) \in Y_p(t_i), \quad i = 1, \dots, N \end{array} \right. \right\}. \quad (2.5)$$

Depicted in red on the left plot in Fig. 1 is the set of all output trajectories satisfying $y(t_i, p) \in Y_p(t_i)$ with $i = 1, \dots, N$, and on the right plot the corresponding set P_e projected onto the (p_1, p_2) space. Obtaining an exact characterization of the set P_e is not possible in general, and one has to resort to approximation techniques that make the problem computationally tractable. The focus in the remainder of the paper is on algorithms approximating P_e using set-inversion techniques.

2.2 Set-Inversion Algorithm

We consider a variant of the Set Inversion Via Interval Analysis (SIVIA) algorithm by (Jaulin and Walter, 1993) in order to approximate the solution set P_e to a desired accuracy. This algorithm has already been exploited in a number of papers in the context of dynamic parameter estimation (e.g., Jaulin, 2002; Raissi et al., 2004; Lin and Stadtherr, 2007a; Kieffer and Walter, 2011; Kletting et al., 2011; Paulen et al., 2013).

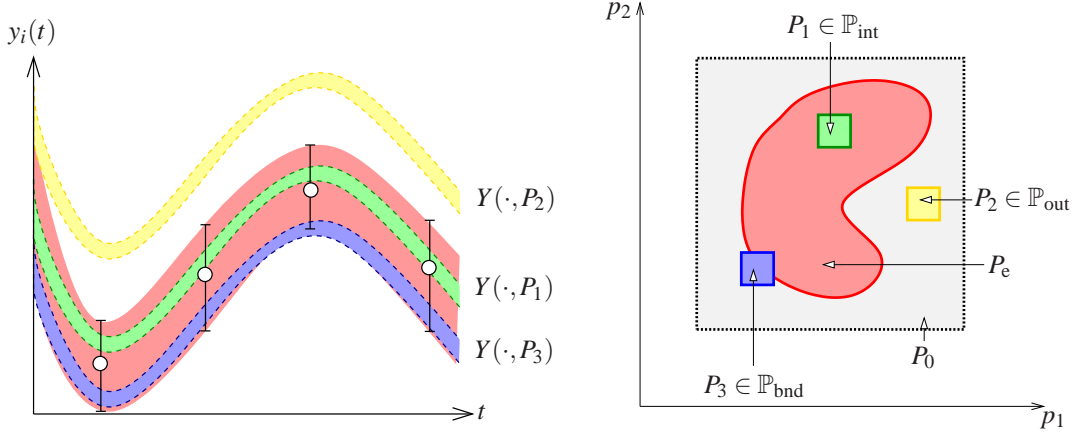


FIG. 1. Illustration of guaranteed parameter estimation concepts in the space of output trajectories (left plot) and in the parameter space (right plot).

Let $Y^{-1}(t, \cdot)$, with $Y^{-1} : [0, t_N] \times \Pi(R^{n_y}) \rightarrow \Pi(R^{n_p})$, denote the inverse of the reachable set mapping $Y(t, \cdot)$ defined earlier in (2.2). It follows that characterizing P_e via (2.5) is equivalent to intersecting the inverse image sets $Y^{-1}(t_i, Y_p(t_i))$ for each $i = 1, \dots, N$:

$$P_e = \left(\bigcap_{i=1}^N Y^{-1}(t_i, Y_p(t_i)) \right) \cap P_0. \quad (2.6)$$

A prototypical set-inversion algorithm based on exhaustive search that uses this property is as follows:

Input: Termination tolerances $\varepsilon_{\text{box}} \geq 0$ and $\varepsilon_{\text{bnd}} \geq 0$

Initialization: Set partitions $\mathbb{P}_{\text{bnd}} = \{P_0\}$, $\mathbb{P}_{\text{int}} = \emptyset$, and $\mathbb{P}_{\text{out}} = \emptyset$; Set iteration counter $k = 0$

Main Loop:

1. Select a parameter box P in the partition \mathbb{P}_{bnd} and remove it from \mathbb{P}_{bnd}
2. Compute enclosures $\bar{Y}(t_i, P) \supseteq Y(t_i, P)$, for each $i = 1, \dots, N$
3. Exclusion Tests:
 - (a) **If** $\bar{Y}(t_i, P) \subseteq Y_p(t_i)$ for all $i \in \{1, \dots, N\}$, insert P into \mathbb{P}_{int}
 - (b) **Else if** $\bar{Y}(t_i, P) \cap Y_p(t_i) = \emptyset$ for some $i \in \{1, \dots, N\}$, insert P into \mathbb{P}_{out}
 - (c) **Else** bisect P and insert subsets back into \mathbb{P}_{bnd}
4. Termination Tests:
 - (a) **If** $V_{\text{bnd}} := \sum_{P \in \mathbb{P}_{\text{bnd}}} \text{volume}(P) \leq \varepsilon_{\text{bnd}}$, **stop**
 - (b) **If** $\text{width}(P) \leq \varepsilon_{\text{box}}$ for all $P \in \mathbb{P}_{\text{bnd}}$, **stop**
5. Increment counter $k += 1$; **Return** to step 1

Output: Partitions \mathbb{P}_{int} , \mathbb{P}_{bnd} , and \mathbb{P}_{out} ; Iteration count k

An illustration of a parameter box belonging to the partition \mathbb{P}_{int} , \mathbb{P}_{bnd} or \mathbb{P}_{out} is shown on the right plot in Fig. 1, together with the corresponding output trajectories on the left plot using a consistent color scheme. Upon termination, this algorithm returns partitions \mathbb{P}_{int} and \mathbb{P}_{bnd} such that

$$\bigcup_{P \in \mathbb{P}_{\text{int}}} P \subseteq P_e \subseteq \bigcup_{P \in \mathbb{P}_{\text{int}} \cup \mathbb{P}_{\text{bnd}}} P. \quad (2.7)$$

The following remarks are in order regarding the set-inversion algorithm:

- Multiple heuristics can be used regarding the selection and the bisection of a parameter box in steps 1 and 3c, respectively. In order for the search to be exhaustive, one can select a parameter box that has the largest width in priority and apply bisection at the mid-point along the least reduced axis of a box for instance.
- Step 2 calls for a procedure capable of computing an enclosure of the output reachable set $Y(\cdot, P)$ for the current parameter box P . The main difficulty of this bounding step lies in the computation of a point-wise-in-time enclosure $\bar{X}(t, P)$ of the state reachable set $X(t, P)$, after which an enclosure $\bar{Y}(t, P) \supseteq Y(t, P)$ can be computed readily by using standard interval analysis (Jaulin et al., 2001; Moore et al., 2009). For instance, Jaulin (2002) first used the theory of differential inequalities which provides a rule for propagating an interval ODE enclosure $\bar{X}(t, P) := [x^L(t), x^U(t)] \in \mathbb{IR}^{n_x}$ of the reachable set $X(t, P)$ in the form of auxiliary ODEs:

$$\dot{x}_i^L(t, P) = \min_{\xi, \rho} \left\{ f_i(\xi, \rho) \left| \begin{array}{l} \xi_i = x_i^L(t) \\ \xi \in [x^L(t), x^U(t)] \\ \rho \in P \end{array} \right. \right\} \quad \text{with} \quad x_i^L(0, P) = \min_{\rho} \{h(\rho) \mid \rho \in P\}, \quad (2.8)$$

$$\dot{x}_i^U(t, P) = \max_{\xi, \rho} \left\{ f_i(\xi, \rho) \left| \begin{array}{l} \xi_i = x_i^U(t) \\ \xi \in [x^L(t), x^U(t)] \\ \rho \in P \end{array} \right. \right\} \quad \text{with} \quad x_i^U(0, P) = \max_{\rho} \{h(\rho) \mid \rho \in P\}, \quad (2.9)$$

for each $i \in \{1, \dots, n_x\}$. In principle, any ODE bounding technique can be used for this step as long as the computed output enclosures $\bar{Y}(t_i, P)$ shrink when the diameter of the parameter host set $\text{diam}(P) := \max_{x, y \in P} |x - y| \rightarrow 0$ in order to guarantee finite termination of the algorithm. This is the case in particular for the higher-order bounding techniques considered in Sect. 3.1.

- Test 4a is an addition to the original SIVIA algorithm (Jaulin and Walter, 1993), which interrupts the iterations when a specified level of approximation of the solution set P_e is reached. The level of approximation is measured here as the total volume V_{bnd} of the boxes in the partition, with corresponding threshold ε_{bnd} . In contrast, stopping the algorithm when a minimum width is reached for all the boxes in \mathbb{P}_{bnd} (Test 4b) does not give any guarantee on the actual approximation level of the solution set boundary because of the overestimation in step 2. We also note that finite termination of the algorithm requires that either $\varepsilon_{\text{box}} > 0$ or $\varepsilon_{\text{bnd}} > 0$.

Variants of this basic algorithm exist that improve the convergence speed by introducing additional exclusion tests. One such test involves checking whether (an enclosure of) the gradient of the objective function in (2.3a) for a given box P does not contain 0, in which case $P \in \mathbb{P}_{\text{out}}$ (Kieffer and Walter, 2011)—this is because there cannot exist any global optimizer of (2.3a) in P in this case. The downside of these strategies is the need to compute bounds on the first-order sensitivities (or adjoints) of

model (2.1), which can cause a significant computational overhead. Reduction of the parameter boxes was also investigated, e.g., by Lin and Stadtherr (2007a) and Kletting et al. (2011) using constraint propagation on Taylor models of the model outputs. The following section describes a new methodology to enhancing the convergence of set-inversion in a guaranteed parameter estimation context and be in a position to tackle more challenging, larger-scale problems.

