Guaranteed bounded-error parameter estimation for nonlinear models with uncertain experimental factors

Luc Jaulin, Eric Walter

Guaranteed Parameter Bounding for Nonlinear Models with Uncertain Experimental Factors

L. Jaulin ◇ and E. Walter ⋆

◇ Laboratoire d'Ingénierie des Systèmes Automatisés, Université d'Angers,

Faculté des Sciences, 2 boulevard Lavoisier, 49045 Angers, France

⋆ Laboratoire des Signaux et Systèmes, CNRS-Supélec,

Plateau de Moulon, 91192 Gif-sur-Yvette, France

October 14, 1998

Abstract

In the context of bounded-error estimation, it is customary to assume that the error between the model output and output data should lie between some known prior bounds. In this paper, it is also assumed that the factors characterizing the experiments that have been carried out (e.g., measurement times) are uncertain, with known prior bounds. An algorithm based on interval analysis is used to characterize the set of all values of the parameter vector to be estimated that are consistent with these hypotheses. This is performed in a guaranteed way, even when the model output is a nonlinear function of the parameters and factors characterizing the experiments.

Key Words: Bounded-error estimation; errors-in-variables; estimation theory; guaranteed estimation; identification algorithms; interval analysis; nonlinear estimation; set inversion.

Any correspondence regarding this paper should be sent to Eric Walter, Laboratoire des Signaux et Systèmes, Supélec, Plateau de Moulon, 91192 Gif-sur-Yvette Cedex, FRANCE.
Phone: (33) 1 69 85 17 21
Fax: (33) 1 69 41 30 60
Email: walter@lss.supelec.fr

Suggested running title: Parameter Bounding with Uncertain Factors

1. Introduction

Bounded-error parametric estimation (or parameter bounding) has received a renewed attention in the last decade, see, *e.g.*, (Walter, 1990; Combettes, 1993; Deller *et al.*, 1993; Norton, 1994, 1995; Milanese *et al.*, 1996; Walter and Pronzato, 1997) and the references therein. Of the reasons for this interest, we shall quote only two. First, the approach can deal with deterministic structural errors, not adequately described by random variables. Second, it is well suited to the guaranteed characterization of parameter uncertainty, a prerequisite for a number of methods in robust control. It is customary to assume that the errors between the model output and output data should lie between some known prior bounds and to try to characterize the set of all values of the parameter vector that are consistent with this hypothesis. Most often, no other type of error is taken into account, with the notable exception of the

errors-in-variables approach, mostly developed in the context of linear estimation, where all components of the regressor vector are assumed to be uncertain. In this paper, we deal with models whose output may be nonlinear in their parameters. It is assumed that the factors characterizing the experimental conditions under which the output data have been collected are uncertain, and that prior bounds are available on their possible values. The simplest example of such a situation is when the model output depends on time and each measurement time t_i is only assumed to belong to some known interval T_i, but the same approach applies if the measurements depend on a space variable or on values of some input factors. Let $\mathbf{t} \in \mathbb{R}^{n_t}$ be the vector of all experimental factors on which the collected output data depend. Our aim is to take into account uncertainty about the values taken by $\mathbf{t}$ and to characterize the set of all values of the parameter vector $\mathbf{p} \in \mathbb{R}^{n_p}$ that are consistent with the available information. This will be performed in a global and guaranteed way by employing the algorithm SIVIA (Set Inverter Via Interval Analysis) introduced in (Jaulin and Walter, 1993a, 1993b). In the context of uncertain experimental factors, proving that a given box in parameter space is inside or outside the solution set (a task that is at the core of SIVIA) will require specific developments. The paper is organized as follows. The next section states the problem to be considered in mathematical terms. A new version of SIVIA adapted to the present context is described in Section 3. It involves two routines INSIDE and OUTSIDE, which are based on tools briefly presented in Section 4. The convergence of the algorithm is analyzed in Section 5 before reporting the results obtained on a test case in Section 6.

2. Problem statement

The model to be considered is some algorithm to compute the model output $\mathbf{y}_m \in \mathbb{R}^{n_\mathrm{y}}$ as a function of $\mathbf{p}$ and $\mathbf{t}$. We assume that the n_p-dimensional parameter vector $\mathbf{p}$ is consistent with the data collected during the ith elementary experiment ($i = 1, ..., m$) if and only if $\mathbf{y}_m(\mathbf{p}, \mathbf{t})$ belongs to $\mathcal{Y}_i$ for some $\mathbf{t}$ in $\mathcal{T}_i$, where $\mathcal{Y}_i$ and $\mathcal{T}_i$ are known sets. $\mathcal{Y}_i$ and $\mathcal{T}_i$ respectively characterize the uncertainty about the ith output data and associated experimental factors; they may be derived from point measurements and bounds on the acceptable errors on these measurements. Bounded-error estimation then aims at characterizing the *posterior feasible set* $\mathcal{S}$, defined as the set of all $\mathbf{p}$'s consistent with all data and given by

$$\mathcal{S} = \{\mathbf{p} \in \mathbb{R}^{n_\mathrm{p}} \mid \forall i \in \{1, ..., \ m\}, \exists \mathbf{t} \in \mathcal{T}_i \mid \mathbf{y}_m(\mathbf{p}, \mathbf{t}) \in \mathcal{Y}_i\}. \tag{2.1}$$

Characterizing $\mathcal{S}$ can be cast into the framework of *errors-in-variable* estimation, to which relatively little attention has been paid in the literature about parameter bounding. When the model output is linear in $\mathbf{p}$, it can be written as

$$\mathbf{y}_m(\mathbf{p}, \mathbf{t}) = \mathbf{R}(\mathbf{t})\mathbf{p}, \tag{2.2}$$

where the matrix of regressors $\mathbf{R}(\mathbf{t})$ is independent of $\mathbf{p}$. If, moreover, $\mathbf{y}_m$ is scalar and $\mathbf{R}(\mathbf{t}) = \mathbf{t}^T$, i.e., each component of the regressor vector corresponds to an uncertain experimental factor, then

$$y_m(\mathbf{p}, \mathbf{t}) = \mathbf{t}^T\mathbf{p}, \tag{2.3}$$

and the problem of characterizing $\mathcal{S}$ can be solved exactly with the technique described in (Cerone, 1991, 1996). Ellipsoidal outer approximations of this set can also

be computed, see, *e.g.*, (Norton, 1987; Clément and Gentil, 1990; Pronzato and Walter, 1994; Veres and Norton, 1996). When the errors on the components of the regressor vector are no longer independent, for instance when dealing with output error models with an autoregressive part, the same approach leads to the construction of sets that can only be guaranteed to contain $\mathcal{S}$. We shall here consider the more general situation where $\mathbf{y}_m(\mathbf{p}, \mathbf{t})$ may be nonlinear in $\mathbf{p}$ and $\mathbf{t}$. To the best of our knowledge, it is the first attempt to provide guaranteed and accurate results in such a nonlinear context. (Estimating parameters in nonlinear models with errors in variables has been considered in the context of least squares for many years, see, *e.g.*, (Schwetlick and Tiller, 1985), but the results are obtained by local methods and thus not guaranteed.) For the sake of simplicity, we shall assume that (i) there is only one experimental factor, which we shall denote by t and call time, although it may have some other meaning; (ii) the system to be considered has only one scalar output at any given time; (iii) the feasible sets $\mathcal{T}_i$ and $\mathcal{Y}_i$ are intervals, denoted by T_i and Y_i, and (iv) the model output $y_m(\mathbf{p}, t)$ is continuous in $\mathbf{p}$ and t. The guaranteed characterization of $\mathcal{S}$, now defined by

$$\mathcal{S} = \{\mathbf{p} \mid \forall i \in \{1, ..., m\}, \exists t \in T_i \mid y_m(\mathbf{p}, t) \in Y_i\}, \tag{2.4}$$

remains a nontrivial problem since deciding whether a given vector $\mathbf{p}$ belongs to $\mathcal{S}$ is already complicated. Define the ith *error function* by

$$e_i(\mathbf{p}, t) = y_m(\mathbf{p}, t) - \text{center}(Y_i), t \in T_i. \tag{2.5}$$

