Guaranteed Parameter Estimation in Nonlinear Dynamic Systems using Improved Bounding Techniques

Radoslav Paulen[1,2,*], Mario Villanueva[2], Miroslav Fikar[1], Benoit Chachuat[2]

***Abstract*— This paper is concerned with guaranteed parameter estimation in 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 outputs match the corresponding measurements within prescribed error bounds. An exhaustive search procedure 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 prespecified threshold on the approximation level is met. Exclusion tests rely on the ability to bound the solution set of the dynamic system for a given parameter subset and the tightness of these bounds is therefore paramount. Equally important is the time required to compute the bounds, thereby defining a trade-off. It is the objective of this paper to investigate this trade-off by comparing various bounding techniques based on interval arithmetic, Taylor model arithmetic and ellipsoidal calculus. When applied to a simple case study, ellipsoidal and Taylor model approaches are found to reduce the number of iterations significantly compared to interval analysis, yet the overall computational time is only reduced for tight approximation levels due to the computational overhead.**

I. Introduction

Process model development has become an integral part of modern process design methodologies as well as for 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, also known as model fitting, normally proceeds by determining parameter values for which the model predictions closely match the observed process. Failure to find an acceptable agreement calls for a revision of the model structure, and the parameter estimation is then repeated.

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 jeopardize a successful and reliable estimation procedure. First of all, structural model mismatch is inherent to the modeling exercise, and it would be an illusion to seek 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 its corresponding measuring apparatus. Moreover, measured data are typically corrupted with noise, for instance Gaussian white noise or more generally colored noise, let alone the presence of systematic offsets or temporal drifts caused by faulty or poorly calibrated sensors.

Among the possible approaches accounting for uncertainty in parameter estimation, the focus in this paper is on *guaranteed* parameter estimation [1], namely the determination of *all* parameter values—referred to as the solution set subsequently—that are consistent with the measurements under given uncertainty scenarios. More specifically, we consider the case that the uncertainty enters the estimation problem in the form of bounded measurement errors. In small-scale applications, the problem of approximating the solution set by a box partition, at an arbitrary precision, has been shown to be tractable by Walter and coworkers [1], [2] using set inversion techniques based on exhaustive search. These authors [2] also identify the computation of bounds on the solutions to the dynamic system as the main bottleneck in terms of convergence speed and accuracy of the resulting solution set approximation.

Computing exact bounds on the solution set of nonlinear parametric ODEs belongs to the class of computationally intensive problems (nonconvex optimization problem). In response to this, approximate methods that overestimate the solution set of parametric ordinary differential equations (ODEs), yet provide sufficiently tight bounds, have been developed over the years. These methods differ in the shape of the enclosing sets, for instance interval boxes [3], [4], Taylor models [5]–[7] or ellipsoids [8], and also in the way these sets are propagated through the flow of the ODEs. Interval bounds derived from Taylor models were first used by Lin and Stadtherr [9] in the context of guaranteed parameter estimation.

The main emphasis in the paper is on so-called bound-then-discretize methods, whereby auxiliary (nonparametric) ODEs are formulated whose solutions enclose those of the original parametric ODEs. Specifically, we consider classical differential inequalities [3], [4] and their combinations with Taylor models as described in [10], as well as the ellipsoidal bounding technique proposed in [8].

The rest of the paper is organized as follows. The problem of guaranteed parameter estimation is stated in Sect. II and the numerical solution procedure is also outlined. The various

[1]Faculty of Chemical and Food Technology, Slovak University of Technology, Bratislava, Slovakia (e-mail: {radoslav.paulen, miroslav.fikar}@stuba.sk)

[2]Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, UK (e-mail: {mario.villanueva10, b.chachuat}@imperial.ac.uk)

[*]Present address: Process Dynamics and Operations Group, Technische Universität Dortmund, Germany (e-mail: radoslav.paulen@bci.tu-dortmund.de)

bounding techniques for parametric ODEs are described in Sect. III. A comparison of these techniques on a simple case study is presented in Sect. IV and the results are discussed. Finally, Sect. V concludes the paper.

II. Guaranteed Parameter Estimation

A. Problem Statement

Consider a model of the observed process described by parametric ODEs of the form

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{p}), \qquad \boldsymbol{x}(t_0) = \boldsymbol{h}(\boldsymbol{p}), \tag{1a}$$

$$\hat{\boldsymbol{y}}(t) = \boldsymbol{g}(\boldsymbol{x}(t), \boldsymbol{p}), \tag{1b}$$

where $\boldsymbol{x}$ denotes the n_x-dimensional vector of process states; $\boldsymbol{p}$, the n_p-dimensional vector of process (a priori unknown) parameters; and $\hat{\boldsymbol{y}}$, the n_y-dimensional vector of model outputs (predictions).