3. Guaranteed Parameter Estimation Methodology using Taylor Models

Kieffer and Walter (2011) have argued that the main computational bottlenecks of the set-inversion algorithm in Sect. 2.2 for guaranteed parameter estimation are the need for tight bounds on the solutions of the dynamic system as well as efficient domain-reduction strategies supporting exclusion tests. The methodology developed through this section aims precisely at addressing these needs. It relies on higher-order techniques based on Taylor models to bound the dynamics (Sect. 3.1), and then takes full advantage of the resulting Taylor model estimators for driving optimization-based domain reduction (Sect. 3.2). Moreover, special strategies are developed that avoid recomputing Taylor models or reduce their expansion orders wherever possible.

3.1 Higher-Order Bounding Strategy

The reachable set of a nonlinear ODE is a nonconvex set in general, and enclosing it within an interval box can lead to significant overestimation due to both the wrapping effect and the dependency problem. One way of propagating nonconvex enclosures $\bar{X}(t, P)$ of the reachable set $X(t, P)$ involves using a q th-order multivariate polynomial $\mathcal{P}_x^q(t, \cdot)$, whose image set on P approximates $X(t, P)$, as

$$\bar{X}(t, P) := \{ \mathcal{P}_x^q(t, p) \mid p \in P \} \oplus \mathcal{R}_x^q(t, P), \quad (3.1)$$

where $\mathcal{R}_x^q(t, P) \supseteq \{x(t, p) - \mathcal{P}_x^q(t, p) \mid p \in P\}$ is the so-called remainder term bounding the approximation error on P . In turn, an interval enclosure can be derived from (3.1) by bounding the polynomial part, for which multiple approaches have been proposed in the literature (Lin and Rokne, 1995; Neumaier, 2002). For simplicity, the approach used in this work considers exact bounding of the linear and diagonal quadratic terms, while overestimating the remaining terms using natural interval extensions (Lin and Stadtherr, 2007b).

The focus hereafter is on Taylor models (Makino and Berz, 1999; Neumaier, 2002; Bompadre et al., 2013), although alternative types of polynomial approximation can be used in principle as long as these constructions can be automated for general factorable functions. In this approach, the polynomial approximant $\mathcal{P}_x^q(t, \cdot)$ matches the q th-order Taylor expansion of $x(t, \cdot)$ on P at a given reference point $\hat{p} \in P$:

$$\forall p \in P, \quad \mathcal{P}_x^q(t, p) := \sum_{\substack{\gamma \in \mathbb{N}^{n_p} \\ |\gamma| \leq q}} \frac{\partial^\gamma x(t, \hat{p})}{\gamma!} (p - \hat{p})^\gamma, \quad (3.2)$$

where multi-index notation is used and $\partial^\gamma x_i(\cdot, \hat{p})$ denotes the state sensitivities $\frac{\partial^{|\gamma|} x_i}{\partial p_1^{\gamma_1} \dots \partial p_{n_p}^{\gamma_{n_p}}}(\cdot, \hat{p})$ at \hat{p} . Note that this construction requires that the right-hand side function f and initial-value function h are at least $(q+1)$ -times continuously-differentiable in all their arguments. Moreover, it requires that a system of state-sensitivity equations of size $O(n_x n_p^q)$ is integrated on the time horizon.

Of the alternatives to compute the pointwise-in-time convex remainder enclosure $\mathcal{R}_x^q(t, P)$, our focus in this paper is on interval and ellipsoidal enclosures, which are summarized below.

TAYLOR MODELS WITH INTERVAL REMAINDER BOUNDS. The method of differential inequalities can be applied to propagate an interval enclosure $\mathcal{R}_x^q(t, P) := [r_x^L(t), r_x^U(t)]$ of the q th-order remainder term by integrating, together with the sensitivity equations for $\partial^\gamma x_i(\cdot, \hat{p})$ with $|\gamma| \leq q$, the following auxiliary ODEs:

$$r_i^L(t) = \min_{\xi, \rho} \left\{ f_i(\mathcal{P}_x^q(t, \rho) + \xi, \rho) - \dot{\mathcal{P}}_x^q(t, \rho) \left| \begin{array}{l} \xi_i = r_{x_i}^{q,L}(t) \\ \xi \in [r_x^L(t), r_x^U(t)] \\ \rho \in P \end{array} \right. \right\} \quad (3.3)$$

with $r_i^L(0) = \min_p \{h(p) - \mathcal{P}_h^q(p) \mid p \in P\}$,

$$r_i^U(t) = \max_{\xi, \rho} \left\{ f_i(\mathcal{P}_x^q(t, \rho) + \xi, \rho) - \dot{\mathcal{P}}_x^q(t, \rho) \left| \begin{array}{l} \xi_i = r_{x_i}^{q,U}(t) \\ \xi \in [r_x^L(t), r_x^U(t)] \\ \rho \in P \end{array} \right. \right\} \quad (3.4)$$

with $r_i^U(0) = \max_p \{h(p) - \mathcal{P}_h^q(p) \mid p \in P\}$,

for each $i \in \{1, \dots, n_x\}$, with \mathcal{P}_h^q denoting the multivariate polynomial in the Taylor expansion of the initial value function h on P at \hat{p} . The resulting enclosures (3.1) enjoy $(q+1)$ th-order convergence to the actual reachable set $X(t, P)$, but the size of the auxiliary bounding system scales as $O(n_x n_p^q)$.

TAYLOR MODELS WITH ELLIPSOIDAL REMAINDER BOUNDS. Likewise, ellipsoidal calculus provides a means of propagating an ellipsoidal enclosure $\mathcal{E}(Q_x^q(t)) := \{Q_x^q(t)^{\frac{1}{2}} v \mid \forall v \in \mathbb{R}^{n_x} : v^T v \leq 1\}$ of the q th-order remainder term by integrating, together with the parametric sensitivity equations up to order q , the following auxiliary ODEs:

$$\begin{aligned} \dot{Q}_x^q(t) &= \left(\frac{\partial f}{\partial x}(\mathcal{P}_x^q(t, \hat{p}), \hat{p}) \right) Q_x^q(t) + Q_x^q(t) \left(\frac{\partial f}{\partial x}(\mathcal{P}_x^q(t, \hat{p}), \hat{p}) \right)^T \\ &+ \sum_{i=1}^{n_x} \kappa_i(t) Q_x^q(t) + \text{diag}(\kappa(t))^{-1} \text{diagrad}(\Omega_f^q[Q_x^q(t), P, \hat{p}])^2 \quad \text{with } Q_x^q(0) = \text{diagrad}(\Omega_h^q[P, \hat{p}])^2. \end{aligned} \quad (3.5)$$

The nonlinearity bounders $\Omega_f^q[Q(t), P, \hat{p}], \Omega_h^q[P, \hat{p}] \in \mathbb{I}\mathbb{R}^{n_x}$ must satisfy

$$\forall (r, \rho) \in \mathcal{E}(Q) \times P, \quad f(\mathcal{P}_x^q(t, \rho) + r, \rho) - \dot{\mathcal{P}}_x^q(t, \rho) - \frac{\partial f}{\partial x}(\mathcal{P}_x^q(t, \rho), \rho) r \in \Omega_f^q[Q, P, \hat{p}], \quad (3.6)$$

and

$$\forall \rho \in P, \quad h(\rho) - \mathcal{P}_h^q(\rho) \in \Omega_h^q[P, \hat{p}], \quad (3.7)$$

and they can be constructed, at a given time t , on application of interval analysis for instance. Moreover, the scaling function κ can be chosen in such a way as to minimize $\text{tr}(Q_x^q(t))$. The resulting enclosures $\{\mathcal{P}_x^q(t, p) \mid p \in P\} \oplus \mathcal{E}(Q_x^q(t))$ enjoy $(q+1)$ th-order convergence to the actual reachable set $X(t, P)$, now at the price of solving an auxiliary bounding system of size $O(n_x n_p^q + n_x^2)$.

The reader is referred to (Villanueva et al., 2014) for more details about the theory and implementation of these methods. A comparison in the context of guaranteed parameter estimation is presented next for a simple case study.

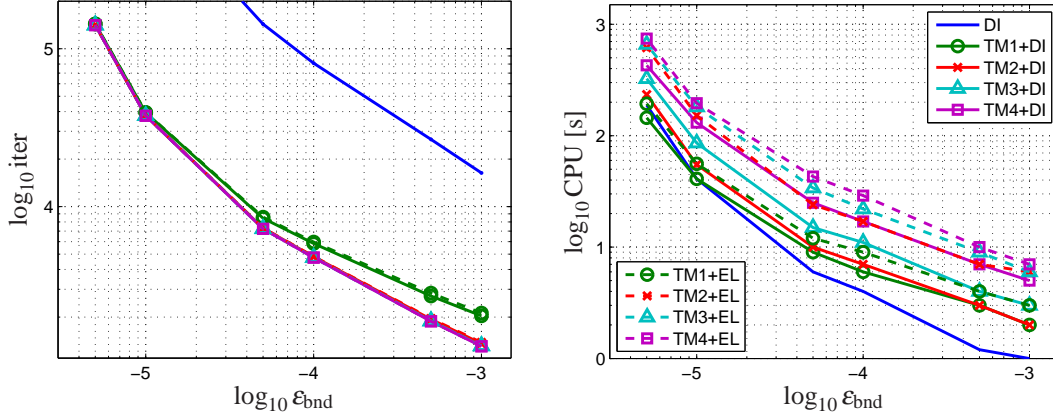


FIG. 2. Performance of guaranteed parameter estimation using various ODE bounding techniques in the set-inversion algorithm. Left: Number of iterations vs. convergence threshold. Right: CPU time vs. convergence threshold.

