12

Lille. CESA 1996 IMAC Multiconference Symposium on Modelling Analysis and Simulation 1156-

GUARANTEED ROBUST NONLINEAR PARAMETER BOUNDING

Luc Jaulin*, Eric Walter** and Olivier Didrit**

* Faculté des Sciences
Université d'Angers
2, Boulevard Lavoisier
49045 Angers Cedex 01
FRANCE
FAX (33) 41 73 53 52
Email: jaulin@univ-angers.fr

** Laboratoire des Signaux et Systèmes
CNRS - École Supérieure d'Électricité
Plateau de Moulon
91192 Gif-sur-Yvette Cedex
FRANCE
FAX (33) 1 69 41 30 60
Email: {walter, didrit}@lss.supelec.fr

ABSTRACT

Bounded-error estimation requires prior bounds on the acceptable errors between the data and corresponding model outputs. When these bounds can not be satisfied for all data points, one may look for the set $\mathbb{S}_q$ of all values of the parameter vector such that all errors but q fall within the acceptable ranges. This corresponds to considering that there may be up to q outliers in the data. New algorithms based on interval analysis make it possible to characterize $\mathbb{S}_q$ in a guaranteed way, either when q takes its minimum possible value or when it belongs to a given range of integers. The methodology readily applies to models nonlinear in their parameters and is illustrated by a two-dimensional example to allow pictures to be drawn.

1. INTRODUCTION

This paper deals with estimating the unknown parameter vector $\mathbf{p}$ of a model from a vector $\mathbf{y}$ of experimental data collected on a system. The vector $\mathbf{p}$ is assumed to belong to some (possibly very large) prior box $\mathbb{p}_0$ in the parameter space. The model $M(\mathbf{p})$ is a set of equations parametrized by $\mathbf{p}$ that generates a vector $\mathbf{y}_m(\mathbf{p})$ of outputs to be compared with the data. Define the error between the data and model output by

$$\mathbf{e}_m(\mathbf{p}) = \mathbf{y} - \mathbf{y}_m(\mathbf{p}). \quad (1)$$

In the context of bounded-error estimation, it is assumed that each component of $\mathbf{e}_m(\mathbf{p})$ must belong to a known interval to be admissible. This corresponds to assuming that $\mathbf{e}_m(\mathbf{p})$ belongs to some prior box $\mathbb{e} \subset \mathbb{R}^{\dim \mathbf{y}}$. For $\mathbf{p}$ to yield an admissible error, $\mathbf{y}_m(\mathbf{p})$ must therefore belong to the set of admissible outputs defined by

$$\mathbb{y} = \mathbf{y} - \mathbb{e} = \{\mathbf{y} - \mathbf{e} \mid \mathbf{e} \in \mathbb{e}\}. \quad (2)$$

The problem to be solved is that of finding the set $\mathbb{S}$ of all values of $\mathbf{p}$ in $\mathbb{p}_0$ corresponding to an admissible output, i.e.

$$\mathbb{S} = \{\mathbf{p} \in \mathbb{p}_0 \mid \mathbf{y}_m(\mathbf{p}) \in \mathbb{y}\} = \mathbf{y}_m^{-1}(\mathbb{y}) \cap \mathbb{p}_0. \quad (3)$$

It can thus be seen as one of *set inversion* (Jaulin and Walter, 1993a). A number of methods are available to characterize $\mathbb{S}$, see, e.g., (Walter, 1990; Norton, 1994, 1995; Milanese *et al.*, 1996) for special issues of journals and a book devoted to this topic. When $\mathbf{y}_m$ is linear in $\mathbf{p}$, $\mathbb{S}$ is a polytope that can be described exactly, or enclosed in a simpler-shaped set such as an ellipsoid, a box or a parallelotope. When $\mathbf{y}_m$ is nonlinear in $\mathbf{p}$, methods based on interval analysis are also available to enclose $\mathbb{S}$ in a union of boxes with an arbitrary precision. All these methods rely on the hypothesis that the prior bounds for the error are correct, which is not always realistic. This paper is concerned with the case where some data points may be *outliers*. Such outliers may for instance result from sensor failure (or any other error occuring during the data collection), from an optimistic choice of the error bounds or from the fact that the model structure is unable to describe the process behavior accurately enough. The associated error should then be allowed to escape the feasible range defined by the prior bounds. Otherwise, the set $\mathbb{S}$ might become unrealistically small or even empty.