Given a set of output measurements $\boldsymbol{y}_\mathrm{m}$ at N time points $t_1, \ldots, t_N$, *classical* parameter estimation seeks for *one* particular instance $\boldsymbol{p}_\mathrm{e}$ of the parameters for which the (possibly weighted) normed difference between these measurements and the corresponding model outputs $\hat{\boldsymbol{y}}$ is minimized. This optimization problem, for instance in the least-square sense, is given by:

$$\boldsymbol{p}_\mathrm{e} \in \arg\min_{\boldsymbol{p}\in\boldsymbol{P}_0} \sum_{i=1}^{N} \|\boldsymbol{y}_\mathrm{m}(t_i) - \hat{\boldsymbol{y}}(t_i)\|_2^2, \tag{2a}$$

$$\text{s.t.} \quad \text{model (1)}, \tag{2b}$$

where the interval box $\boldsymbol{P}_0 := [\boldsymbol{p}_0^L, \boldsymbol{p}_0^U]$ 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, $\boldsymbol{y}_\mathrm{p}$, are only known within some bounded measurement errors $\boldsymbol{e} \in \boldsymbol{E} := [\boldsymbol{e}^L, \boldsymbol{e}^U]$, so that

$$\boldsymbol{y}_\mathrm{p}(t_i) \in \boldsymbol{y}_\mathrm{m}(t_i) + [\boldsymbol{e}^L, \boldsymbol{e}^U] =: \boldsymbol{Y}_i. \tag{3}$$

Here, the main objective is to estimate the set $\boldsymbol{P}_\mathrm{e}$ of *all* possible parameter values $\boldsymbol{p}$ such that $\hat{\boldsymbol{y}}(t_i) \in \boldsymbol{Y}_i$ for every $i = 1, \ldots, N$; that is,

$$\boldsymbol{P}_\mathrm{e} := \left\{ \boldsymbol{p} \in \boldsymbol{P}_0 \;\middle|\; \begin{array}{l} \exists \boldsymbol{x}, \hat{\boldsymbol{y}} : \\ \dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{p}),\ \boldsymbol{x}(t_0) = \boldsymbol{h}(\boldsymbol{p}), \\ \hat{\boldsymbol{y}}(t_i) = \boldsymbol{g}(\boldsymbol{x}(t_i), \boldsymbol{p}), \\ \hat{\boldsymbol{y}}(t_i) \in \boldsymbol{Y}_i, \\ \forall t \in [t_0, t_N], \forall i \in \{1, \ldots, N\} \end{array} \right\}. \tag{4}$$

Depicted in red on the top plot of Figure 1 is the set of all output trajectories satisfying $\hat{\boldsymbol{y}}(t_i) \in \boldsymbol{Y}_i$, $i = 1, \ldots, N$, and on the bottom plot the corresponding set $\boldsymbol{P}_\mathrm{e}$ projected in the (p_1, p_2) space.

In the case that $\boldsymbol{P}_\mathrm{e}$ is nonempty, this set turns to be the same as the set of all global minimizers of the problem

$$\min_{\boldsymbol{p}\in\boldsymbol{P}_0} \sum_{i=1}^{N} \|\boldsymbol{Y}_i - \hat{\boldsymbol{y}}(t_i)\|_2^2, \tag{5a}$$

$$\text{s.t.} \quad \text{model (1)}, \tag{5b}$$

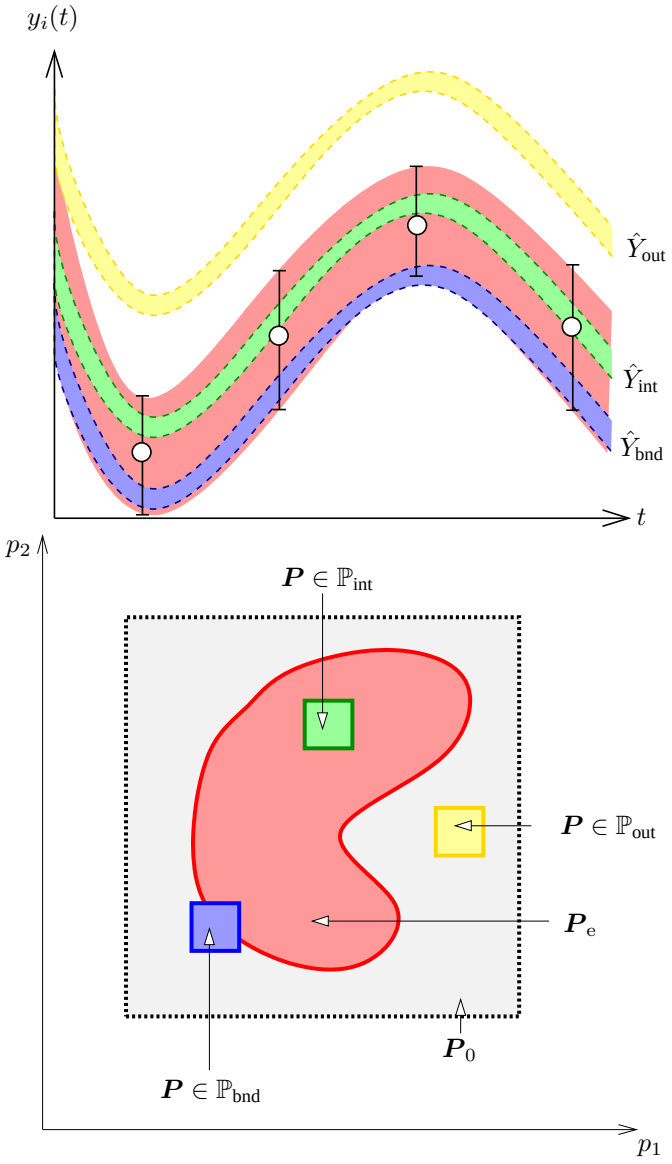

Fig. 1. Illustration of guaranteed parameter estimation concepts in the space of output trajectories (top plot) and parameters (bottom plot).

with the 2-norm of a compact set $\boldsymbol{X} \subset \mathbb{R}^n$ defined as

$$\|\boldsymbol{X}\|_2 := \max_{\boldsymbol{x},\boldsymbol{y}\in\boldsymbol{X}} \|\boldsymbol{x} - \boldsymbol{y}\|_2. \tag{6}$$

Obtaining an exact characterization of the set $\boldsymbol{P}_\mathrm{e}$ is not possible in general, and one has to resort to approximation techniques to make the problem computationally tractable. Existing algorithms providing such approximations use either set inversion techniques based on (4), or global optimization methods based on (5). The focus in the remainder of the paper is on the former.