CASE STUDY. Consider the following dynamic model involving two state variables $x = (x_1, x_2)^T$ and three uncertain parameters $p = (p_1, p_2, p_3)^T \in [0.01, 1]^3$ (Kieffer and Walter, 2011):

$$\dot{x}_1(t) = -(p_1 + p_3)x_1(t) + p_2x_2(t) \quad \text{with } x_1(0) = 1, \quad (3.8a)$$

$$\dot{x}_2(t) = p_1x_1(t) - p_2x_2(t) \quad \text{with } x_2(0) = 0. \quad (3.8b)$$

This system has a single output variable y , which corresponds to the state variable x_2 , $y(t, p) := x_2(t, p)$, with $N = 15$ measurements corresponding to the time instants $t_i = 1, \dots, 15$. Synthetic experimental data are generated by simulating the model (3.8) with parameter values $p^* = (0.6, 0.15, 0.35)^T$, and then rounding the output $y(t_i)$ up or down to the nearest value by retaining two significant digits only; then, measurement error ranges of $\pm 5 \times 10^{-3}$ are added around these values.

The guaranteed parameter estimation algorithm (Sect. 2.2) is implemented in a C++ program that uses the library MC++ (<http://projects.coin-or.org/MC++>) for computations involving Taylor models. Moreover, the code calls the ODE integration methods in the GNU Scientific Library (GSL) to bound the parametric ODEs based on the techniques outlined in Sect. 3.1. All the numerical results presented subsequently use the explicit embedded Runge-Kutta-Fehlberg (4,5) method, with both relative and absolute tolerances set to 10^{-7} , and are obtained on a workstation with Intel Core i7-3770 processors at 3.40 GHz and running 64-bit Linux.

The performance of guaranteed parameter estimation is investigated for continuous-time ODE bounding techniques propagating Taylor models of orders $q = 1, \dots, 4$ with interval or ellipsoidal remainders (Sect. 3.1) and compared to standard differential inequalities. In order to allow for fair comparisons, the termination criterion is defined in terms of the level of accuracy ε_{bnd} of the solution set (Test 4a) in the range $10^{-3} \rightarrow 5 \times 10^{-6}$ —the termination criterion in terms of the minimum box size ε_{box} (Test 4b) is set to zero, on the other hand. The results are shown in Fig. 2 in terms of the number of iterations (left plot) and CPU time (right plot).

It is evident that classical differential inequalities (DI) require by far the largest number of iterations, at any accuracy level. For accuracies of $\varepsilon_{\text{bnd}} = 10^{-5}$ and $\varepsilon_{\text{bnd}} = 5 \times 10^{-6}$, respectively 932, 454 and 3,612,968 iterations are needed. Memory storage of such a high number of parameter boxes during the course of the algorithm can become a serious issue with an increasing number of uncertain parameters,

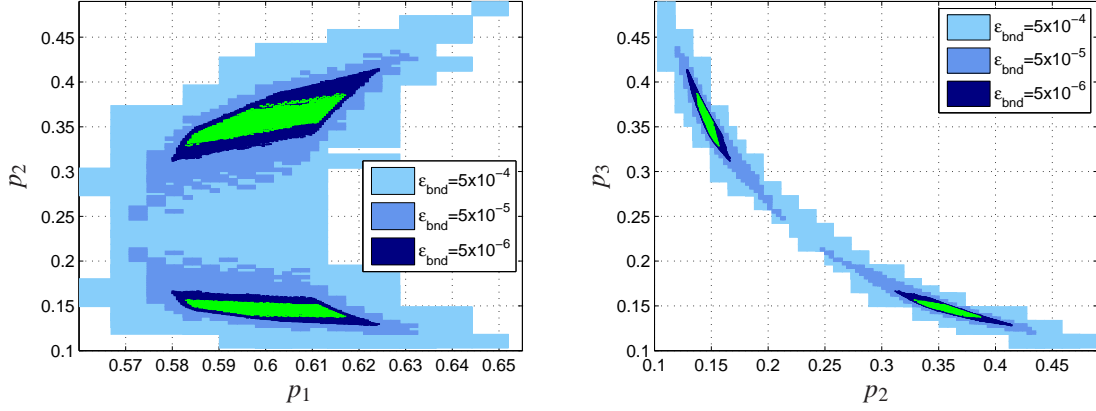


FIG. 3. Outer approximations of the sets of guaranteed parameter estimates for different levels of accuracy ε_{bnd} . Inner approximation of the set of guaranteed parameter estimates for $\varepsilon_{\text{bnd}} = 5 \times 10^{-6}$ plotted in green. Left: Projections onto (p_1, p_2) space. Right: Projections onto (p_2, p_3) space.

calling for less conservative bounding techniques. Despite the large number of iterations however, this approach allows for the fastest computations for accuracies down to $\varepsilon_{\text{bnd}} \approx 10^{-5}$ due to its simplicity. At higher accuracy levels, bounding techniques based on Taylor models are seen to exhibit faster convergence as the extra computational burden of these higher-order bounding techniques is overpowered by a dramatic reduction in overall number of iterations (more than an order of magnitude). The shortest run-time is obtained with first-order Taylor model with interval remainder bounds (labelled TM1+DI in Fig. 2) for $\varepsilon_{\text{bnd}} < 10^{-5}$ here.

In terms of overall number of iterations, the performance of the set-inversion algorithm between first-order Taylor models (both variants TM1+DI and TM1+EL), on the one hand, and between all Taylor models of second-, third- and fourth-order (both variants TM2+DI, TM3+DI, TM4+DI and TM2+EL, TM3+EL, TM4+EL), on the other hand, is about the same. The lower performance of first-order Taylor models compared to higher-order Taylor models can be attributed to the fact that first-order Taylor models compute convex enclosures and are thus limited for the approximation of (potentially) nonconvex reachable sets. In terms of the overall run-time though, first-order Taylor models with interval remainders are found to outperform the other bounding techniques based on higher-order Taylor models in this case study. Bearing in mind the trade-off between a smaller number of iterations and a larger processing time needed for a single iteration, it is expected that higher-order bounding techniques will become advantageous for dynamic models of higher complexity or with more uncertain parameters nonetheless.

Finally, the left and right plots in Fig. 3 show projections of the approximate solution sets— \mathbb{P}_{int} and \mathbb{P}_{bnd} are shown using the same color scheme as in Fig. 1 above—onto the (p_1, p_2) and (p_2, p_3) subspaces, respectively, for different levels of accuracy ε_{bnd} . Observe first that the ‘true’ parameter values p^* used to generate the pseudo-experimental data are part of the solution set, for each reported level ε_{bnd} . Moreover, the solution set for this problem turns out to be disconnected, thus suggesting a possible structural identifiability problem. Interestingly, this non-connectedness of the solution set can only be detected when the accuracy level ε_{bnd} is already lower than 5×10^{-5} . This clearly supports the need for developing strategies that can accelerate the convergence of the set-inverse algorithm, such as

box-reduction and other CPU-time-reduction approaches.

3.2 Domain Reduction and CPU-Time Reduction Strategies

OPTIMIZATION-BASED DOMAIN REDUCTION. An advantage inherent to using Taylor models for bounding the reachable set of a dynamic system is that the multivariate polynomial part captures the parametric dependencies in the ODE solutions. For given restrictions (constraints) on the state or output variables, it becomes possible in turn to exclude part of the parameter set for which these restrictions cannot be met—the so-called constraint propagation approach. Lin and Stadtherr (2007a) and Kletting et al. (2011) used this idea in order to contract the parameter boxes at each iteration of the set-inversion algorithm. Notwithstanding its effectiveness, this approach only exploits the dependencies contained in the linear parts of the Taylor models. In contrast, this paper proposes an optimization-based domain-reduction approach that fully exploits the dependencies in the Taylor models of the predicted outputs.

Given a parameter box $P := [p^L, p^U]$ as well as q th-order Taylor model enclosures $\bar{Y}(t_i, P) := \{\mathcal{P}_y^q(t_i, p) \mid p \in P\} \oplus \mathcal{R}_y^q(t, P) \supseteq Y(t_i, P)$, the lower and upper parameter bounds p_j^L and p_j^U for each $j = 1, \dots, n_p$ can be tightened by solving optimization problems of the form

$$p_j^L = \max \{p_j \mid \{\mathcal{P}_y^q(t_i, p)\} \oplus \mathcal{R}_y^q(t, P) \supseteq Y_p(t_i), \forall p \in P, \forall i = 1, \dots, N\}, \quad (3.9a)$$

$$p_j^U = \min \{p_j \mid \{\mathcal{P}_y^q(t_i, p)\} \oplus \mathcal{R}_y^q(t, P) \supseteq Y_p(t_i), \forall p \in P, \forall i = 1, \dots, N\}. \quad (3.9b)$$

This way, a reduced box P is obtained after solving $2 \times n_p$ optimization problems—one problem for the lower bound and one for the upper bound of each parameter. In the case of Taylor models with interval remainder bounds for instance, $\mathcal{R}_y^q(t, P) := [r_y^{q,L}(t, P), r_y^{q,U}(t, P)]$, the inclusion constraints in (3.9) can be equivalently rewritten in the form

$$\mathcal{P}_y^q(t_i, p) + r_y^{q,L}(t, P) \leq \text{Sup}\{Y_p(t_i)\}, \quad (3.10a)$$

$$\mathcal{P}_y^q(t_i, p) + r_y^{q,U}(t, P) \geq \text{Inf}\{Y_p(t_i)\}, \quad (3.10b)$$

for each $i = 1, \dots, N$.