If e_i^{max} denotes the radius of Y_i, the posterior feasible set can also be defined by

$$\mathcal{S} = \{\mathbf{p} \mid \forall i \in \{1, ..., m\}, \exists t \in T_i \text{ such that } |e_i(\mathbf{p}, t)| \leq e_i^{max}\}. \tag{2.6}$$

As an illustrative test case, consider an estimation problem derived from (Milanese and Vicino, 1991) and (Jaulin and Walter, 1993b). Table 1 gives ten pairs (T_i, Y_i) of feasible interval data. In (Milanese and Vicino, 1991) and (Jaulin and Walter, 1993b), the measurement times were assumed to be known exactly. In contrast, it is assumed here that errors of up to ± 1 on the measurement times may have taken place.

Possible location for Table 1.

Figure 1 presents the resulting data. Gray boxes indicate the uncertainties associated with each pair of output and time data.

Possible location for Figure 1.

The model output is

$$y_m(\mathbf{p}, t) = 20 \exp(-p_1 t) - 8 \exp(-p_2 t). \tag{2.7}$$

The set $\mathcal{S}$ to be characterized is the set of all values of $\mathbf{p} = (p_1, p_2)^T$ such that this model output goes through all ten boxes of Figure 1. Treatment of this test case is deferred to Section 6.

3. Set inverter

The algorithm SIVIA presented here is adapted from (Jaulin and Walter, 1993a) to allow the use of more elaborate subroutines in order to decide whether a given box is

included in or excluded from the feasible region, as required by the present context. To describe it, a few definitions are needed. A *box* or *vector interval* $\mathbf{X}$ of $\mathbb{R}^n$ is the Cartesian product of n scalar intervals:

$$\mathbf{X} = \left[x_1^-, x_1^+\right] \times \cdots \times \left[x_n^-, x_n^+\right] = X_1 \times \cdots \times X_n. \tag{3.1}$$

The set of all boxes of $\mathbb{R}^n$ is denoted by $\mathbb{IR}^n$. In the sequel, intervals are written with italic upper-case letters, boxes with bold upper-case letters, and vectors with bold lower-case letters. The *width* $w(\mathbf{X})$ of a box $\mathbf{X}$ is the length of its largest side(s). To *bisect* a box $\mathbf{X}$ means to cut it along a symmetry plane normal to a side of maximum length. This generates two boxes $\mathbf{X}_1$ and $\mathbf{X}_2$ such that $\mathbf{X} = \mathbf{X}_1 \cup \mathbf{X}_2$. A *stack of boxes* is a last-in-first-out list of boxes. The principle of the algorithm is to partition some prior box of interest (or prior feasible set) into three sets of nonoverlapping boxes, namely those that have been proved to be included in $\mathcal{S}$, those that have been proved not to intersect it and those for which nothing could be proved and that are deemed small enough not to be analyzed any further. A basic tool for this purpose is interval analysis, very briefly presented in Section 4.

SIVIA calls two subroutines: INSIDE, which attempts to prove that boxes belong to $\mathcal{S}$, and OUTSIDE, which attempts to prove that boxes have a void intersection with $\mathcal{S}$. Upon completion, SIVIA has generated two subpavings (lists of nonoverlapping boxes). The first one $\mathcal{S}^-(\varepsilon)$ contains all boxes that have been proved to belong to $\mathcal{S}$, and the second one $\Delta\mathcal{S}(\varepsilon)$ all boxes for which nothing could be proved. Both subpavings depend on a positive real ε representing the required accuracy of the characterization. These subpavings provide the following bracketing of the solution set:

$$\mathcal{S}^-(\varepsilon) \subset \mathcal{S} \subset \mathcal{S}^-(\varepsilon) \cup \Delta\mathcal{S}(\varepsilon). \tag{3.2}$$

The dependency in ε will be omitted whenever possible. SIVIA can be summarized as follows:

SIVIA

Input: prior box of interest $\mathbf{P}_0$; required accuracy ε;

Initialization: $Stack = \{\mathbf{P}_0\}$; $\mathcal{S}^- := \emptyset$; $\Delta\mathcal{S} := \emptyset$;

Repeat

Unstack into $\mathbf{P}$;

If (INSIDE($\mathbf{P}$) returns "$\mathbf{P} \subset \mathcal{S}$"), then $\{\mathcal{S}^- := \mathcal{S}^- \cup \mathbf{P}\}$,

else if OUTSIDE($\mathbf{P}$) reports failure,

If $w(\mathbf{P}) > \varepsilon$, then {bisect $\mathbf{P}$; stack the two resulting boxes},

else $\Delta\mathcal{S} := \Delta\mathcal{S} \cup \mathbf{P}$;

Until $Stack = \emptyset$;

Output: $\mathcal{S}^-$, $\Delta\mathcal{S}$.

The subroutine INSIDE aims at proving that a box $\mathbf{P}$ belongs to the posterior feasible set $\mathcal{S}$. In our present context of uncertain measurement times, INSIDE attempts to find m reals $t_i \in T_i$, $i \in \{1, \ldots, m\}$ such that $y_m(\mathbf{P}, t_i) \subset Y_i$. For a given $\mathbf{P}$, the strategy chosen, illustrated by Figure 2, is to take t_i as $\check{t}_i(\text{center}(\mathbf{P}))$, where $\check{t}_i(\cdot)$ is defined by

$$\check{t}_i(\mathbf{p}) = \underset{t \in T_i}{\arg\min}\ (e_i(\mathbf{p}, t))^2. \tag{3.3}$$

Possible location of Figure 2.

An intuitive and partial justification for this strategy is as follows: the smaller

$(e_i(\text{center}(\mathbf{P}), t_i))^2$ is, the more likely it is that $|e_i(\mathbf{P}, t_i)| \leq e_i^{\max}$, or equivalently that $y_m(\mathbf{P}, t_i) \subset Y_i$, will be satisfied.