B. Algorithmic Procedure

We consider a variant of the Set Inversion Via Interval Analysis (SIVIA) algorithm by [1] in order to approximate as closely as possible the solution set $\boldsymbol{P}_\mathrm{e}$. Let $\boldsymbol{F}_t$ denote the

mapping associated to the ODE model (1), such that

$$\forall \boldsymbol{p} \in \boldsymbol{P}_0 : \ \boldsymbol{F}_t(\boldsymbol{p}) = \hat{\boldsymbol{y}}(t), \tag{7}$$

for each $t \in [t_0, t_N]$. It follows that characterizing $\boldsymbol{P}_{\rm e}$ via (4) is equivalent to intersecting the inverse images $\boldsymbol{F}_{t_i}^{-1}(\boldsymbol{Y}_i)$ of $\boldsymbol{Y}_i$ via $\boldsymbol{F}_{t_i}$ for each $i = 1, \ldots, N$,

$$\boldsymbol{P}_{\rm e} = \left(\bigcap_{i=1}^{N} \boldsymbol{F}_{t_i}^{-1}(\boldsymbol{Y}_i) \right) \cap \boldsymbol{P}_0. \tag{8}$$

The algorithm proceeds as follows:

Input: Termination tolerances $\epsilon_{\rm box} > 0$ and $\epsilon_{\rm bnd} > 0$

Initialization: Set partitions $\mathbb{P}_{\rm bnd} = \{\boldsymbol{P}_0\}$, $\mathbb{P}_{\rm int} = \mathbb{P}_{\rm out} = \emptyset$; Set iteration count $k = 0$

Main Loop:

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

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

An illustration of parameter subboxes belonging to the partitions $\mathbb{P}_{\rm int}$, $\mathbb{P}_{\rm bnd}$, and $\mathbb{P}_{\rm out}$ is shown on the bottom plot in Figure 1, together with the corresponding output trajectories on the top plot using the same color code. Upon termination, this algorithm returns partitions $\mathbb{P}_{\rm int}$ and $\mathbb{P}_{\rm bnd}$ such that

$$\bigcup_{\boldsymbol{P} \in \mathbb{P}_{\rm int}} \boldsymbol{P} \ \subseteq \ \boldsymbol{P}_{\rm e} \ \subseteq \ \left(\bigcup_{\boldsymbol{P} \in \mathbb{P}_{\rm int}} \boldsymbol{P} \right) \cup \left(\bigcup_{\boldsymbol{P} \in \mathbb{P}_{\rm bnd}} \boldsymbol{P} \right). \tag{9}$$

A number of remarks are in order:

- Multiple heuristics can be used regarding the selection of a parameter box in step 1 or the bisection in step 3c. Here, we select the widest parameter box in step 1 and bisect at the mid-point along the least reduced axis in step 3c.
- Step 2 involves bounding the solution set of the parametric ODEs (1) for the current parameter box $\boldsymbol{P}$. Note that the algorithm will terminate finitely if the output enclosures $\hat{\boldsymbol{Y}}(t_i)$ shrink as $\|\boldsymbol{P}\| \to 0$. This is the case for all of the bounding techniques detailed later on in Sect. III.
- Test 4b is an addition to the original SIVIA algorithm [1], which interrupts the iterations when a specified level of approximation of the solution set $\boldsymbol{P}_{\rm e}$ is reached. The level of approximation is measured here as the total volume $V_{\rm bnd}$ of the boxes in the partition $\mathbb{P}_{\rm bnd}$, with corresponding threshold $\epsilon_{\rm bnd}$. In contrast, stopping the algorithm when a minimum box width is reached does not give any guarantee on the approximation level because of the overestimation in step 2.

Variants of this algorithm exist that improve its convergence rate by introducing additional exclusion tests. One such test involves checking whether the interval gradient of the objective function in (5) for a given subbox $\boldsymbol{P}$ does not contain $\boldsymbol{0}$, in which case $\boldsymbol{P} \in \mathbb{P}_{\rm out}$—this is because there cannot exist any global optimizer of (5) in $\boldsymbol{P}$ in this case. Interval contractors based on interval gradients were also proposed in [2]. The downside of these heuristics is the need to compute enclosures of the first-order sensitivities of model (1), which can cause a significant computational overhead.

III. Bounding Parametric ODEs

This section describes various techniques for enclosing the solution set of parametric ODEs, as needed in step 2 of the guaranteed parameter estimation algorithm. The main difficulty is to compute bounds on the state trajectories $\boldsymbol{x}$, since bounds on the output trajectories $\boldsymbol{y}$ are easily computed from the state and parameter bounds when the function $\boldsymbol{g}$ is tree-decomposable.

A. Interval Bounds

It is well known that interval bounds for (1) can be computed by application of the following classical result from the theory of differential inequalities [3], [4].