The robust set estimator OMNE (for Outlier Minimal Number Estimator) has been designed to cater for this situation (Lahanier, Walter and Gomeni, 1987; Walter and Piet-Lahanier, 1988; Piet-Lahanier and Walter, 1990; Pronzato and Walter, 1996). An estimate $\mathbb{S}'$ is the set of all values of $\mathbf{p}$ in $\mathbb{p}_0$ that are associated with the globally minimum value of a criterion $j(\mathbf{p})$. This criterion is the number of components $y_{m,i}(\mathbf{p})$ of $\mathbf{y}_m(\mathbf{p})$ that do not fall within the feasible range $\mathbb{y}_i$ defined by their prior bounds i.e. such that $y_{m,i}(\mathbf{p}) \notin \mathbb{y}_i$. It is piecewise constant, and its gradient is therefore zero wherever it is differentiable. Moreover, $\mathbb{S}'$ is usually not a singleton, may be nonconvex or even disconnected, and has a nonzero volume. These features are illustrated by Figure 1 in a one-dimensional case. The algorithm used so far to characterize $\mathbb{S}'$ was based on a random scanning of the parameter space and no guarantee could be provided about the results obtained. One of the purposes of this paper is to describe a method to characterize $\mathbb{S}'$ in a guaranteed way.

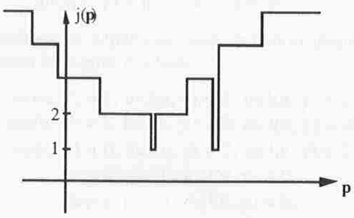

Figure 1: The criterion to be minimized is not continuous, and the set of all global minimizers is disconnected and not countable.

Section 2 briefly introduces interval analysis and defines the new notions of *inclusion* and *separation degrees*. These notions are used in Section 3 to present a new algorithm to build a set of boxes in the parameter space the union of which is guaranteed to contain $\mathbb{S}'$. This involves the resolution of a finite sequence of set-inversion problems by the algorithm SIVIA (Set Inversion Via Interval Analysis) developed by Jaulin and Walter (1993a). The procedure is illustrated on a test-case from the literature. In Section 4, a new algorithm for simultaneously characterizing the isocriteria of j(.) for several values of the number of data points considered as outliers is presented. The algorithms described in Sections 3 and 4 both lead to global and guaranteed results.

2. INTERVAL ANALYSIS

The following notions will be used for the description of the algorithms. A *box*, or *vector interval*, $\mathbf{x}$ of $\mathbb{R}^n$ is a vector whose components $\mathbf{x}_i$ (i = 1, ..., n) are scalar intervals:

$$\mathbf{x} = [x_1^-, x_1^+] \times ... \times [x_n^-, x_n^+] = \mathbf{x}_1 \times ... \times \mathbf{x}_n = [\mathbf{x}^-, \mathbf{x}^+] \,, \quad (4)$$

where $\mathbf{x}^- = (x_1^-, x_2^-, ..., x_n^-)^T$ and $\mathbf{x}^+ = (x_1^+, x_2^+, ..., x_n^+)^T$. The set of all boxes of $\mathbb{R}^n$ is denoted by $\mathbb{IR}^n$. A *principal plane* of $\mathbf{x}$ is a symmetry plane normal to a side of maximum length. Let $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^p$ be a vector function, the set-valued function $\mathbb{f} : \mathbb{IR}^n \to \mathbb{IR}^p$ is an *inclusion function* of $\mathbf{f}$ if

$$\mathbf{f}(\mathbf{x}) \subset \mathbb{f}(\mathbf{x}) \quad (5)$$

for any $\mathbf{x}$ of $\mathbb{IR}^n$. This inclusion function is *convergent* if, for any sequence of boxes $\mathbf{x}$ of $\mathbb{IR}^n$,

$$w(\mathbf{x}) \to 0 \quad \Rightarrow \quad w(\mathbb{f}(\mathbf{x})) \to 0, \quad (6)$$

where $w(\mathbf{x})$ is the *width* of $\mathbf{x}$, i.e. the length of its largest side(s). A number of methods exist for deriving an inclusion function associated with any function computable in a finite number of steps (see, e.g., (Moore, 1979; Ratschek and Rokne, 1988)). Consider, for instance, the function $\mathbf{f}$ from $\mathbb{R}^2$ to $\mathbb{R}^{10}$ given by the following algorithm:

Input: p_1, p_2
For i := 1 to 10 do $f_i := 20 \exp(-p_1 i) - 8 \exp(-p_2 i)$
Output: f_i, i ∈ {1, ..., 10}.