Remark 1. *This strategy is not necessarily successful. Assume, for instance, that* $m = 1$ *and* $e_1(p, t_1) = \exp(-pt_1)$, *with* $T_1 = [0.5, 1]$, $P = [-1, 3]$ *and* $e_1^{\max} = 2$. *Here* $\check{t}_1(\text{center}(P)) = 1$, *and* $|e_1(P, 1)| = [\exp(-3), \exp(1)]$, *the upper bound of which is larger than* $e_1^{\max}$, *so* P *is not proved to be feasible. On the other hand, taking* $t_1 = 0.5$ *would make it possible to prove that* P *is feasible, as* $|e_1(P, 0.5)| = [\exp(-1.5), \exp(0.5)]$, *the upper bound of which is smaller than* $e_1^{\max}$. *The actual justification for the strategy advocated in this paper is therefore the possibility to prove convergence when it is used.*

INSIDE requires an algorithm for computing $\check{t}_i(\mathbf{p})$ in a guaranteed way and sufficient conditions to prove that, for a given box $\mathbf{P}$ and a given real t, $y_m(\mathbf{P}, t)$ belongs to Y_i. We shall see in Section 4 that such an algorithm can be provided by interval analysis. INSIDE can be summarized as follows:

INSIDE

Input: $\mathbf{P}$;
For $i := 1$ to m do
  Compute $\check{t} := \check{t}_i\,(\text{center}(\mathbf{P}))\,;$
  If it can be proved that $y_m(\mathbf{P}, \check{t}) \subset Y_i$ then next i,
  else return ("Failed");
EndFor;
Return ("$\mathbf{P} \subset \mathcal{S}$").

The subroutine OUTSIDE aims at proving that a box $\mathbf{P}$ does not intersect $\mathcal{S}$. In

our present context of uncertain measurement times, OUTSIDE attempts to take advantage of the following implication

$$\exists\, i \in \{1, ..., m\} \mid y_m(\mathbf{P}, T_i) \cap Y_i = \emptyset \Rightarrow \mathbf{P} \cap \mathcal{S} = \emptyset. \tag{3.4}$$

It can be summarized as follows:

OUTSIDE

Input: $\mathbf{P}$;

For $i := 1$ to m do

Step 1 $Stack := \{T_i\}$;

Step 2 Unstack into T;

Step 3 If it can be proved that $y_m(\mathbf{P}, T) \cap Y_i = \emptyset$, go to Step 6;

Step 4 If $w(T) < w(\mathbf{P})$, next i;

Step 5 Bisect T and stack the two resulting intervals;

Step 6 If $stack \neq \emptyset$, go to Step 2;

Step 7 Return "$\mathbf{P} \cap \mathcal{S} = \emptyset$";

EndFor;

Return ("Failed").

For each $i \in \{1, ..., m\}$, OUTSIDE tries to partition T_i into subintervals such that $y_m(\mathbf{P}, T) \cap Y_i = \emptyset$ for any subinterval T of T_i. If, for at least one i, OUTSIDE succeeds, then $\mathbf{P} \cap \mathcal{S} = \emptyset$. All sub-intervals of T_i still to be studied are stored in a stack. At Step 3, OUTSIDE uses interval analysis, see Section 4, to prove that $y_m(\mathbf{P}, T) \cap Y_i = \emptyset$. The test at Step 4 is introduced to avoid splitting T *ad infinitum* and to introduce some relation with the splitting policy followed for $\mathbf{P}$ by SIVIA.

Remark 2. *INSIDE and OUTSIDE have voluntarily been kept simple to facilitate convergence analysis. Their performance can be improved by adding new tests. For instance, in the* for *loop of OUTSIDE, if* $y_m(center(\mathbf{P}), center(T)) \in Y_i$*, then OUTSIDE will necessarily fail to prove that* $y_m(\mathbf{P}, T) \cap Y_i = \emptyset$*. Between Steps 3 and 4 of OUTSIDE, we may thus add the following step:*

Step 3_{bis} *If* $y_m(center(\mathbf{P}), center(T)) \in Y_i$ *then next* i*.*

4. Interval analysis

Interval analysis is a numerical tool originally developed to quantify the effect of finite-precision floating-point arithmetic on results obtained from a computer (Moore, 1979). At present, it is also used, *e.g.*, to find *all* solutions of a set on nonlinear equations in several indeterminates (Hansen, 1992), to locate *all* global optimizers of a multimodal cost function (Hansen, 1992) and to prove formal inequalities (Moore, 1979). As mentioned in Section 3, INSIDE and OUTSIDE will use interval analysis to provide

- (IA1) a global optimization algorithm for computing $\check{t}_i(\text{center}(\mathbf{P}))$,
- (IA2) sufficient conditions for proving that $y_m(\mathbf{P}, t) \subset Y_i$ and
- (IA3) sufficient conditions for proving that $y_m(\mathbf{P}, T) \cap Y_i = \emptyset$.

A fundamental notion is that of an inclusion function. Consider a continuous function f mapping R^n into R. An interval-valued function F, defined from $\mathbb{IR}^n$ into $\mathbb{IR}$, is

an *inclusion function* of f if:

$$\forall \mathbf{X} \in \mathbb{IR}^n, \quad f(\mathbf{X}) \subset F(\mathbf{X}), \tag{4.1}$$

Since f is continuous, the set $f(\mathbf{X}) = \{f(\mathbf{x}) \mid \mathbf{x} \in \mathbf{X}\}$ is a sub-interval of $F(\mathbf{X})$. In what follows, the inclusion function F will be said to be *convergent* if, for any sequence of boxes $\mathbf{X}(k)$ of $\mathbb{IR}^n$ that converges to a real vector $\mathbf{x}$, the sequence of intervals $F(\mathbf{X}(k))$ converges to the real scalar $f(\mathbf{x})$. The inclusion function F is *inclusion monotonic* if for any $\mathbf{X}$ and $\mathbf{Y}$ in $\mathbb{IR}^n$,

$$\mathbf{X} \subset \mathbf{Y} \Rightarrow F(\mathbf{X}) \subset F(\mathbf{Y}). \tag{4.2}$$

These notions of inclusion monotonicity and convergence are only needed to ensure the convergence of the algorithm to be presented. The computation of a convergent and inclusion-monotonic inclusion function associated with any continuous function defined by an explicit formal expression or by a finite algorithm is in principle very simple (Moore, 1979), and routinely performed by widely available software, see, *e.g.*, (Hammer *et al.*, 1995). Note, however, that rounding errors only make it possible to obtain a *guaranteed* approximation of such an inclusion function in practice.