Theorem 1: Consider the parametric ODEs (1), where $\boldsymbol{f} : \boldsymbol{D} \times \boldsymbol{P} \to \mathbb{R}^{n_x}$ is a continuous vector function and satisfy a uniqueness condition on $\boldsymbol{D} \times \boldsymbol{P}$, with $\boldsymbol{D} \subset \mathbb{R}^{n_x}$. Let the functions $\boldsymbol{x}^L, \boldsymbol{x}^U : \mathbb{R} \to \mathbb{R}^{n_x}$ be continuous on some open set containing $[t_0, t_N]$ and satisfy $[\boldsymbol{x}^L(t), \boldsymbol{x}^U(t)] \subset \boldsymbol{D}$ for all $t \in [t_0, t_N]$. If $\boldsymbol{x}^L(t_0) \leq \boldsymbol{h}(\boldsymbol{p}) \leq \boldsymbol{x}^U(t_0)$ for all $\boldsymbol{p} \in \boldsymbol{P}$ and

$$\dot{x}_i^L(t) \leq \min \left\{ f_i(\boldsymbol{z}, \boldsymbol{p}) \mid \boldsymbol{p} \in \boldsymbol{P}, \boldsymbol{z} \in [\boldsymbol{x}^L, \boldsymbol{x}^U], z_i = x_i^L(t) \right\}, \tag{10}$$

$$\dot{x}_i^U(t) \geq \max \left\{ f_i(\boldsymbol{z}, \boldsymbol{p}) \mid \boldsymbol{p} \in \boldsymbol{P}, \boldsymbol{z} \in [\boldsymbol{x}^L, \boldsymbol{x}^U], z_i = x_i^U(t) \right\}, \tag{11}$$

for almost all $t \in [t_0, t_N]$ and $i = 1, \ldots, n_x$, then $\boldsymbol{x}(t) \in [\boldsymbol{x}^L(t), \boldsymbol{x}^U(t)]$ for all $(t, \boldsymbol{p}) \in [t_0, t_N] \times \boldsymbol{P}$. ◇

In practice, bounds on the right-hand sides of the differential inequalities can be obtained via natural interval extensions. Alternatively, centered forms [11] can be used to obtain tighter bounds on $\hat{\boldsymbol{Y}}$. Although simple to implement, the bounds obtained with the method of differential inequalities only converge to the exact state bounds at a linear rate as $\|P\| \to 0$ in general. This motivates the use of Taylor model-based approaches, which enjoy higher-order convergence.

B. Taylor Model-based Bounds

Given a non-empty set $\boldsymbol{P} \in \mathbb{R}^{n_p}$ and a $\mathcal{C}^{q+1}$ function $\phi : \boldsymbol{X} \to \mathbb{R}$, with $q \geq 0$, the pair $(\mathcal{P}^q_{\phi,\boldsymbol{P}}, \mathcal{R}^q_{\phi,\boldsymbol{P}})$ is called a *qth-order Taylor model of* ϕ *on* $\boldsymbol{P}$ if [12]

- the n_p-variate polynomial $\mathcal{P}^q_{\phi,\boldsymbol{P}}$ is such that

$$\forall \boldsymbol{p} \in \boldsymbol{P}: \ \mathcal{P}^q_{\phi,\boldsymbol{P}}(\boldsymbol{p}) = \sum_{\substack{\boldsymbol{\kappa}\in\mathbb{N}^{n_p},\\ |\boldsymbol{\kappa}|\leq q}} \frac{\partial^{\boldsymbol{\kappa}}\phi(\boldsymbol{p}^*)}{\boldsymbol{\kappa}!}(\boldsymbol{p}-\boldsymbol{p}^*)^{\boldsymbol{\kappa}}, \tag{12}$$

  for some $\boldsymbol{p}^* \in \boldsymbol{P}$, with multi-index notation used;
- the remainder interval $\mathcal{R}^q_{\phi,Y}$ is such that

$$\forall \boldsymbol{p} \in \boldsymbol{P}: \ \phi(\boldsymbol{p}) - \mathcal{P}^q_{\phi,\boldsymbol{P}}(\boldsymbol{p}) \in \mathcal{R}^q_{\phi,\boldsymbol{P}}. \tag{13}$$

Similar to interval analysis, Taylor models can be constructed recursively for factorable functions, which are defined by a finite recursive composition of binary sums, binary products, and a given library of univariate intrinsic functions such as $\exp(\cdot)$, $\sqrt{\cdot}$, etc. This recursive procedure is initiated with a known Taylor model, which can be the Taylor model of a variable or, in the case of a composite function, a Taylor model of the inner function.

An extension of the theory of differential inequalities was recently proposed in [10], which allows their combination with Taylor models.

Theorem 2: Consider the parametric ODEs (1), where $\boldsymbol{f}: \boldsymbol{D}\times\boldsymbol{P} \to \mathbb{R}^{n_x}$ is a $\mathcal{C}^{q+1}$ vector function, with $\boldsymbol{D} \subset \mathbb{R}^{n_x}$ and $q \geq 0$. For each $i = 1,\ldots,n_x$ and each $t \in [t_0,t_N]$, let $\mathcal{P}^q_{x_i(t),\boldsymbol{P}}: \boldsymbol{P} \to \mathbb{R}$ be the n_p-variate polynomial of order q matching the truncated Taylor expansion of $x_i(t)$ at some $\boldsymbol{p}^* \in \boldsymbol{P}$. Let also $r^L, r^U: \mathbb{R} \to \mathbb{R}^{n_x}$ be differentiable functions on some open set containing $[t_0,t_N]$, which satisfy $\mathcal{P}^q_{x_i(t),\boldsymbol{P}}(\boldsymbol{p})+[r^L(t),r^U(t)] \subset \boldsymbol{D}$ for all $(t,\boldsymbol{p}) \in [t_0,t_N]\times\boldsymbol{P}$. If $r^L(t_0) \leq \boldsymbol{h}(\boldsymbol{p}) - \mathcal{P}^q_{\boldsymbol{x}(t_0),\boldsymbol{P}}(\boldsymbol{p}) \leq r^U(t_0)$ for all $\boldsymbol{p} \in \boldsymbol{P}$ and