Since the range of the polynomial part of a q th-order Taylor model, $\{\mathcal{P}_y^q(t_i, p) \mid p \in P\}$, turns out to be a nonconvex set for $q \geq 2$ in general, the bound-reduction problems (3.9) themselves be nonconvex. Instead of trying to solve these problems directly to global optimality, we construct polyhedral relaxations in the form of linear programs (LPs), similar to the approach used for bound contraction in branch-and-bound search (see, e.g., Zamora and Grossmann, 1999; Neumaier, 2004; Tawarmalani and Sahinidis, 2004). This relaxation procedure follows three steps:

1. **Decomposition.** The multivariate polynomials $\mathcal{P}_y(t_i, \cdot)$, $i = 1, \dots, N$, are decomposed into factored form, comprised of binary sums, binary products and univariate composition terms only, via the introduction of auxiliary variables (Smith and Pantelides, 1999; Tawarmalani and Sahinidis, 2004). Here, the constraints in the reformulated optimization problem are either linear or contain a single bilinear term $p_j p_k$ or integer power term $(p_j)^k$.
2. **Relaxation.** The nonconvex terms in the reformulated problem are relaxed so as to obtain a convex optimization problem. Here, this relaxation involves replacing both the bilinear and integer power terms with their convex/concave envelopes, e.g., based on McCormick relaxations (McCormick, 1976).

3. **Polyhedral Outer Approximation.** Since the convex/concave envelopes of power terms are nonlinear in general, polyhedral outer-approximations are constructed via linearization at a number of points. These are usually so chosen as to meet a given level of accuracy (Tawarmalani and Sahinidis, 2004).

By construction, the relaxed optimization problems are fully linear, making it possible to exploit the robustness, efficiency and speed of state-of-the-art LP solvers such as GUROBI or CPLEX. We also note that further improvements could be obtained by tightening the relaxations, for instance using reformulation-linearization technique (RLT) (Sherali, 2002; Sherali et al., 2012) or exploiting intermediate substructures in the factored optimization problem (Zorn and Sahinidis, 2014; Misener and Floudas, 2014).

In practice, the domain-reduction procedure can be performed as an extra step in the set-inversion algorithm, between Steps 2 and 3. Moreover, in case the reduction of a parameter box P is larger than a given threshold, for instance $\geq 20\%$ in volume, it can be repeated multiple times. It is important to bear in mind that repeating the reduction several times requires recomputing the enclosures $\bar{Y}(t_i, P)$ of the model outputs on the reduced box P though. This defines a clear trade-off between the extra computational burden and the reduction in the size of the partition \mathbb{P}_{bnd} , which is of course problem-dependent. An illustration of the effectiveness of this approach is presented below.

CPU-TIME REDUCTION. When combined with domain-reduction techniques, Taylor models can improve the convergence speed of the set-inversion algorithm significantly. But because Taylor models can also cause a large computational overhead, this benefit is mostly noticeable at an early stage of the set-inversion procedure, when many boxes can be fathomed or greatly reduced. This calls for further CPU-time-reduction strategies in order to make guaranteed parameter estimation more competitive for high-order Taylor models.

In the basic set-inversion algorithm of Sect. 2.2, the enclosures $\bar{Y}(t_i, P)$ are recomputed at every iteration because of the overestimation inherent to ODE bounding techniques. In this context, a simple CPU-time-reduction strategy involves reusing the enclosures computed at a parent node (i.e., for a larger parameter box P) as soon as the overestimation at all sampling times and for all output variables has become smaller than a given threshold $\epsilon_{\text{cvg}} > 0$. For Taylor models, such overestimation is directly measured by the remainder term \mathcal{R}_x^q , and a possible re-usability condition thus reads

$$\forall i \in \{1, \dots, N\}, \quad \text{diam}(\mathcal{R}_x^q(t_i, P)) \leq \epsilon_{\text{cvg}}. \quad (3.11)$$

As soon as this condition is met, the corresponding Taylor models $(\mathcal{P}_x^q(t_i, \cdot), \mathcal{R}_x^q(t_i, P))$ can indeed be stored and used later on in any child node $P' \subseteq P$, effectively by-passing the ODE bounding step 2. Variants of this approach can of course be used that consider relative convergence criteria and scaling for instance.

In addition to reusing Taylor models at children nodes, a further CPU-time-reduction strategy involves reducing the order of the Taylor models, which can lead to significant savings in connection to the relaxation and solution of the optimization-based domain-reduction problems (3.9). A simple order-reduction procedure is as follows:

Input: Convergence threshold $\epsilon_{\text{cvg}} > 0$; parameter box P ; q th-order Taylor models $(\mathcal{P}_x^q(t_i, \cdot), \mathcal{R}_x^q(t_i, P))$ of $x(t_i, \cdot)$ on P satisfying (3.11)

Initialization: Set reduced order $\rho = q$

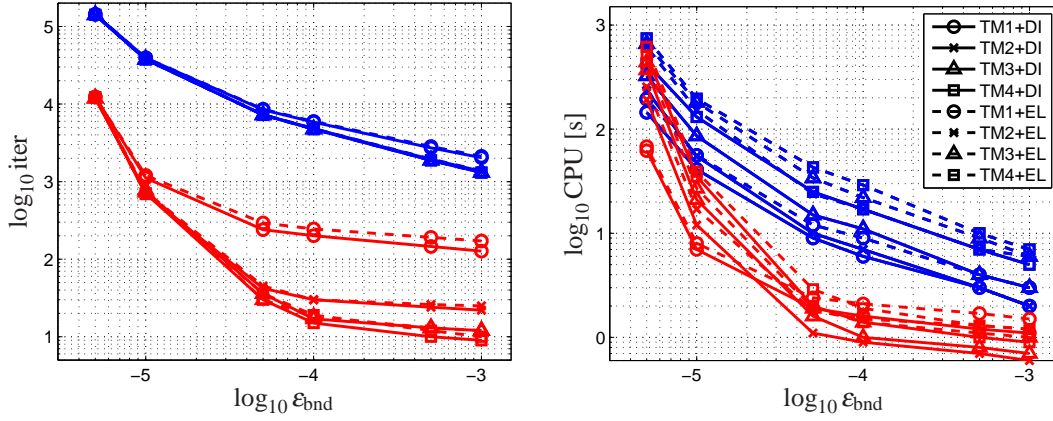


FIG. 4. Performance of guaranteed parameter estimation with (red lines) and without (blue lines) the use of domain reduction and with ODE bounding techniques based on Taylor models of orders $q = 1, \dots, 4$. Left: Number of iterations vs. convergence threshold. Right: CPU time vs. convergence threshold.

Main Loop:

1. Compute enclosures $B^p(t_i, P) \supseteq \left\{ \sum_{\substack{\gamma \in \mathbb{N}^{n_p} \\ |\gamma| = p}} \frac{\partial^\gamma x(t_i, \hat{p})}{\gamma!} \mid p \in P \right\}$ for all $i \in \{1, \dots, N\}$
2. **If** $\text{diam}(B^p(t_i, P)) > \varepsilon_{\text{cvg}}$ for some $i \in \{1, \dots, N\}$, **stop**
3. Reduce order $\rho = \rho - 1$; **Return** to step 1

Output: Reduced Taylor model order ρ

In particular, bounding of all the monomials of a given order ρ in step 1 can be achieved using interval analysis or other less conservative strategies (Lin and Rokne, 1995; Neumaier, 2002). Regarding the convergence threshold ε_{cvg} finally, we like to note that a larger threshold will lead to reusing Taylor models from parent nodes earlier as well as reducing their order faster, but too large a threshold can prevent convergence of the set-inversion algorithm if the stopping criterion is based solely on the total volume threshold ε_{bnd} (Step 4a).

CASE STUDY (CONTINUED). We continue the case study of the dynamic system (3.8) in order to investigate the effect of optimization-based domain reduction and CPU-time reduction. Guaranteed parameter estimation is applied with and without the use of domain reduction as an extra step in the set-inversion algorithm (reduction threshold of 20% and maximum of 10 reduction loops at each iteration). Taylor models of orders $q = 1, \dots, 4$ are considered for enclosing the output reachable set $\bar{Y}(\cdot, P)$, and the termination criteria remain the same as defined previously.

The number of iterations and the CPU time required by the set-inversion algorithm to terminate with different Taylor model orders and with or without the use of domain reduction are reported on the left and right plots of Fig. 4, respectively, as a function of the termination tolerance ε_{bnd} . It is evident that the number of iterations decreases significantly when domain reduction is used—here by at least one order of magnitude for all considered tolerance levels ε_{bnd} . Moreover, the higher the order of the Taylor

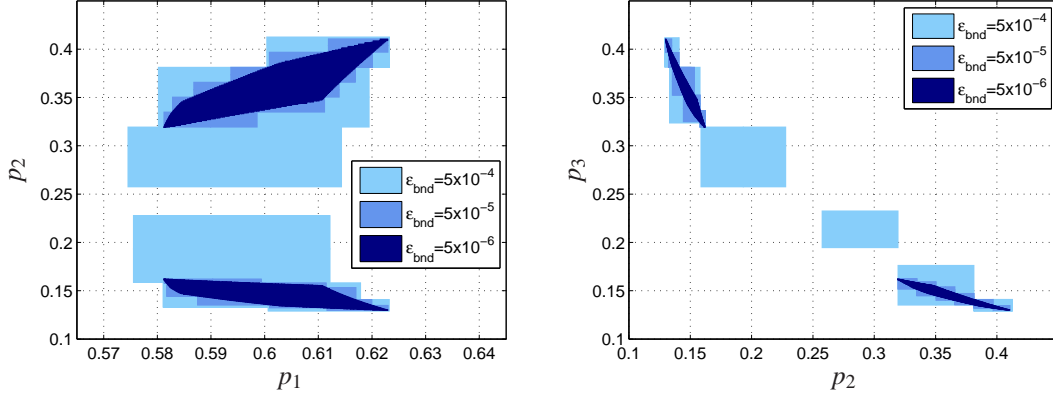


FIG. 5. Outer approximations of the sets of guaranteed parameter estimates for different levels of accuracy ϵ_{bnd} using domain reduction and second-order Taylor models. Left: Projections onto (p_1, p_2) space. Right: Projections onto (p_2, p_3) space.

model, the smaller the number of iterations required by the algorithm to converge for a given accuracy level. In terms of overall CPU time, the use of domain reduction is found to be mostly beneficial at an early stage of the set-inversion procedure, where many boxes can be significantly reduced or even eliminated using optimization-based domain reduction.