Example 1. *An inclusion function for the model output given by (2.7) is*

$$Y_m(\mathbf{P}, T) = 20 \exp(-P_1 * T) - 8 \exp(-P_2 * T). \tag{4.3}$$

If, for instance, $T = [2, 4], P_1 = [0, 0.1]$ *and* $P_2 = [1, 1.1]$*, then*

$$\begin{aligned} Y_m(\mathbf{P}, T) &= 20\exp(-[2,4] * [0,0.1]) - 8\exp(-[2,4] * [1,1.1]) \\ &= 20\exp(-[0,0.4]) - 8\exp(-[2,4.4]) \\ &= 20\exp([-0.4,0]) - 8\exp([-4.4,-2]) \\ &= 20[0.67,1] - 8[0.012,0.14] \\ &= [13.4,20] - [0.096,1.12] = [12.2,19.9]. \end{aligned} \tag{4.4}$$

Note that from Table 1 and (4.4), $Y_m(\mathbf{P}, T_4) \cap Y_4 = [12.2, 19.9] \cap [-0.95, 1.15] = \emptyset$*. Therefore, all parameter vectors* $\mathbf{p}$ *in the box* $[0, 0.1] \times [1, 1.1]$ *are unfeasible.*

The algorithm GOP1 (1-dimensional Global OPtimizer), to be presented now, locates a global minimizer of a function $f : R \mapsto R$ over an interval T_0 in a guaranteed way. It is assumed that f is twice differentiable over T_0 and that its second derivative is nonzero at the minimum. GOP1 is a very simple algorithm, sufficient for showing the feasibility of the approach advocated here, but that could not compete with sophisticated interval-based algorithms such as Hansen's (Hansen, 1992), which can deal with a larger class of problems and larger dimensions. GOP1 calls a classical dichotomy routine, able to locate the global minimizer of any inverse-unimodal function over an interval in a guaranteed way. We assume that convergent inclusion functions F, F', F'' are available for f and its first two derivatives f' and f''. $f(\check{t})$ is the lowest value of f known at present, and $\check{t}$ is the corresponding argument. Interval inequalities should be understood as: $X > a \Leftrightarrow x^- > a$ and $X < a \Leftrightarrow x^+ < a$.

GOP1

Inputs: function $f : R \mapsto R$; prior interval of interest T_0;

Initialization: $T := T_0$; $\check{t} :=$ center(T); $Stack = \emptyset$;

Step 1 If $F(T) > f(\check{t})$, go to Step 5;

Step 2 If $(0 \notin F'(T)$ or $F''(T) < 0)$, then

$\{\check{t} :=$arg min$_{\{\check{t},t^-,t^+\}}(f)$; go to Step 5$\}$;

Step 3 If $F''(T) > 0$, then

$\{\check{t}_{dicho} =$ Dichotomy(f, T); $\check{t} :=$arg min$_{\{\check{t},\check{t}_{dicho}\}}(f)$; go to Step 5$\}$;

Step 4 Bisect T and stack the two resulting intervals;

Step 5 If $Stack \neq \emptyset$ then {unstack into T; go to Step 1};

Step 6 Return $\check{t}$ and $f(\check{t})$.

If the condition $F(T) > f(\check{t})$ in Step 1 is satisfied, the current interval T cannot contain any global minimizer. If the condition $(0 \notin F'(T)$ or $F''(T) < 0)$ of Step 2 is satisfied, the global minimizer of f over T is necessarily one of the two extreme values t^- and t^+ of T. On Step 3, if the condition $F''(T) > 0$ holds, f is convex over T and therefore inverse-unimodal, so dichotomy will be able to locate the global minimizer of f over T in a guaranteed way with arbitrary precision. At the end of Step 4, the stack contains all intervals to which global minimizers may still belong.

We now have all the tools required by INSIDE and OUTSIDE that were enumerated at the beginning of Section 4, namely

- (IA1) a way to compute $\check{t}_i($center$(\mathbf{P}))$ in a guaranteed way by calling GOP1 for $f(t) = (e_i($center$(\mathbf{P}), t))^2$ and $T = T_i$;

- (IA2) a sufficient condition for proving that $y_m(\mathbf{P},t) \subset Y_i$, namely $F(\mathbf{P}) \subset Y_i$, where $F(\mathbf{P})$ is an inclusion function for $f(\mathbf{p}) = y_m(\mathbf{p},t)$;
- (IA3) a sufficient condition for proving that $y_m(\mathbf{P},T) \cap Y_i = \emptyset$, namely $Y_m(\mathbf{P},T) \cap Y_i = \emptyset$, where $Y_m(\mathbf{P},T)$ is an inclusion function for $y_m(\mathbf{p},t)$.

The next section will study the convergence of the resulting algorithm when the required accuracy ε tends to zero.

5. Convergence analysis

The algorithm presented in Section 3 generates a bracketing of $\mathcal{S}$ in the sense of (3.2). Recall that $\Delta\mathcal{S}(\varepsilon)$ contains all parameter vectors for which no conclusion could be reached as to their feasibility. The purpose of this section is to study the properties of $\Delta\mathcal{S}(\varepsilon)$ when ε tends to zero. Define

$$\overset{o}{\mathcal{S}} = \{\mathbf{p} \mid \forall i \in \{1,...,m\}, \exists t \in T_i \text{ such that } |e_i(\mathbf{p},t)| < e_i^{max}\}. \tag{5.1}$$

Except in degenerate cases, $\overset{o}{\mathcal{S}}$ is the interior of $\mathcal{S}$. The theorem to be presented shows that if ε is infinitely small, only vectors $\mathbf{p}$ that belong to $\mathcal{S}$ and are outside $\overset{o}{\mathcal{S}}$ will belong to the uncertainty layer $\Delta\mathcal{S}(\varepsilon)$. For such vectors, $y_m(\mathbf{p},t)$ is consistent with all data boxes and tangent (from the outside) to at least one of them.

Theorem 1. *If $Y_m(\mathbf{P},T)$ is an inclusion-monotonic and convergent inclusion function for the model output function $y_m(\mathbf{p},t)$, SIVIA generates $\mathcal{S}^-(\varepsilon)$ and $\Delta\mathcal{S}(\varepsilon)$ such*

that

$$\begin{aligned} &(i) \quad \mathbf{p} \in \overset{o}{\mathcal{S}} \quad \Rightarrow \exists \varepsilon > 0 \mid \mathbf{p} \in \mathcal{S}^{-}(\varepsilon), \\ &(ii) \quad \mathbf{p} \notin \mathcal{S} \quad \Rightarrow \exists \varepsilon > 0 \mid \mathbf{p} \notin \mathcal{S}^{-}(\varepsilon) \cup \Delta\mathcal{S}(\varepsilon). \end{aligned} \tag{5.2}$$

Proof: Part (i): If the contraposite were true, *i.e.*, if there existed $\mathbf{p} \in \overset{o}{\mathcal{S}}$ such that $\forall \varepsilon > 0, \mathbf{p} \in \Delta\mathcal{S}(\varepsilon)$, then for any $\varepsilon > 0$, there would exist a box $\mathbf{P}$ with a width smaller than ε and containing $\mathbf{p}$ such that INSIDE($\mathbf{P}$) fails. This is impossible because of the theorem of Section 5.1 below. Part (ii): If the contraposite were true, *i.e.*, if there existed a vector $\mathbf{p}$ outside $\mathcal{S}$ such that $\forall \varepsilon > 0, \mathbf{p} \in \mathcal{S}^{-}(\varepsilon) \cup \Delta\mathcal{S}(\varepsilon)$, then $\forall \varepsilon > 0$, there would exist a box $\mathbf{P}$ with a width smaller than ε and containing $\mathbf{p}$ such that OUTSIDE($\mathbf{P}$) fails. This is impossible because of the theorem of Section 5.2 below.