$$\begin{aligned}\dot{r}^L_i(t) \leq \ \min\Big\{ f_i\left(\mathcal{P}^q_{x_i(t),\boldsymbol{P}}(\boldsymbol{p})+\boldsymbol{z},\boldsymbol{p}\right) - \dot{\mathcal{P}}^q_{x_i(t),\boldsymbol{P}}(\boldsymbol{p}) \mid \\ \boldsymbol{p}\in\boldsymbol{P}, \boldsymbol{z}\in[r^L(t),r^U(t)], z_i = r^L_i(t)\Big\},\end{aligned} \tag{14}$$

$$\begin{aligned}\dot{r}^U_i(t) \geq \ \max\Big\{ f_i\left(\mathcal{P}^q_{x_i(t),\boldsymbol{P}}(\boldsymbol{p})+\boldsymbol{z},\boldsymbol{p}\right) - \dot{\mathcal{P}}^q_{x_i(t),\boldsymbol{P}}(\boldsymbol{p}) \mid \\ \boldsymbol{p}\in\boldsymbol{P}, \boldsymbol{z}\in[r^L(t),r^U(t)], z_i = r^U_i(t)\Big\},\end{aligned} \tag{15}$$

for almost every $t \in [t_0,t_N]$ and all $i = 1,\ldots,n_x$, then $(\mathcal{P}^q_{\boldsymbol{x}(t),\boldsymbol{P}}, [r^L(t),r^U(t)])$ is a qth-order Taylor model of $\boldsymbol{x}(t)$ on $\boldsymbol{P}$ for all $t \in [t_0,t_N]$. ◇

Taylor model methods are appealing in that the overestimation in the remainder term $\mathcal{R}^q_{x_i(t),\boldsymbol{P}}$ converges to zero at order (at least) $q+1$ as $\|P\| \to 0$ [12]. However, in order to derive interval bounds from a Taylor model, the polynomial part must be bounded. Because computing the exact range of a multivariate polynomial is $\mathcal{NP}$ hard, approximate range bounders are used in practice; for instance, the linear and diagonal quadratic terms can be bounded exactly, and the remaining terms estimated using natural interval extensions, as proposed in [6].

C. Ellipsoidal-based Bounds

Given a positive semi-definite matrix $\boldsymbol{Q} \in \mathbb{S}^{n_x}_+$ and a vector $\boldsymbol{c} \in \mathbb{R}^{n_x}$, the set

$$\mathcal{E}(\boldsymbol{Q},\boldsymbol{c}) := \{\boldsymbol{c} + \boldsymbol{Q}^{\frac{1}{2}}\boldsymbol{v} \mid \exists \boldsymbol{v} \in \mathbb{R}^{n_x}: \boldsymbol{v}^{\mathsf{T}}\boldsymbol{v} \leq 1\}, \tag{16}$$

defines an n_x-dimensional ellipsoid centered at $\boldsymbol{c}$ and with shape matrix $\boldsymbol{Q}$. A technique for computing ellipsoidal enclosures for the solution set of nonlinear parametric ODEs was recently proposed by [8].

Theorem 3: Consider the parametric ODEs (1), where $\boldsymbol{f}: \boldsymbol{D}\times\boldsymbol{P} \to \mathbb{R}^{n_x}$ is a $\mathcal{C}^2$ vector function, with $\boldsymbol{D} \subset \mathbb{R}^{n_x}$. Let $\boldsymbol{Q}: [t_0,t_N] \to \mathbb{S}^{n_x}_+$, $\boldsymbol{S}: [t_0,t_N] \to \mathbb{R}^{n_x\times n_p}_+$ and $\boldsymbol{x}^*: [t_0,t_N] \to \mathbb{R}^{n_x}$, and define $\boldsymbol{X}(t)$ such that

$$X_i(t) = \sqrt{Q_{ii}}[-1,1] + \boldsymbol{S}_i(t)[\boldsymbol{P}-\boldsymbol{p}^*] + x^*_i(t), \tag{17}$$

for each $i = 1,\ldots,n_x$ and all $t \in [t_0,t_N]$. Suppose that $\boldsymbol{Q}$, $\boldsymbol{S}$ and $\boldsymbol{x}^*$ satisfy

$$\begin{aligned}\dot{\boldsymbol{Q}}(t) \succeq \ & \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}(\boldsymbol{x}^*(t),\boldsymbol{p}^*)\,\boldsymbol{Q}(t) + \boldsymbol{Q}(t)\,\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}(\boldsymbol{x}^*(t),\boldsymbol{p}^*)^{\mathsf{T}} \\ & + \frac{\boldsymbol{1}^{\mathsf{T}}\mathrm{rad}\boldsymbol{R}_t}{\sqrt{\mathrm{tr}\boldsymbol{Q}(t)}}\,\boldsymbol{Q}(t) + \mathrm{diag}[\mathrm{rad}\boldsymbol{R}_t]\,\sqrt{\mathrm{tr}\boldsymbol{Q}(t)},\end{aligned} \tag{18}$$