A certain trade-off is observed in terms of CPU time on the right plot of Fig. 4, whereby higher-order Taylor models can cause a significant computational overhead. On the whole, first- or second-order Taylor models with interval remainder bounds are found to enable the fastest computations in this case study. Although higher-order Taylor models reduce the overestimation, lower-order Taylor models eventually become computationally advantageous as the parameter boxes shrink. Another trade-off is observed in terms of the overhead caused by the application of domain reduction (construction and solution of relaxed LP problems). These trends show a clear need for CPU-time-reduction strategies in connection to Taylor model-based ODE bounding. Nonetheless, when used in combination with domain reduction, Taylor model-based ODE bounders now greatly outperforms classical differential inequalities (see Fig. 2).

The plots in Fig. 5 show the projections of the solution set outer-approximation onto the (p_1, p_2) and (p_2, p_3) subspaces, for increasing accuracy levels of $\epsilon_{\text{bnd}} = 5 \times 10^{-4}$, 5×10^{-5} , and 5×10^{-6} , using optimization-based domain reduction and second-order Taylor models with interval remainder terms for ODE bounding. In comparing outer-approximations of the guaranteed parameter set P_e for various accuracy levels, it is found that setting $\epsilon_{\text{bnd}} = 5 \times 10^{-5}$ already provides a tight approximation of P_e , with only 34 boxes and a run-time of about 2 sec.

As expected, a much tighter approximation is obtained by setting $\epsilon_{\text{bnd}} = 5 \times 10^{-6}$, yet this is at the price of a much finer box partition comprising 11,250 boxes here and a corresponding run-time of over 60 sec. For the sake of comparison we also note that, when no domain reduction is used, the partition comprises over 2,200 boxes with $\epsilon_{\text{bnd}} = 5 \times 10^{-5}$ and over 70,000 boxes with $\epsilon_{\text{bnd}} = 5 \times 10^{-6}$. These results also suggest that the efficiency of the set-inversion algorithm in computing highly-accurate set approximations could be improved significantly if affine cuts were enabled in addition to simple bounds contraction during the domain-reduction procedure. Such cuts would provide the extra flexibility needed to closely approximate the actual parameter set and will be the topic of future research.

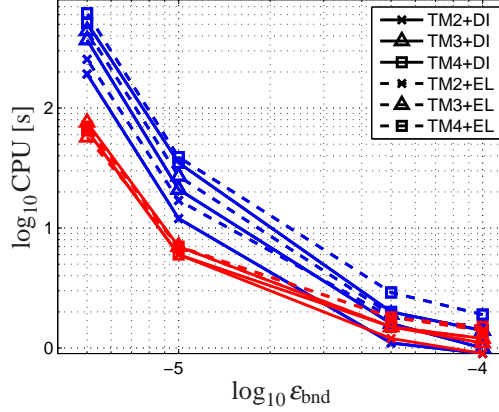


FIG. 6. Performance of guaranteed parameter estimation with (red lines) and without (blue lines) CPU-time-reduction strategies, in combination with domain-reduction strategy and ODE bounding techniques based on Taylor models of orders $q = 2, \dots, 4$: CPU time vs. convergence threshold.

Finally, we investigate the effect of CPU-time reduction, by considering both strategies of reusing and reducing the order of Taylor models computed at parent nodes. A convergence threshold of $\epsilon_{\text{cvg}} = 10^{-4}$ (determined heuristically) is used here.

Computational time requirements for the set-inversion algorithm to converge are shown in Fig. 6 for various termination tolerances ϵ_{bnd} . Not reported on this plot are the CPU times for first-order Taylor models since the corresponding improvement is marginal—convergence of first-order Taylor models within $\epsilon_{\text{cvg}} = 10^{-4}$ is only achieved for very small parameter boxes in this case. For higher-order Taylor models, it is evident that the CPU-time-reduction strategies are effective. The best performance is achieved when second-order Taylor models with interval remainder bounds are used, but third- and fourth-order Taylor models lead to comparable run-times nonetheless.

With all the proposed improvements used together, guaranteed parameter estimation of the dynamic system (3.8) can be solved to within $\epsilon_{\text{bnd}} = 5 \times 10^{-6}$ in less than 60 seconds. This is a three-fold reduction compared to the classical method of differential inequalities.

4. Guaranteed Parameter Estimation for an Anaerobic Digestion Process

This section illustrates the benefits of using high-order ODE bounding, optimization-based domain reduction, and CPU-time reduction in the context of guaranteed parameter estimation for a case study in anaerobic digestion. We consider a six-state model representing the dynamics of an anaerobic digester, as originally proposed by Bernard et al. (2001). Enclosing the solutions of this model in the presence of parametric uncertainty is challenging due to the presence of complex and liquid-gas transfer and pH self-regulation mechanisms. Moreover, the system exhibits both fast dynamics acting on a time-scale of minutes/hours, and slow dynamics acting on a time-scale of days.

$$\dot{X}_1 = (\mu_1(S_1) - \alpha D) X_1, \quad (4.1a)$$

$$\dot{X}_2 = (\mu_2(S_2) - \alpha D) X_2, \quad (4.1b)$$

$$\dot{S}_1 = D(S_1^{\text{in}} - S_1) - k_1 \mu_1(S_1) X_1, \quad (4.1c)$$

$$\dot{S}_2 = D(S_2^{\text{in}} - S_2) + k_2\mu_1(S_1)X_1 - k_3\mu_2(S_2)X_2, \quad (4.1d)$$

$$\dot{Z} = D(Z^{\text{in}} - Z), \quad (4.1e)$$

$$\dot{C} = D(C^{\text{in}} - C) - q_{\text{CO}_2} + k_4\mu_1(S_1)X_1 + k_5\mu_2(S_2)X_2. \quad (4.1f)$$

The states X_1 and X_2 stand for the concentrations of acidogenic and methanogenic biomass, respectively; S_1 , the organic substrate concentration (COD other than VFA); S_2 , the volatile fatty acids (VFA) concentration; Z , the total alkalinity concentration (TALK); and C , the total inorganic carbon concentration (TIC). Moreover, D represents the dilution rate; S_1^{in} , S_2^{in} , Z^{in} and C^{in} are the inlet concentrations of organic substrate, VFA, TALK and TIC, respectively; α is the fraction of biomass in the liquid phase (i.e., not attached to a support); and k_1, \dots, k_6 are pseudo-stoichiometric yield coefficients.

The specific growth rates of acidogenic bacteria, μ_1 , and methanogenic bacteria, μ_2 , are assumed to follow Michaelis-Menten and Haldane kinetics,

$$\mu_1(S_1) := \bar{\mu}_1 \frac{S_1}{S_1 + K_{S_1}}, \quad (4.1g)$$

$$\mu_2(S_2) := \bar{\mu}_2 \frac{S_2}{S_2 + K_{S_2} + S_2^2/K_{I_2}}, \quad (4.1h)$$

with maximum growth rates $\bar{\mu}_1$ and $\bar{\mu}_2$, half-saturation constants K_{S_1} and K_{S_2} , and inhibition constant K_{I_2} (methanogenic bacteria only). Finally, the molar flowrate of CO_2 , q_{CO_2} , is given by

$$q_{\text{CO}_2} := k_L a (C + S_2 - Z - K_H P_{\text{CO}_2}), \quad (4.1i)$$

$$\text{with } P_{\text{CO}_2} := \frac{\phi_{\text{CO}_2} - \sqrt{\phi_{\text{CO}_2}^2 - 4K_H P_t (C + S_2 - Z)}}{2K_H} \quad (4.1j)$$

$$\phi_{\text{CO}_2} := C + S_2 - Z + K_H P_t + \frac{k_6}{k_L a} \mu_2(S_2)X_2, \quad (4.1k)$$

where $k_L a$ denotes the liquid-gas transfer constant, K_H is Henry's constant, and P_t is the total pressure.

Nominal values for all the parameters are taken from (Bernard et al., 2001). We apply guaranteed parameter estimation to estimate the kinetic parameters describing biomass growth. These parameters are listed in Table 1 with their nominal values and the considered variation ranges. The rest of the parameters as well as the initial conditions used are reported in Table 2 for sake of completeness.

Table 1. Estimated parameters of the anaerobic digestion model (4.1).

Parameter	Nominal value	Range	Unit
$\bar{\mu}_1$	1.2	[1.15, 1.25]	/day
K_{S_1}	7.1	[6.7, 7.3]	g(COD)/L
$\bar{\mu}_2$	0.74	[0.735, .75]	/day
K_{S_2}	9.28	[9.2, 9.5]	mmol/L
K_{I_2}	256	[235.0, 265.0]	mmol/L

In order to apply guaranteed parameter estimation, pseudo-experimental data are generated by simulating the model (4.1) with nominal parameter values from Tables 1 and 2 over a four-day period. The profiles used for the dilution rate and for the influent concentrations are those reported in Table 3.