5.1. Inclusion test theorem

Theorem 2. *(Inclusion test) If* $\mathbf{p} \in \overset{o}{\mathcal{S}}$, $\exists \varepsilon > 0$ *such that, for any box* $\mathbf{P}$ *containing* $\mathbf{p}$ *with a width smaller than* ε, *INSIDE concludes* $\mathbf{P} \subset \mathcal{S}$. $\diamondsuit$

To prove this theorem, we shall use the three following lemmas. In the sequel, a *cube* is a box with a nonzero volume, all sides of which have the same length. The set of all cubes with center $\mathbf{p}$ is denoted by Cube($\mathbf{p}$).

Lemma 3. *Let* T_0 *be a given interval and* $Z(\mathbf{P}, t)$ *be a convergent inclusion function for the zero function* $z(\mathbf{p}, t) \equiv 0, \forall \mathbf{p} \in R^{n_p}, \forall t \in T_0$. *We have*

$$\forall \varepsilon > 0, \forall \mathbf{p} \in \mathbb{R}^{n_p}, \exists \mathbf{P} \in Cube(\mathbf{p}) \mid \forall t \in T_0, w(Z(\mathbf{P}, t)) < \varepsilon \tag{5.3}$$

$\diamondsuit$

Remark 3. *The interval function $F(\mathbf{P},t)$ is a special case of the classical inclusion function $F(\mathbf{P},T)$, where the width of T is zero. The notion of convergence for $F(\mathbf{P},t)$ thus follows from that for $F(\mathbf{P},T)$: $F(\mathbf{P},t)$ will be said to be convergent if for all sequences of boxes $\mathbf{P}(k) \in \mathbb{IR}^{n_p}$ converging to the vector $\mathbf{p}$ and for all sequences of reals $t(k)$ converging to the real t, the interval sequence $F(\mathbf{P}(k),t(k))$ converges to the real $f(\mathbf{p},t)$. Note that if $F(\mathbf{P},T)$ is a convergent inclusion function, $F(\mathbf{P},t)$ is also convergent.*

Proof The proof is by contradiction. Assume that $\exists \varepsilon > 0$, $\exists \mathbf{p} \in \mathbb{R}^{n_p} \mid \forall \mathbf{P} \in \text{Cube}(\mathbf{p})$, $\exists t \in T_0, w(Z(\mathbf{P},t)) \geq \varepsilon$. Let $\{\mathbf{P}(k), k \in \mathbb{N}\}$ be a sequence of Cube($\mathbf{p}$) converging to $\mathbf{p}$. There exists a sequence of real $\{t(k)\} \in T_0$ such that $w(Z(\mathbf{P}(k),t(k))) \geq \varepsilon$. Since T_0 is a compact set, the infinite sequence $\{t(k)\}$ has at least one accumulation point $\bar{t}$. From $\{t(k)\}$, it is therefore possible to extract a subsequence $\{t(k_j), j \in \mathbb{N}\}$ with $k_j < k_{j+1}$ converging to $\bar{t}$. Since $w(Z(\mathbf{P}(k_j),t(k_j))) \geq \varepsilon$, since $\mathbf{P}(k_j) \rightarrow \mathbf{p}$ and since $t(k_j) \rightarrow \bar{t}$, the inclusion function $Z(\mathbf{P},t)$ cannot be convergent. $\diamondsuit$

Let $E_i(\mathbf{P},t)$ be the inclusion function for $e_i(\mathbf{p},t)$ defined by $E_i(\mathbf{P},t) = Y_m(\mathbf{P},t) -$ center(Y_i). Since $Y_m(\mathbf{P},t)$ is inclusion monotonic and convergent, so is $E_i(\mathbf{P},t)$.

Lemma 4. *For all $i \in \{1,\ldots,m\}$, we have*

$$\forall \varepsilon_i > 0, \forall \mathbf{p} \in \mathbb{R}^{n_p}, \exists \mathbf{P} \in \mathit{Cube}(\mathbf{p}) \mid \forall t \in T_i, ||E_i(\mathbf{P},t)| - |e_i(\mathbf{p},t)|| < \varepsilon_i \qquad (5.4)$$

Proof It is cumbersome but trivial to show that the interval function $Z(\mathbf{P},t) = ||E_i(\mathbf{P},t)| - |e_i(\text{center}(\mathbf{P}),t)||$ is a convergent inclusion function for the zero function $z(\mathbf{p},t) \equiv 0, \mathbf{p} \in \mathbb{R}^{n_\mathrm{p}}, t \in T_i$. From Lemma 3, there exists $\mathbf{P} \in \text{Cube}(\mathbf{p})$ such that $\forall t \in T_i$, $w(Z(\mathbf{P},t)) < \varepsilon_i$. Since $0 \in Z(\mathbf{P},t)$, we have $Z(\mathbf{P},t) < \varepsilon_i$, *i.e.*, $||E_i(\mathbf{P},t)| - |e_i(\mathbf{p},t)|| < \varepsilon_i$. $\diamondsuit$

Lemma 5. *If* $\mathbf{p} \in \overset{o}{\mathcal{S}}$, *then* $\forall i \in \{1,...,m\}$, $\exists \eta_i > 0$ *such that, for any box* $\mathbf{Q}$ *containing* $\mathbf{p}$ *with a width smaller than* η_i, $Y_m(\mathbf{Q}, \check{t}_i(\text{center}(\mathbf{Q}))) \subset Y_i$. $\diamondsuit$

Proof If $\mathbf{p} \in \overset{o}{\mathcal{S}}$, $\forall i \in \{1,...,m\}, \exists t \in T_i$ such that $|e_i(\mathbf{p},t)| < e_i^{max}$ and therefore, the number

$$\varepsilon_i = \frac{1}{3}(e_i^{max} - |e_i(\mathbf{p},t)|) \tag{5.5}$$

is strictly positive. Let $\mathbf{P}$ be a box of $\text{Cube}(\mathbf{p})$ satisfying (5.4) and $\mathbf{Q}$ be a box containing $\mathbf{p}$, with center $\mathbf{q}$ and width smaller than $\eta_i = \frac{w(\mathbf{P})}{2}$. $\mathbf{Q}$ is therefore included in $\mathbf{P}$. Since $\mathbf{q} \in \mathbf{P}$, $e_i(\mathbf{q},t) \in E_i(\mathbf{P},t)$, and Lemma 4 implies

$$|e_i(\mathbf{q},t)| < |e_i(\mathbf{p},t)| + \varepsilon_i. \tag{5.6}$$

Taking (3.3) into account, we get

$$|e_i(\mathbf{q}, \check{t}_i(\mathbf{q}))| < |e_i(\mathbf{q},t)|. \tag{5.7}$$

Let $e_i^+(\mathbf{P},t)$ be the upper bound of the interval $|E_i(\mathbf{P},t)|$. According to Lemma 4, the two following inequalities are satisfied

$$|e_i(\mathbf{p}, \check{t}_i(\mathbf{q}))| < |e_i(\mathbf{q}, \check{t}_i(\mathbf{q}))| + \varepsilon_i, \tag{5.8}$$

$$e_i^+(\mathbf{P}, \check{t}_i(\mathbf{q})) < |e_i(\mathbf{p}, \check{t}_i(\mathbf{q}))| + \varepsilon_i. \tag{5.9}$$