$$\dot{\boldsymbol{S}}(t) = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}(\boldsymbol{x}^*(t),\boldsymbol{p}^*)\,\boldsymbol{S}(t) + \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{p}}(\boldsymbol{x}^*(t),\boldsymbol{p}^*), \tag{19}$$

$$\dot{\boldsymbol{x}}^*(t) = \boldsymbol{f}(\boldsymbol{x}^*(t),\boldsymbol{p}^*), \tag{20}$$

for almost every $t \in [t_0,t_N]$, as well as the initial conditions $\boldsymbol{Q}(t_0) = \mathrm{diag}[\mathrm{rad}\boldsymbol{R}_0]^2$, $\boldsymbol{S}(t_0) = \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{p}}(\boldsymbol{p}^*)$ and $\boldsymbol{x}(t_0) = \boldsymbol{h}(\boldsymbol{p}^*)$, where $\boldsymbol{R}_t$ and $\boldsymbol{R}_0$ denote, respectively, the remainder term in a 1st-order Taylor model of $\boldsymbol{f}$ on $\boldsymbol{X}(t)\times\boldsymbol{P}$ at $(\boldsymbol{x}^*(t),\boldsymbol{p}^*)$ and the remainder term in a 1st-order Taylor model of $\boldsymbol{h}$ on $\boldsymbol{P}$ at $\boldsymbol{p}^*$. Then, $\boldsymbol{x}(t) \in \boldsymbol{X}(t)$ for all $(t,\boldsymbol{p}) \in [t_0,t_N]\times\boldsymbol{P}$. ◇

The convergence rate of such ellipsoidal bounds to the exact state bounds is typically quadratic as $\|P\| \to 0$, and they are advantageous in terms of computational overhead compared with Taylor model-based bounds. A comparison of these three bounding techniques in a context of guaranteed parameter estimation is presented in the next section for a simple case study.

IV. Case Study

A. Problem Definition

A dynamic model involving two state variables $\boldsymbol{x} = (x_1,x_2)^{\mathsf{T}}$ and three uncertain model parameters $\boldsymbol{p} = (p_1,p_2,p_3)^{\mathsf{T}} \in [0.01,1]^3$ is considered [2]:

$$\dot{x}_1(t) = -(p_1+p_3)x_1(t) + p_2x_2(t), \quad x_1(0) = 1, \tag{21a}$$

$$\dot{x}_2(t) = p_1x_1(t) - p_2x_2(t), \quad x_2(0) = 0. \tag{21b}$$

The system has a single output variable $\hat{y}$, which corresponds to the state variable x_2, $\hat{y}(t) = x_2(t)$, with $N = 15$ measurements corresponding to the times $t_i = 1,\ldots,15$. Synthetic experimental data are generated by simulating the model (21) with parameter values $\boldsymbol{p}^0 = (0.6,0.15,0.35)^{\mathsf{T}}$, and then rounding the output $y(t_i)$ up or down to the nearest value by retaining two significant digits only.

B. Numerical Implementation

The guaranteed parameter estimation algorithm in Sect. II-B is implemented in a C++ program, which uses the PROFIL/BIAS library [13] for (validated) interval computations and the MC++ library (`http://www3.imperial.`

ac.uk/people/b.chachuat/research) 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. III. All of the numerical results presented next use the explicit embedded Runge-Kutta-Fehlberg (4,5) method, with both the relative and absolute tolerances set to 1×10^{-9}. All the reported results are obtained on workstation Intel Xeon CPU X5660 with 2.80 GHz and 16 GB RAM.

Exclusion tests based on gradient information (see Sect. II-B) are also implemented as a means to enhance the convergence of the algorithm. However, this was found to have an adverse effect on the overall execution time due to the need for bounding first-order sensitivity equations of the dynamic model (21). Further research is clearly warranted in order to reduce the associated computational overhead, possibly by using adjoint-based sensitivity. In particular, the results reported hereafter do not consider such gradient-based exclusion tests.

C. Comparative Analysis

The algorithm is run with ODE bounding techniques based on classical differential inequalities, ellipsoidal techniques, and differential inequalities with Taylor models, for Taylor model orders of $q = 1, \ldots, 4$. In order to allow for fair comparisons, termination criteria are defined in terms of the level of accuracy $\epsilon_{\rm bnd}$ of the solution set in the range 5×10^{-6} to 1×10^{-3}, while termination criteria in terms of the minimum box size $\epsilon_{\rm box}$ are set to 0.

The number of iterations and CPU time for the various methods are reported on the top and bottom plots of Figure 2, respectively, as a function of the required accuracy $\epsilon_{\rm bnd}$. It is evident that classical differential inequalities require by far the largest number of iterations at any accuracy level. Despite this large number of iterations, this method remains the fastest for accuracies $\epsilon_{\rm bnd} > 10^{-4}$ mainly due to its simplicity. At higher accuracy levels, bounding techniques based on ellipsoidal calculus and Taylor model arithmetic, which exhibit faster convergence rates, are seen to become competitive. At an accuracy level of $\epsilon_{\rm bnd} = 1 \times 10^{-5}$, for instance, the use of classical differential inequalities requires over 900,000 iterations and 15 hours to terminate; moreover, a case with $\epsilon_{\rm bnd} = 5 \times 10^{-6}$ could not be run till completion with this bounding technique.