Table 2. Constant parameters and initial states of the anaerobic digestion model (4.1).

Parameter	Value	Unit	Parameter	Value	Unit
k_1	42.14	g(COD)/g(cell)	$X_1(0)$	0.5	g(VSS)/L
k_2	116.5	mmol/g(cell)	$X_2(0)$	1.0	g(VSS)/L
k_3	268.0	mmol/g(cell)	$S_1(0)$	1.0	g(COD)/L
k_4	50.6	mmol/g(cell)	$S_2(0)$	5.0	mmol/L
k_5	343.6	mmol/g(cell)	$C(0)$	40.0	mmol/L
k_6	453.0	mmol/g(cell)	$Z(0)$	50.0	mmol/L
k_{La}	19.8	day ⁻¹	P_i	1	atm
K_H	16	mmolL ⁻¹ atm ⁻¹	α	0.5	–

Moreover, three outputs are considered to carry out the estimation, namely S_1 , S_2 , and C , with measurements every 4 hours. In order to simulate the effect of measurement noise, the simulated values are rounded up or down to the nearest values by retaining, respectively, 2, 1 and 1 significant digits only; then, measurement error ranges of, respectively, ± 0.01 , ± 0.1 and ± 0.1 are added around these values.

Table 3. Dilution rate and inlet concentration profiles corresponding to the pseudo-experimental data.

Input	Day 1	Day 2	Day 3	Day 4
D [/day]	0.25	1.00	1.00	0.25
S_1^{in} [g(COD)/L]	2.38	2.38	4.76	2.38
S_2^{in} [mmol/L]	80.0	80.0	160.0	80.0
Z^{in} [mmol/L]	50.0	50.0	100.0	50.0
C^{in} [mmol/L]	5.0	5.0	10.0	5.0

In the remainder of this section, we investigate guaranteed parameter estimation with different bounding techniques and with both optimization-based domain-reduction and CPU-time-reduction strategies in order to demonstrate the proposed improvements on a real-life problem. As previously in the simple case study, a 20% threshold and a maximum of 10 reduction loops are defined for the optimization-based domain-reduction strategy, and an absolute convergence threshold of $\varepsilon_{\text{cvg}} = 10^{-4}$ is defined in connection to the CPU-time-reduction strategy. Problems of increasing complexity with 2, 3, 5 and 7 estimated parameters are addressed in Sect. 4.1–Sect. 4.3 below.

4.1 Case Study 1 – Two-Parameter Guaranteed Parameter Estimation

We consider the estimation of the parameters $\bar{\mu}_1$ and K_{S_1} , while the rest of the parameters from Tables 1 and 2 are fixed at their nominal values. The set-inversion algorithm is used with continuous-time ODE bounding techniques propagating Taylor models of orders $q = 1, \dots, 4$ with interval or ellipsoidal remainders. The termination criterion is defined as $\varepsilon_{\text{bnd}} = 10^{-4}$, whereas ε_{box} is set to zero, and the maximum number of iterations and maximal computational time are set to 1,000,000 iterations and 10 hours, respectively.

Fig. 7 shows both the inner- and outer-approximation of the set of guaranteed parameter estimates for the selected termination criteria. We start by noting that the true parameter values lie inside the approximation of the set P_e and that the selected termination criteria appear to be appropriate in view

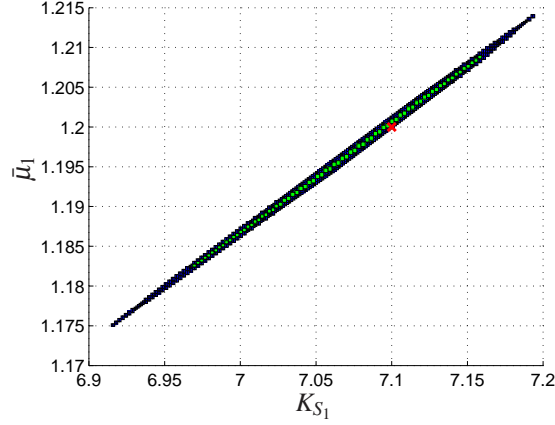


FIG. 7. Guaranteed parameter set approximation (\mathbb{P}_{int} in green, \mathbb{P}_{bnd} one in blue) for Case Study 1. The red cross indicates the ‘true’ (nominal) parameter values.

Table 4. Iteration counts and run-times of the set-inversion algorithm for Case Study 1.

Bounding method	Domain reduction	CPU-time reduction	Number of iterations	CPU time [s]
TM1+EL	✗	✗	8,738	801
TM1+DI	✓	✗	4,280	1682
TM4+EL	✓	✗	3,690	5748
TM4+EL	✓	✓	5,229	49

of the approximation level. Such a shape of the guaranteed parameter set is characteristic of the large correlations between the parameters $\bar{\mu}_1$ and K_{S_1} , according to (4.1g), and shows that $\mu_1(S_1) \approx \frac{\bar{\mu}_1}{K_{S_1}} X_1$ in this case.

When the method of differential inequalities is used to bound the reachable set, the algorithm stops after 1,000,000 iterations, without reaching the desired level of approximation—The volume of the partition \mathbb{P}_{bnd} is about 2.5×10^{-4} then. This behavior is attributed to the inability of the method of differential inequalities to generate tight bounds for the anaerobic digestion model, even for very small parameter uncertainty. In contrast, higher-order ODE bounding techniques enable convergence of the set-inversion algorithm, as summarized in Table 4. Using Taylor models in combination with domain reduction, the algorithm is found to converge within a few thousand iterations (2nd and 3rd row), yet this remains insufficient to override the extra computational burden associated with domain reduction (1st row). The use of domain reduction becomes advantageous only when combined with CPU-time reduction (4th row), then leading to dramatic reduction of the run-time down to 49 s. Note that the number of iterations increases in the latter case compared to a run with the same settings but without CPU-time-reduction strategies, a behavior that is indeed expected and attributed to the approximation introduced by the finite convergence threshold ε_{cvg} .

Table 5. Iteration counts and run-times of the set-inversion algorithm for Case Study 2.

Bounding method	Domain reduction	CPU-time reduction	Number of iterations	CPU time [s]
TM1+EL	✗	✗	23,838	3,584
TM1+DI	✓	✗	4,095	1,154
TM4+EL	✓	✗	3,423	9,593
TM4+EL	✓	✓	3,488	111

4.2 Case Study 2 – Three-Parameter Guaranteed Parameter Estimation

Next, we consider the estimation of the parameters $\bar{\mu}_2$, K_{S_2} and K_{I_2} , while the rest of the parameters from Tables 1 and 2 are fixed at their nominal values. The set-inversion algorithm is run with the exact same settings as previously in Sect. 4.1, to the exception of the termination criterion ε_{bnd} that is now set to 5×10^{-5} .

Fig. 8 shows the outer-approximation of the set of guaranteed parameter estimates for the selected termination criteria. The true parameter values lie inside the approximation of the set P_e and the selected termination criteria is deemed appropriate by visual inspection of the approximation level. Here again, the shape of the guaranteed parameter set is expected given the large correlations between the parameters $\bar{\mu}_2$, K_{S_2} and K_{I_2} according to (4.1h).

When the method of differential inequalities is used to bound the reachable set, the algorithm stops after 1,000,000 iterations, without reaching the desired level of approximation—The volume of the partition \mathbb{P}_{bnd} is about 1.7×10^{-3} then. This behavior is again due to the inability of the method of differential inequalities to generate tight bounds for the anaerobic digestion model, even for very small parametric uncertainty. In contrast, higher-order ODE bounding techniques enable convergence of the set-inversion algorithm, as summarized in Table 5. [Using first-order Taylor models with ellipsoidal remainders but no other improvement, the set-inversion algorithm takes about 24,000 iterations to converge \(1st row\). This is to be compared with a few thousand iterations when domain domain reduction is used \(2nd, 3rd and 4th rows\), similar to the previous 2-parameter case despite the extra parameter.](#) This suggests that the domain reduction might become more and more advantageous as the number of uncertain parameters increases, a trend that will confirm later on in Sect. 4.3. An expected behavior here is the reduction in the number of iterations as higher-order Taylor models are used. Finally, the effect of the CPU-time-reduction strategy is rather dramatic, with a run-time reduction about 2 orders of magnitude lower in the case of fourth-order Taylor models with ellipsoidal remainder bounds. It is noteworthy that the shortest runtime in this case is even lower, down to 41 s when fourth-order Taylor models with interval remainder bounds are used.

4.3 Case Study 3 – Five- and Seven-Parameter Guaranteed Parameter Estimation

We now consider the estimation of the parameters $\bar{\mu}_1$, K_{S_1} , $\bar{\mu}_2$, K_{S_2} and K_{I_2} simultaneously, leaving the other parameters at their nominal values in Table 2. The set-inversion algorithm is run with the exact same settings as previously in Sect. 4.1 and Sect. 4.2, apart from the termination criterion ε_{bnd} that is now set to 5×10^{-7} .