Since $\mathbf{Q} \subset \mathbf{P}$ and $|E_i(\mathbf{P}, t|$ is inclusion monotonic,

$$e_i^+(\mathbf{Q}, \check{t}_i(\mathbf{q})) < e_i^+(\mathbf{P}, \check{t}_i(\mathbf{q})). \tag{5.10}$$

Adding inequalities (5.6)-(5.10) under (5.5), we get $e_i^+(\mathbf{Q}, \check{t}_i(\mathbf{q})) < |e_i(\mathbf{p}, t)| + 3\varepsilon_i = e_i^{max}$. Therefore $|E_i(\mathbf{Q}, \check{t}_i(\mathbf{q}))| = |Y_m(\mathbf{Q}, \check{t}_i(\mathbf{q})) - \text{center}(Y_i)| < e_i^{max}$,which implies that $Y_m(\mathbf{Q}, \check{t}_i(\mathbf{q})) \subset Y_i$. ♢

Proof of Theorem 2 From Lemma 5, if $\mathbf{p} \in \overset{o}{\mathcal{S}}$, for all i in $\{1, ..., m\}$, there exists $\eta_i > 0$ such that, for any box $\mathbf{Q}$ containing $\mathbf{p}$ with a width smaller than η_i, $Y_m(\mathbf{Q}, \check{t}_i(\text{center}(\mathbf{Q}))) \subset Y_i$. Therefore any box $\mathbf{P}$ containing $\mathbf{p}$ with a width smaller than $\varepsilon = \min(\eta_1, \eta_2, ..., \eta_m)$ is such that $Y_m(\mathbf{P}, \check{t}_i(\text{center}(\mathbf{P}))) \subset Y_i$. For such a $\mathbf{P}$, INSIDE concludes that $\mathbf{P} \subset \mathcal{S}$. ♢

5.2. Exclusion test theorem

Theorem 6. *(Exclusion test) If* $\mathbf{p} \notin \mathcal{S}$, $\exists \varepsilon > 0$ *such that for any box* $\mathbf{P}$ *containing* $\mathbf{p}$ *and with a width smaller than* ε, *OUTSIDE(*$\mathbf{P}$*) concludes that* $\mathbf{P} \cap \mathcal{S} = \emptyset$. ♢

Proof The proof is again by contradiction. Let $\{\varepsilon(k), k \in \mathbb{N}\}$ be a sequence of positive reals converging to zero. If Theorem 6 is false, then

$$\forall k \in \mathbb{N}, \exists \mathbf{P}(k) \ni \mathbf{p}, w(\mathbf{P}(k)) < \varepsilon(k) \text{ such that OUTSIDE fails.} \tag{5.11}$$

This means that, for each i, OUTSIDE jumps to the next i at Step 4. It has thus found an interval $T(k)$ such that $w(T(k)) < w(\mathbf{P}(k))$ and such that the condition at

Step 3 is not satisfied, *i.e.*,

$$\forall i, \exists T(k) \subset T_i, w(T(k)) < w(\mathbf{P}(k)) \mid Y_m(\mathbf{P}(k), T(k)) \cap Y_i \neq \emptyset. \tag{5.12}$$

Since the sequence $\{T(k), k \in \mathbb{N}\}$ is enclosed in the compact set T_i and has its width converging to zero, it has at least one accumulation point $\bar{t}$, and there exists a subsequence $\{T(k_j),\ j \in \mathbb{N}, \text{ and } k_j < k_{j+1}\}$ containing $\bar{t}$. The sequence of intervals $Y_m(\mathbf{P}(k_j), T(k_j))$ thus converges to the real $y_m(\mathbf{p}, \bar{t})$. Now, according to (5.12), $Y_m(\mathbf{P}(k_j), T(k_j)) \cap Y_i \neq \emptyset$, which implies $y_m(\mathbf{p}, \bar{t}) \in Y_i$. Therefore $\mathbf{p} \in \mathcal{S}$. $\diamondsuit$

6. Test case

Consider again the test case introduced at the end of Section 2. For $\varepsilon = 0.01$ and $\mathbf{P}_0 = [0, 1.2] \times [0, 0.5]$, SIVIA generates the subpavings represented on Figure 3. The dark gray boxes have been proved to be included in $\mathcal{S}$ and correspond to $\mathcal{S}^-$. The light gray boxes have been proved to have an empty intersection with $\mathcal{S}$. No conclusion has been reached for the black boxes, which belong to the uncertainty layer $\Delta\mathcal{S}$.

Possible location for Figure 3.

Taking advantage of Remark 2, this paving is obtained in 38 seconds, instead of about 8 minutes, on a Pentium 133-based personnal computer.

7. Conclusions

In the context of bounded-error estimation, the actual values of experimental factors may also be uncertain. To take this into account, we have assumed that the output data and these experimental factors are only known to belong to some given intervals. The problem considered was the guaranteed characterization of the set of all parameter vectors of a given model structure that are consistent with this hypothesis. A new algorithm, based on interval analysis, has been proposed for this purpose. For the sake of simplicity, it has been described in the particular case where there is a single experimental factor. Even in this simple case, we know of no other method that allows a guaranteed characterization of the posterior feasible set for the parameters in this nonlinear context. The fundamental notion of inclusion function allows the approach to deal with a huge class of problems. Convergence analysis shows that in almost any situation, the set of feasible parameter vector can be characterized with an accuracy that is only limited by the effect of rounding. The main limitation of the approach is that its complexity is exponential with the dimension of parameter space. The approach readily generalizes to problems involving q experimental factors and r outputs. The posterior feasible set is then defined by

$$\mathcal{S} = \{\mathbf{p} \in R^{n_\mathrm{p}} \mid \forall i \in \{1, ..., m\}, \exists \mathbf{t} \in \mathbf{T}_i \subset \mathbb{R}^q \mid \mathbf{y}_m(\mathbf{p}, \mathbf{t}) \in \mathbf{Y}_i \subset \mathbb{R}^r\}\,, \tag{7.1}$$

and vector inclusion functions and a global optimization algorithm for multivariable functions must be employed. The algorithm used in this paper has voluntarily been kept simple, so as to facilitate convergence analysis, and there is ample room for further improvements of its efficiency.

Acknowledgement

The authors wish to thank an anonymous reviewer for comments which have helped them to improve the presentation of the paper.

References

Cerone, V. (1991). Parameter bounds for models with bounded errors in all variables, in *Prep. 9th IFAC/IFORS Symp. Identification and System Parameter Estimation*, Budapest, pp. 1518–1523.

Cerone, V. (1996). Errors-in-variables models in parameter bounding, *in* M. Milanese, J. Norton, H. Piet-Lahanier and E. Walter (Eds.), *Bounding Approaches to System Identification,* Plenum, New York, pp. 289–306.

Clément, T., and S. Gentil (1990). Recursive membership set estimation for output error models, *Mathematics and Computers in Simulation*, **32**, 505–513.

Combettes, P.L. (1993). The foundations of set theoretic estimation, *Proc. of the IEEE*, **81**(2), 182–208.

Deller, J.R., M. Nayeri and S.H. Odeh, (1993). Least-square identification with error bounds for real-time signal processing and control, *Proc. of the IEEE*, **81**(6), 815–849

Hammer, R., M. Hocks, U. Kulisch, and D. Ratz (1995). *C++ Toolbox For Verified Computing*, Springer-Verlag, Heidelberg.