A significant reduction in the number of iterations is obtained when differential inequalities with Taylor models are used. As expected, the reduction gets larger as the Taylor model order increases. However, the observed trend is also that the reduction becomes marginal as Taylor model orders greater than $q = 3$ are used. At the same time, it is evident from the bottom plot of Figure 2 that higher-order Taylor model incur a significant computational overhead. In terms of the overall computational efficiency, the best trade-off is obtained for $q = 2$ here. Note that the performance with 1st-order Taylor models is comparable, even though the number

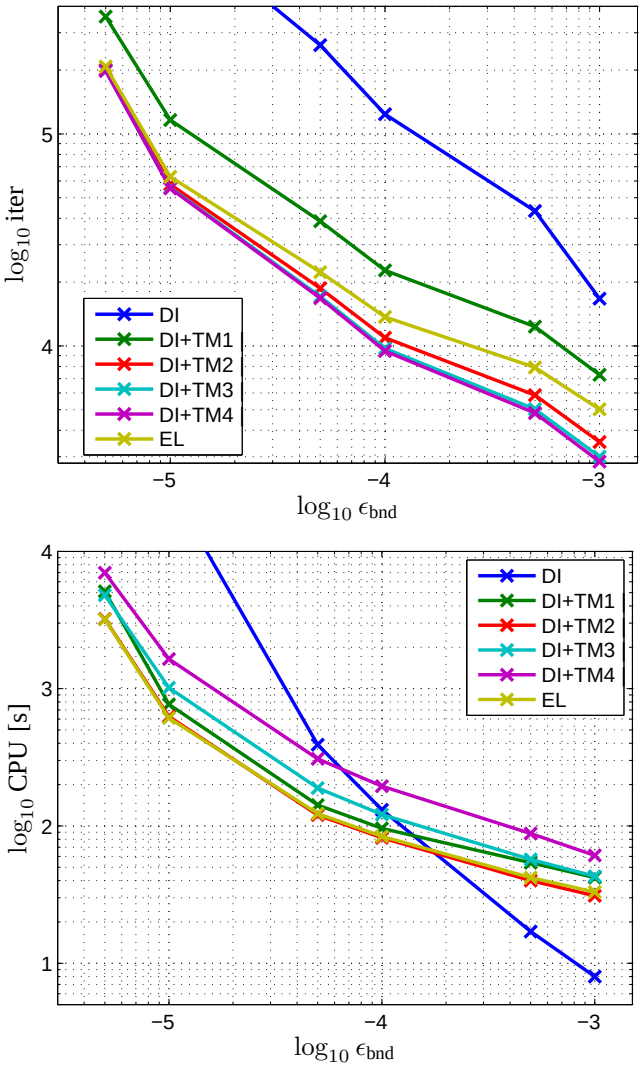

Fig. 2. Comparison of guaranteed parameter estimation algorithm for various bounding techniques. *Top:* Number of iterations vs. convergence threshold. *Bottom:* CPU time vs. convergence threshold.

of iterations is much larger compared to 2nd-order Taylor models.

The performance with ellipsoidal techniques is also comparable with that of 2nd-order Taylor model, both in terms of computational effort and number of iterations. This suggests a good potential for improvement of calculating bounds on the solutions of parametric ODEs by combining these two techniques. A bounding of Taylor model remainders using ellipsoidal calculus might be investigated in the future work.

The top and bottom plots in Figure 3 show the projections of the approximate solution set—$\mathbb{P}_{\rm int}$ and $\mathbb{P}_{\rm bnd}$ are shown using the same color code as in Fig. 1—on the (p_1, p_2) and (p_2, p_3) subspaces, respectively, for different levels of accuracy $\epsilon_{\rm bnd}$. Note first that, for each reported $\epsilon_{\rm bnd}$, the parameter values $\boldsymbol{p}^0$ used to generate the experimental data are part of the solution set. The solution set for this problem turns out to be disconnected, thereby suggesting a potential structural identifiability problem. Interestingly, this non-

connectedness of the solution sets can only be detected when the accuracy level ϵ_{bnd} is 5×10^{-5} or less, which again calls for the use of Taylor model and ellipsoidal techniques whose convergence rate is faster. More generally, these results also support the use of guaranteed parameter estimation as a tool for analyzing structural identifiability. Another interesting observation is that $\mathbb{P}_{\text{int}}$ remains empty when the threshold level ϵ_{bnd} is greater than 1×10^{-6}.