Fig. 9 shows the outer-approximation of the set of guaranteed parameter estimates for the selected termination criterion. The true parameter values lie inside the approximation of the set P_e and the approximation level, although coarse, validates the chosen termination criterion. Large correlations between the parameters $\bar{\mu}_1$ and K_{S_1} , on the one hand, and between $\bar{\mu}_2$, K_{S_2} and K_{I_2} , on the other hand,

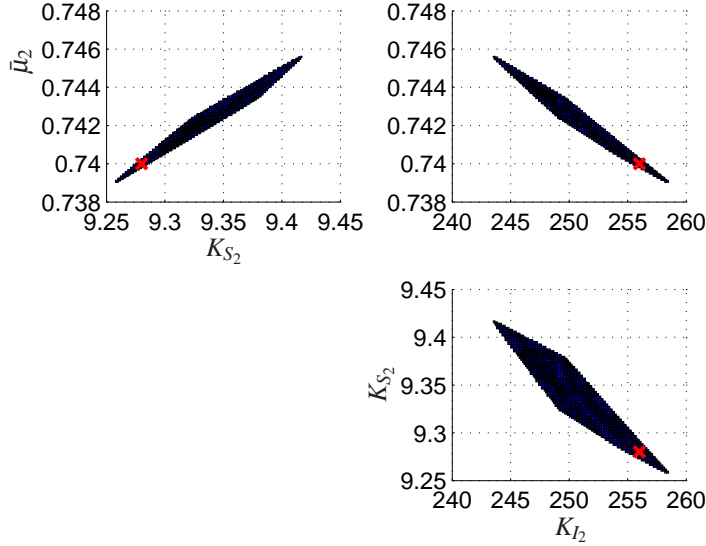


FIG. 8. Outer approximation of the set of guaranteed parameter estimates for Case Study 2. Projections onto the subspaces $(K_{S_2}, \bar{\mu}_2)$, $(K_{I_2}, \bar{\mu}_2)$, and (K_{I_2}, K_{S_2}) . The red crosses indicate the true (nominal) parameter values.

are observed, which is in complete agreement with the results shown earlier in Fig. 7 and Fig. 8. In contrast, rather small cross-correlations are observed between these two parameter subsets, as illustrated for instance for the parameters K_{S_1} and $\bar{\mu}_2$ in the top-right plot of Fig. 9.

Table 6 presents a comparison of the performance of various ODE bounding techniques and other improvement strategies. As previously, early termination is obtained with the method of differential inequalities after 1,000,000 iterations (without a single parameter box being fathomed here), and the algorithm now fails to converge after 10 hours with Taylor models as well when domain reduction is not applied. With respect to Taylor models combined with domain reduction, the benefit of higher-order ODE bounds in terms of the number of iterations is becoming more obvious in this 5-parameter problem—for instance, 10 times more iterations are needed with a first-order Taylor model compared to a fourth-order one. Yet, this large reduction is still not enough to overpower the extra computational burden of a single iteration with a higher-order Taylor model. Only when used in combination with CPU-time-reduction strategies are fourth-order Taylor models found to become competitive, with a runtime down to about 2,100 s. Finally, it is noteworthy that the shortest runtime in this case is close to 1,400 s, which is obtained for fourth-order Taylor model with interval remainder bounds and with all the developed reduction strategies.

Concerning the anaerobic digestion application, a more realistic parameter estimation problem should of course consider the initial biomass concentrations to be uncertain as well. Adding both initial concentrations $X_1(0)$ and $X_2(0)$ to the five uncertain kinetic parameters in Table 1 yields a total of seven parameters. In order to carry out the computations, only a small level of uncertainty of ± 0.001 g(VSS)/L is considered for $X_1(0)$ and $X_2(0)$ here, and the termination criterion ε_{bnd} is decreased to 5×10^{-11} . The set-inversion algorithm appears to be tractable only with fourth-order Taylor models (or higher) and only when combined with domain-reduction and CPU-time-reduction strategies—

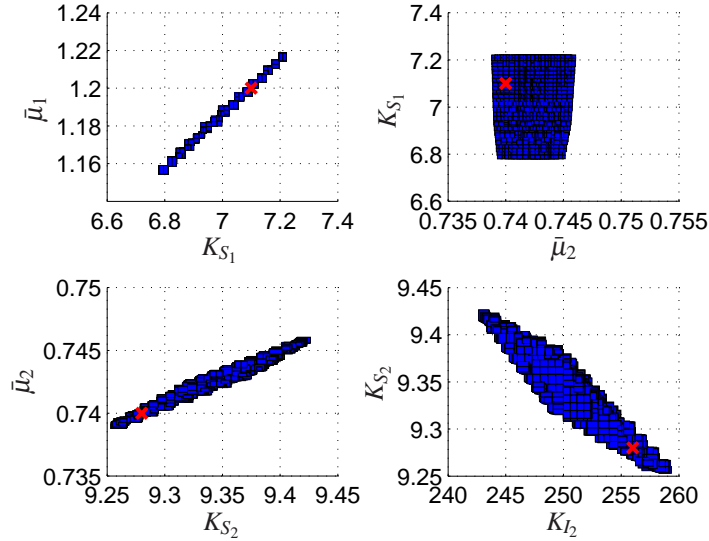


FIG. 9. Outer approximation of the set of guaranteed parameter estimates for Case Study 3. Projections onto the subspaces $(K_{S_1}, \bar{\mu}_1)$, $(\bar{\mu}_2, K_{S_1})$, $(K_{S_2}, \bar{\mu}_2)$, and $(K_{J_2}, \bar{\mu}_2)$ only. The red crosses indicate the true (nominal) parameter values.

Table 6. Iteration counts and run-times of the set-inversion algorithm for Case Study 3.

Bounding method	Domain reduction	CPU-time reduction	Number of iterations	CPU time [s]
TM1+EL	✗	✗	219,178	36,000 ^a
TM1+DI	✓	✗	41,148	8,295
TM4+EL	✓	✗	3,010	18,346
TM4+EL	✓	✓	3,130	2,118

^aDid not converge to the specified termination criterion within the maximum allowable time.

^aDid not converge to the specified termination criterion within the maximum allowable time.

Convergence is achieved after 4,225 iterations and a corresponding runtime just above 19,000 s (5.3 h) in this case. These results confirm the advantage of high-order ODE bounding techniques and improved domain- and CPU-time-reduction strategies in addressing real-life problems possessing complex dynamics and more than a handful of uncertain parameters.

4.4 Discussion

In order to bring guaranteed parameter estimation to the next level and allow for problems with more than ten parameters, the case studies in this paper suggest that improving the performance of existing ODE bounding techniques remains key, i.e., to enable tight bounds on larger parameter ranges while reducing the computational burden at the same time. It is also found that both domain-reduction and CPU-time-reduction strategies can have a dramatic effect on the performance, and it would thus appear important to develop new ‘smart’ heuristics to further enhance the search; see, e.g., the recent work by Caprara and Locatelli (2010); Locatelli (2014). The use of tight termination tolerances in the set-inversion algorithm typically results in a very large accumulation of parameter boxes, similar in essence

to the cluster effect in global optimization (Du and Kearfott, 1994; Neumaier, 2004). Strategies to help mitigate this behavior are also clearly warranted.

A practical limitation for the guaranteed parameter estimation problem as formulated in (2.5) is the need for consistent measurement data and bounds throughout the entire time series; otherwise, there may not be any model response matching the output measurements within the specified error bounds, in which case the parameter set \mathcal{P}_e is empty. In most applications based on real data, this calls for data preprocessing, for instance using data reconciliation techniques, in order to get rid of the outliers. For instance, these outliers could be due to over-optimistic noise bounds or to sensor failures at given time instants. To handle this situation, it is possible to ‘protect’ the estimator against at most n outliers, by allowing for a number of output variables to be outside of their prior feasible intervals (see, e.g., Jaulin et al., 2001; Kieffer and Walter, 2005). Another situation whereby the parameter set \mathcal{P}_e may be empty is in the presence of significant model mismatch. Taking guaranteed parameter estimation to the next level in order to address complex, large-scale problems of practical applicability, calls for the development of further robustification strategies as well as alternative guaranteed parameter estimation paradigms (Csáji et al., 2012; Kieffer and Walter, 2014).

5. Conclusions

The focus of this paper has been on the problem of guaranteed parameter estimation, which seeks to determine all parameter values of a dynamic model that are consistent with some experimental data, within specified error bounds. Set-inversion techniques based on exhaustive search are considered, with special emphasis on high-order bounding techniques for uncertain dynamic systems combined with efficient strategies for enhancing convergence speed. Specifically, a methodology is developed which starts by computing state/output bounds in the form of Taylor models that capture the parametric dependencies, and then takes full advantage of these Taylor models to perform optimization-based domain reduction. On top of this, strategies are implemented in order to decrease the overall run-time, whereby the Taylor models computed at a parent node can be reused and their order automatically reduced wherever possible. The potential of these new developments have been demonstrated both on a simple case study and on a real-life problem in anaerobic digestion with up to seven uncertain parameters, showing clear improvements of the set-inversion algorithm for guaranteed parameter estimation.

Acknowledgements

RP gratefully acknowledges the contribution of the European Commission under research project MOBOCON (Grant agreement number 291458). MV and BC acknowledge financial support from Marie Curie under grant PCIG09-GA-2011-293953. MV is also grateful to CONACYT for a doctoral scholarship.

References

- Belotti, P., Lee, J., Liberti, L., Margot, F. and Wächter, A. (2009). “Branching and bounds tightening techniques for non-convex MINLP”. *Optimization Methods & Software*, vol. 24 (4-5), pp. 597–634.
- Bernard, O., Hadj-Sadok, Z., Dochain, D., Genovesi, A. and Steyer, J.P. (2001). “Dynamical model development and parameter identification for an anaerobic wastewater treatment process”. *Biotechnology & Bioengineering*, vol. 75 (4), pp. 424–438.