Hansen, E. (1992). *Global Optimization Using Interval Analysis*, Marcel Dekker, New York.

Jaulin, L., and E. Walter (1993a). Set estimation via interval analysis for nonlinear bounded-error estimation, *Automatica*, **29**(4), 1053–1064.

Jaulin, L., and E. Walter (1993b). Guaranteed nonlinear parameter estimation from bounded-error data via interval analysis, *Mathematics and Computers in Simulation*, **35**, 123–137.

Milanese, M., and A. Vicino (1991), Estimation theory for nonlinear models and set membership uncertainty. *Automatica*, **27**, 403–408.

Milanese, M., J. Norton, H. Piet-Lahanier and E. Walter (Eds.) (1996). *Bounding Approaches to System Identification*, Plenum, New York.

Moore, R.E. (1979). *Methods and Applications of Interval Analysis*, SIAM, Philadelphia.

Norton, J.-P. (1987). Identification of parameter bounds for ARMAX models from records with bounded noise, *Int. J. Control*, **45**, 375–390.

J.-P. Norton (Ed.) (1994). Special issue on bounded-error estimation, **1**, *Int. J. Adapt. Control Signal Proc.*, **8**(1), 1–118.

J.-P. Norton (Ed.) (1995). Special issue on bounded-error estimation, **2**, *Int. J. Adapt. Control Signal Proc.*, **9**(1), 1–132.

Pronzato, L., and E. Walter (1994). Minimal volume ellipsoids, *Int. J. Adapt. Control Signal Process.*, **8**(1), 15–30.

Schwetlick, H., and V. Tiller (1985). Numerical methods for estimating parameters in nonlinear models with errors in the variables, *Technometrics*, **27**(1), 17–24.

Veres, S.M., and J.P. Norton (1996), Parameter-bounding algorithms for linear errors-in-variables models, in M. Milanese, J. Norton, H. Piet-Lahanier and E. Walter (Eds.), *Bounding Approaches to System Identification,* Plenum, New York, pp. 275–288.

Walter, E., (Ed.) (1990). Parameter identifications with error bounds, special issue of *Mathematics and Computers in Simulation*, **32**(5&6), 447-638.

Walter, E., and L. Pronzato (1997). *Identification of Parametric Models from Experimental Data*, Springer, London, pp. 257–283.

Figure captions

Figure 1: The boxes indicate the uncertainty associated with each output data point and associated measurement time. A parameter vector is feasible if and only if $y_m(\mathbf{p}, t)$ goes through all ten boxes.

Figure 2: Illustration of the strategy chosen for INSIDE. The interval $y_m(\mathbf{P}, \check{t}_i(\mathbf{p}))$ is included in the data interval Y_i. Therefore the parameter vector $\mathbf{p}$ is consistent with the ith datum.

Figure 3: Paving generated by SIVIA to bracket the solution set $\mathcal{S}$ for the test case. The frame box is $\mathbf{P}_0 = [0, 1.2] \times [0, 0.5]$. Dark and light gray boxes are inside and outside $\mathcal{S}$, respectively. $\Delta\mathcal{S}$ is in black.

Table caption

Table 1: Feasible intervals for collected data in the test case.