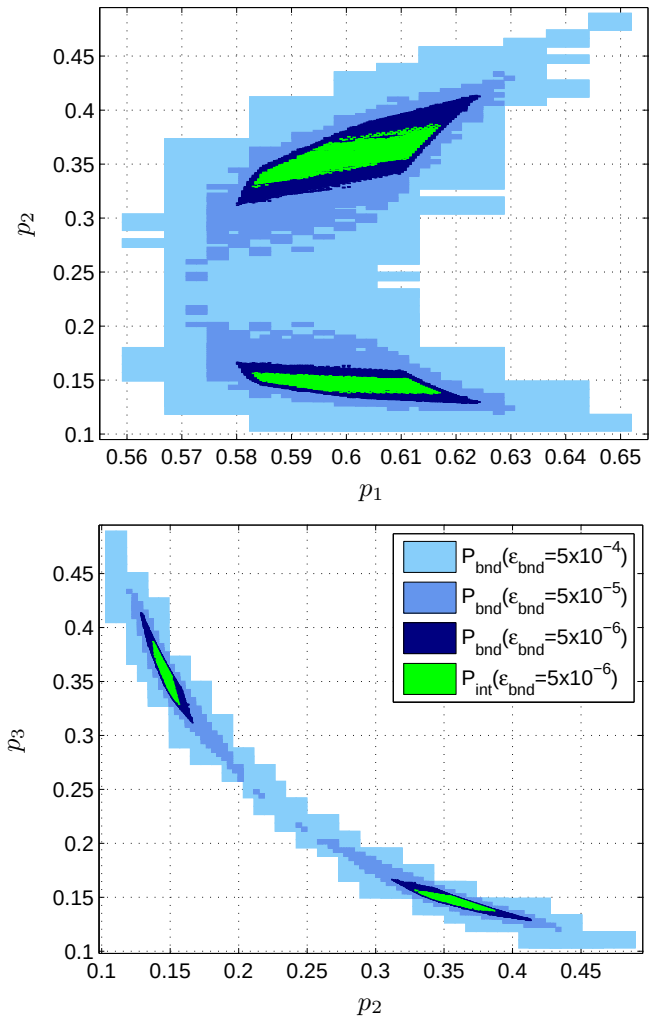

Fig. 3. Guaranteed parameter set for the case study ($\epsilon_{\text{bnd}} = 5 \times 10^{-6}$. *Top:* Projections into (p_1, p_2) space (top plot) and (p_2, p_3) space (bottom plot).

V. Conclusions

This paper has considered the problem of guaranteed parameter estimation, where one seeks for all parameter values of a dynamic model that are consistent with some experimental data, within specified error bounds. The focus has been on set inversion techniques based on exhaustive search, for which the ability to compute tight bounds on the solution set of the dynamic model is critical. Various bounding techniques based on interval analysis, ellipsoidal calculus and Taylor model arithmetic have been compared on a case study. Due to its simplicity, the former proved to be the fastest for obtaining coarse approximations of the solution set. On the other hand, the latter two techniques were found to greatly reduce the number of iterations and the overall computational time when more accurate approximations are sought. From these results, which ought to be confirmed on more complex parameter estimation problems, the development of hybrid approaches combining various bounding techniques appears promising. In order to further enhance the convergence of these algorithms, the application of efficient domain reduction techniques and other heuristics from the area of branch-and-bound in global optimization will also be investigated as a part of future work.

Acknowledgments

The authors are thankful to Dr. Boris Houska for useful suggestions. RP and MF gratefully acknowledge the contribution of the Scientific Grant Agency of the Slovak Republic under the grant 1/0053/13 and the Slovak Research and Development Agency under the project APVV-0551-11. MV and BC acknowledge financial support from Marie Curie grant PCIG09-GA-2011-293953. MV is also grateful to CONACYT for a doctoral scholarship.

References

[1] L. Jaulin and E. Walter, "Set inversion via interval analysis for nonlinear bounded-error estimation," *Automatica*, vol. 29, no. 4, pp. 1053–1064, 1993.

[2] M. Kieffer and E. Walter, "Guaranteed estimation of the parameters of nonlinear continuous-time models: contributions of interval analysis," *International Journal of Adaptive Control and Signal Processing*, vol. 25, no. 3, pp. 191 – 207, 2011.

[3] W. Walter, *Differential and integral inequalities*. Springer-Verlag, 1970.

[4] A. Singer and P. Barton, "Bounding the solutions of parameter dependent nonlinear ordinary differential equations," *SIAM Journal on Scientic Computing*, vol. 27, no. 6, pp. 2167–2182, 2006.

[5] M. Berz and G. Hoffstaetter, "Computation and application of Taylor polynomials with interval remainder bounds," *Reliable Computing*, vol. 4, no. 1, pp. 83–97, 1998.

[6] Y. Lin and M. A. Stadtherr, "Validated solutions of initial value problems for parametric ODEs," *Appl. Numer. Math.*, vol. 57, no. 10, pp. 1145–1162, 2007.

[7] A. Sahlodin and B. Chachuat, "Convex/concave relaxations of parametric ODEs using Taylor models," *Computers & Chemical Engineering*, vol. 35, no. 5, pp. 844 – 857, 2011.

[8] B. Houska, F. Logist, J. Van Impe, and M. Diehl, "Robust optimization of nonlinear dynamic systems with application to a jacketed tubular reactor," *Journal of Process Control*, vol. 22, no. 6, pp. 1152–1160, 2012.

[9] Y. Lin and M. A. Stadtherr, "Guaranteed state and parameter estimation for nonlinear continuous-time systems with bounded-error measurements," *Industrial & Engineering Chemistry Research*, vol. 46, no. 22, pp. 7198–7207, 2007.

[10] B. Chachuat and M. Villanueva, "Bounding the solutions of parametric ODEs: when Taylor models meet differential inequalities," in *22nd European Symposium on Computer Aided Process Engineering*, vol. 30. Elsevier, 2012, pp. 1307–1311.

[11] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter, *Applied Interval Analysis*. Springer Verlag, 2001.

[12] A. Bompadre, A. Mitsos, and B. Chachuat, "Convergence analysis of Taylor and McCormick-Taylor models," *Journal of Global Optimization*, in press (DOI:10.1007/s10898-012-9998-9).

[13] O. Knüppel, "PROFIL/BIAS – A fast interval library," *Computing*, vol. 53, pp. 277–287, 1994.