- Berz, M. and Makino, K. (1998). “Verified integration of ODEs and flows using differential algebraic methods on high-order Taylor series”. *Reliable Computing*, vol. 4, pp. 361–369.
- Bompadre, A., Mitsos, A. and Chachuat, B. (2013). “Convergence analysis of Taylor and McCormick-Taylor models”. *Journal of Global Optimization*, vol. 57 (1), pp. 75–114.
- Caprara, A. and Locatelli, M. (2010). “Global optimization problems and domain reduction strategies”. *Mathematical Programming*, vol. 1 (125), pp. 123–137.
- Chachuat, B. and Villanueva, M.E. (2012). “Bounding the solutions of parametric ODEs: when Taylor models meet differential inequalities”. *22nd European Symposium on Computer Aided Process Engineering*, vol. 30 (I.D.L. Bogle and M. Fairweather, eds.). Elsevier, pp. 1307–1311.
- Csáji, B.C., Campi, M.C. and Weyer, E. (2012). “Non-asymptotic confidence regions for the least-squares estimate”. *Proceedings of the 16th IFAC Symposium on System Identification (SYSID 2012)*. Brussels, Belgium, p. 227232.
- Du, K. and Kearfott, R.B. (1994). “The cluster problem in multivariate global optimization”. *Journal of Global Optimization*, vol. 5 (3), pp. 253–265.
- Houska, B., Logist, F., Van Impe, J. and Diehl, M. (2012). “Robust optimization of nonlinear dynamic systems with application to a jacketed tubular reactor”. *Journal of Process Control*, vol. 22 (6), pp. 1152–1160.
- Houska, B., Villanueva, M.E. and Chachuat, B. (2013). “A validated integration algorithm for nonlinear ODEs using Taylor models and ellipsoidal calculus.” *Proceedings of the 52nd IEEE Conference on Decision and Control (CDC)*. Florence, Italy, pp. 484–489.
- Jaulin, L. (2002). “Nonlinear bounded-error state estimation of continuous-time systems”. *Automatica*, vol. 38, pp. 1079–1082.
- Jaulin, L., Kieffer, M., Didrit, O. and Walter, E. (2001). *Applied Interval Analysis*. Springer-Verlag, London.
- Jaulin, L. and Walter, E. (1993). “Set inversion via interval analysis for nonlinear bounded-error estimation”. *Automatica*, vol. 29 (4), pp. 1053–1064.
- Kieffer, M. and Walter, E. (2011). “Guaranteed estimation of the parameters of nonlinear continuous-time models: contributions of interval analysis”. *International Journal of Adaptive Control & Signal Processing*, vol. 25 (3), pp. 191–207.
- Kieffer, M. and Walter, E. (2014). “Guaranteed characterization of exact non-asymptotic confidence regions as defined by LSCR and SPS”. *Automatica*, vol. 50 (2), pp. 507–512.
- Kieffer, M. and Walter, E. (2005). “Interval analysis for guaranteed nonlinear parameter and state estimation”. *Mathematical & Computer Modelling of Dynamical Systems*, vol. 11 (2), pp. 171–181.
- Kletting, M., Kieffer, M. and Walter, E. (2011). “Two approaches for guaranteed state estimation of nonlinear continuous-time models”. *Modeling, Design, and Simulation of Systems with Uncertainties, Mathematical Engineering*, vol. 3. Springer, pp. 199–220.

- Kurzhanski, A.B. and Varaiya, P. (2002). “Reachability analysis for uncertain systems - The ellipsoidal technique”. *Dynamics of Continuous, Discrete & Impulsive System, Series B*, vol. 9 (3), pp. 347–368.
- Lin, Q. and Rokne, J.G. (1995). “Methods for bounding the range of a polynomial”. *Journal of Computational & Applied Mathematics*, vol. 58, pp. 193–199.
- Lin, Y. and Stadtherr, M.A. (2007a). “Guaranteed state and parameter estimation for nonlinear continuous-time systems with bounded-error measurements”. *Industrial & Engineering Chemistry Research*, vol. 46 (22), pp. 7198–7207.
- Lin, Y. and Stadtherr, M.A. (2007b). “Validated solutions of initial value problems for parametric ODEs”. *Applied Numerical Mathematics*, vol. 57 (10), pp. 1145–1162.
- Locatelli, M. (2014). “Alternative branching rules for some nonconvex problems”. *Optimization Methods & Software*, vol. in press, p. (DOI: 10.1080/10556788.2014.885521).
- Lohner, R.J. (1992). “Computation of guaranteed enclosures for the solutions of ordinary initial and boundary value problems”. *Computational Ordinary Differential Equations*, vol. 1 (J.R. Cash and I. Gladwell, eds.). Clarendon Press, pp. 425–436.
- Makino, K. and Berz, M. (1999). “Efficient control of the dependency problem based on Taylor model methods”. *Reliable Computing*, vol. 5 (1), pp. 3–12.
- McCormick, G.P. (1976). “Computability of global solutions to factorable nonconvex programs: Part I – Convex underestimating problems”. *Mathematical Programming*, vol. 10, pp. 147–175.
- Misener, R. and Floudas, C.A. (2014). “A framework for globally optimizing mixed-integer signomial programs”. *Journal of Optimization Theory & Applications*, vol. in press, pp. (DOI: 10.1007/s10957-013-0396-3).
- Moore, R.E. (1992). “Parameter sets for bounded-error data”. *Mathematics & Computers in Simulation*, vol. 34 (2), p. 113119.
- Moore, R.E., Kearfott, R.B. and Cloud, M.J. (2009). *Introduction to Interval Analysis*. SIAM, Philadelphia, PA.
- Nedialkov, N.S., Jackson, K.R. and Corliss, G.F. (1999). “Validated solutions of initial value problems for ordinary differential equations”. *Applied Mathematics & Computation*, vol. 105 (1), pp. 21–68.
- Neher, M., Jackson, K.R. and Nedialkov, N.S. (2007). “On Taylor model based integration of ODEs”. *SIAM Journal on Numerical Analysis*, vol. 45, pp. 236–262.
- Neumaier, A. (2002). “Taylor forms – Use and limits”. *Reliable Computing*, vol. 9 (1), pp. 43–79.
- Neumaier, A. (2004). “Complete search in continuous global optimization and constraint satisfaction”. *Acta Numerica*, vol. 13, pp. 271–369.
- Paulen, R., Villanueva, M.E., Fikar, M. and Chachuat, B. (2013). “Guaranteed parameter estimation in nonlinear dynamic systems using improved bounding techniques”. *Proceedings of the 2013 European Control Conference (ECC’13)*. Zürich, Switzerland, pp. 4514–4519.

- Raissi, T., Ramdani, N. and Candau, Y. (2004). “Set membership state and parameter estimation for systems described by nonlinear differential equations”. *Automatica*, vol. 40, pp. 1771–1777.
- Rauh, A., Hofer, E.P. and Auer, E. (2006). “VALENCIA-IVP: A comparison with other initial value problem solvers”. *Proceedings of the 12th GAMM-IMACS International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics (SCAN’2006)*. Duisburg, Germany, p. 36.
- Sahlodin, A.M. (2012). *Global Optimization of Dynamic Process Systems using Complete Search Methods*. Ph.D. thesis, McMaster University.
- Sahlodin, A.M. and Chachuat, B. (2011). “Convex/concave relaxations of parametric ODEs using Taylor models”. *Computers & Chemical Engineering*, vol. 35 (5), pp. 844–857.
- Sherali, H.D. (2002). “Tight relaxations for nonconvex optimization problems using the reformulation-linearization/convexification technique (RLT)”. *Handbook of Global Optimization*, vol. 2 (P.M. Pardalos and H.E. Romeijn, eds.). Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 1–63.
- Sherali, H.D., Dalkiran, E. and Liberti, L. (2012). “Reduced RLT representations for nonconvex polynomial programming problems”. *Journal of Global Optimization*, vol. 52 (3), pp. 447–469.
- Smith, E.M.B. and Pantelides, C.C. (1999). “A symbolic reformulation/spatial branch-and-bound algorithm for the global optimisation of nonconvex MINLPs”. *Computers & Chemical Engineering*, vol. 23 (4-5), pp. 457–478.
- Tawarmalani, M. and Sahinidis, N.V. (2004). “Global optimization of mixed-integer nonlinear programs: A theoretical and computational study”. *Mathematical Programming*, vol. 99 (3), pp. 563–591.
- Villanueva, M.E., Houska, B. and Chachuat, B. (2014). “Unified framework for the propagation of continuous-time enclosures for parametric nonlinear ODEs”. *Journal of Global Optimization*, vol. in press, DOI: 10.1007/s10898-014-0235-6.
- Villanueva, M.E., Paulen, R., Houska, B. and Chachuat, B. (2013). “Enclosing the reachable set of parametric ODEs using Taylor models and ellipsoidal calculus”. *23rd European Symposium on Computer Aided Process Engineering*, vol. 32 (A. Kraslawski and I. Turunen, eds.). Elsevier, pp. 979–984.
- Walter, E. (ed.) (1990). *Parameter Identifications with Error Bound, Mathematics & Computers in Simulation*, vol. 32. Elsevier.
- Walter, W. (1970). *Differential and integral inequalities*. Springer-Verlag, Berlin.
- Zamora, J.M. and Grossmann, I.E. (1999). “A branch and contract algorithm for problems with concave univariate, bilinear and linear fractional terms”. *Journal of Global Optimization*, vol. 14, pp. 217–249.
- Zorn, K. and Sahinidis, N.V. (2014). “Global optimization of general non-convex problems with intermediate bilinear substructures”. *Optimization Methods & Software*, vol. 29 (3), pp. 442–462.