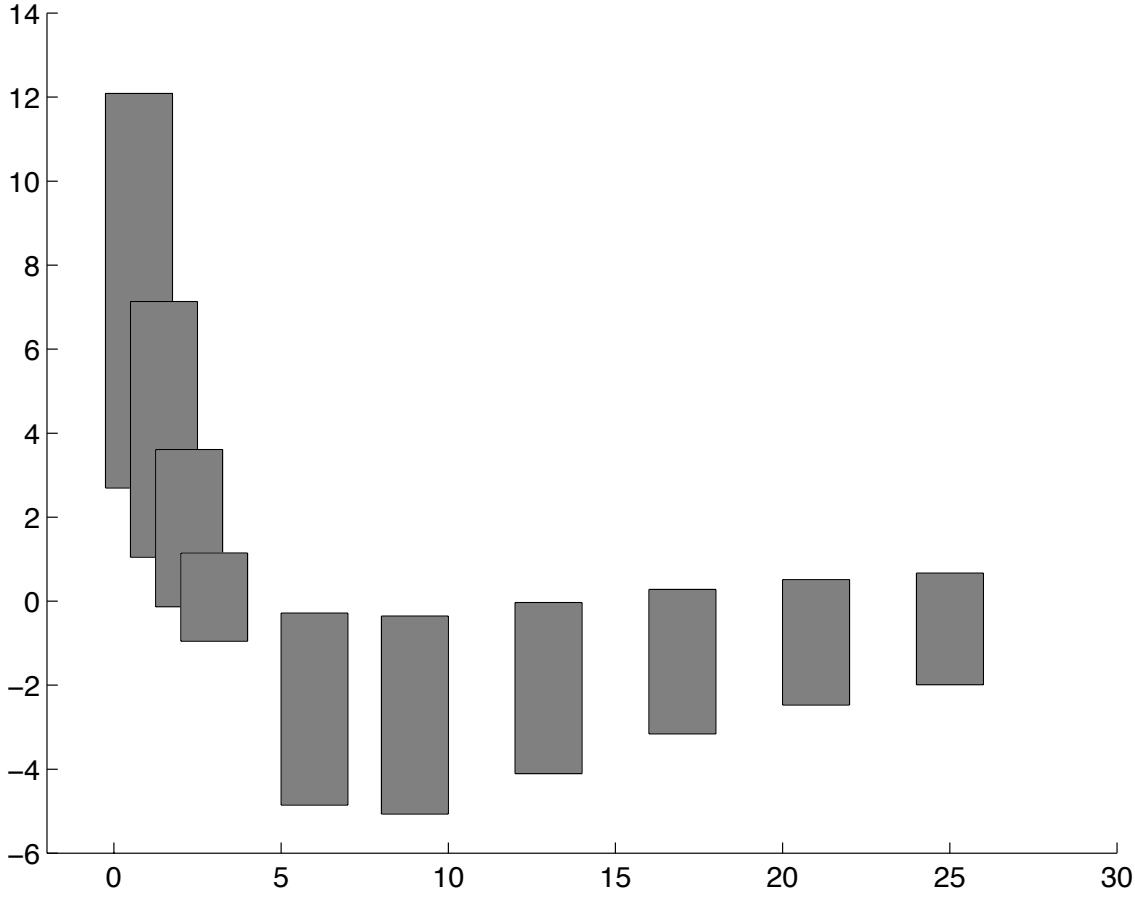

Figure 1

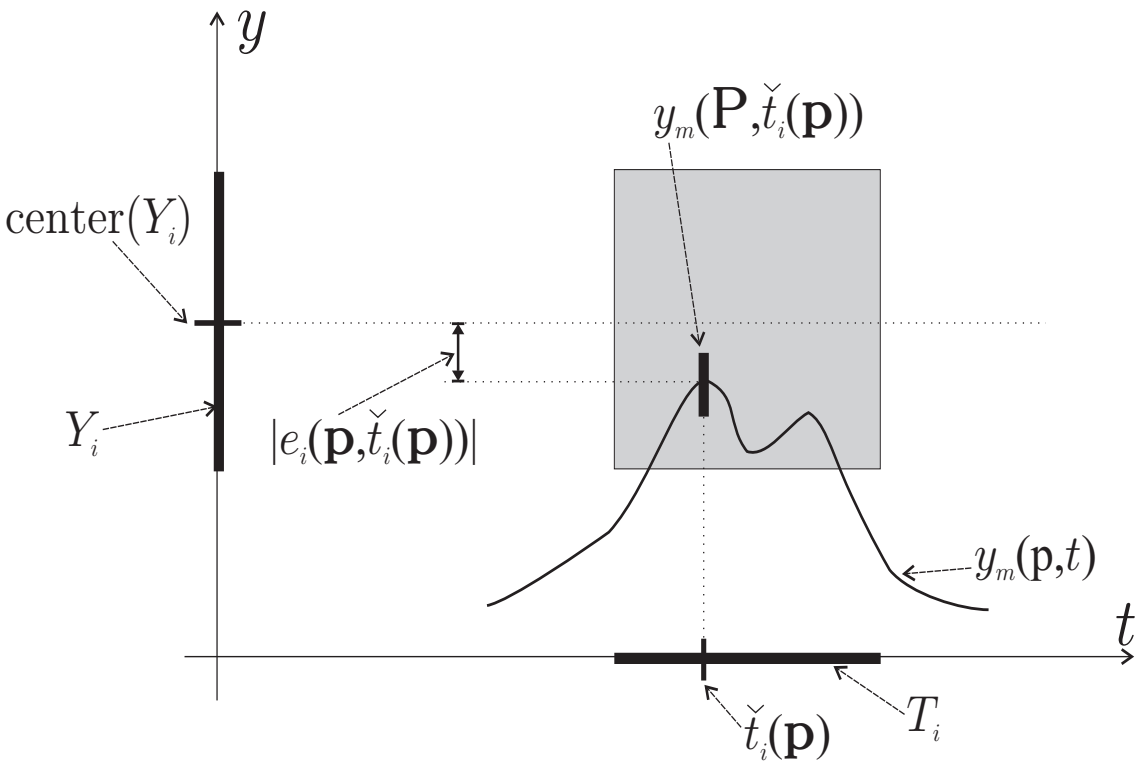
y
$y_m(\mathrm{P},\check{t}_i(\mathbf{p}))$
$\mathrm{center}(Y_i)$
Y_i
$|e_i(\mathbf{p},\check{t}_i(\mathbf{p}))|$
$y_m(\mathrm{p},t)$
t
$\check{t}_i(\mathbf{p})$
T_i

Figure 2

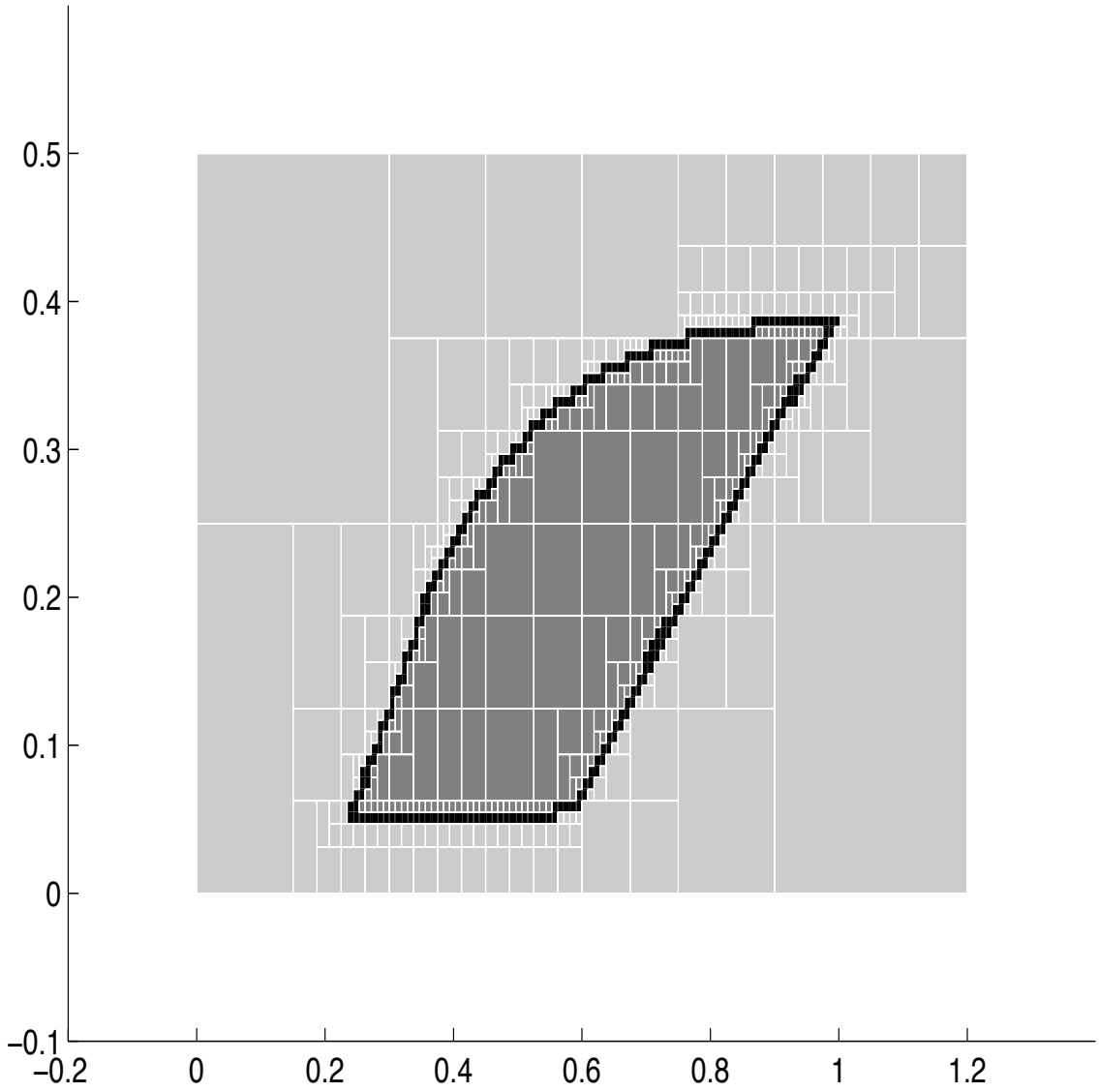

Figure 3

Table 1

i	T_i	Y_i
1	$[-0.25, 1.75]$	$[2.7, 12.1]$
2	$[0.5, 2.5]$	$[1.04, 7.14]$
3	$[1.25, 3.25]$	$[-0.13, 3.61]$
4	$[2, 4]$	$[-0.95, 1.15]$
5	$[5, 7]$	$[-4.85, -0.29]$
6	$[8, 10]$	$[-5.06, -0.36]$
7	$[12, 14]$	$[-4.1, -0.04]$
8	$[16, 18]$	$[-3.16, 0.3]$
9	$[20, 22]$	$[-2.5, 0.51]$
10	$[24, 26]$	$[-2, 0.67]$