A convergent inclusion function $\mathbb{f}$ for $\mathbf{f}$ is given by the following interval algorithm:

Input: $\mathbf{p}_1$, $\mathbf{p}_2$
For i := 1 to 10 do $\mathbb{f}_i := 20\, \mathbf{exp}(-\mathbf{p}_1 i) - 8\, \mathbf{exp}(-\mathbf{p}_2 i)$
Output: $\mathbb{f}_i$, i ∈ {1, ..., 10},

where $\mathbf{exp}(.)$ denotes an inclusion function for the exponential function. The interval operations involved are as follows. If $\mathbf{x} = [x^-, x^+]$ and $\mathbf{y} = [y^-, y^+]$ are two scalar intervals and α is a real,

$$\mathbf{exp}(\mathbf{x}) = \exp(\mathbf{x}) = [\exp(x^-), \exp(x^+)], \quad (7)$$

$$\alpha.\mathbf{x} = [\alpha.x^-, \alpha.x^+] \text{ if } \alpha \geq 0, \quad (8)$$

$$\alpha.\mathbf{x} = [\alpha.x^+, \alpha.x^-] \text{ if } \alpha \leq 0 \text{ and} \quad (9)$$

$$\mathbf{x} + \mathbf{y} = [x^- + y^-, x^+ + y^+]. \quad (10)$$

A *subpaving* of $\mathbb{R}^n$ is a set of non-overlapping boxes of $\mathbb{IR}^n$, with nonzero width. If $\mathbb{A}$ is the subset of $\mathbb{R}^n$ generated by the union of all boxes of the subpaving $\mathbb{K}$, then $\mathbb{K}$ is a *paving* of $\mathbb{A}$. The new notions of separation and inclusion degrees will be helpful for the algorithms presented in Sections 3 and 4. The *separation degree between two scalar intervals* $\mathbf{x}$ and $\mathbf{y}$ of $\mathbb{IR}$ is defined by

$$\text{sep}(\mathbf{x}, \mathbf{y}) = 1 \text{ if } \mathbf{x} \cap \mathbf{y} = \emptyset$$
$$\text{and sep}(\mathbf{x}, \mathbf{y}) = 0 \text{ if } \mathbf{x} \cap \mathbf{y} \neq \emptyset. \quad (11)$$

The *separation degree between two boxes* $\mathbf{x}$ and $\mathbf{y}$ of $\mathbb{IR}^n$ is defined by

$$\text{sep}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \text{sep}(\mathbf{x}_i, \mathbf{y}_i). \quad (12)$$

Lemma 1: If $\mathbf{x}$ and $\mathbf{y}$ are two boxes of $\mathbb{IR}^n$, then $\text{sep}(\mathbf{x}, \mathbf{y}) > 0 \Leftrightarrow \mathbf{x} \cap \mathbf{y} = \emptyset$.

Proof: $\text{sep}(\mathbf{x}, \mathbf{y}) > 0 \Leftrightarrow \exists\, i \mid \text{sep}(\mathbf{x}_i, \mathbf{y}_i) = 1 \Leftrightarrow \exists\, i \mid \mathbf{x}_i \cap \mathbf{y}_i = \emptyset \Leftrightarrow \mathbf{x} \cap \mathbf{y} = \emptyset$. ◊

The *inclusion degree of a scalar interval* $\mathbf{x}$ *into a scalar interval* $\mathbf{y}$ is defined by

$$\text{incl}(\mathbf{x}, \mathbf{y}) = 1 \text{ if } \mathbf{x} \subset \mathbf{y}$$
$$\text{and incl}(\mathbf{x}, \mathbf{y}) = 0 \text{ otherwise.} \quad (13)$$

If $\mathbf{x}$ and $\mathbf{y}$ are two boxes of $\mathbb{IR}^n$, then *the inclusion degree of* $\mathbf{x}$ *into* $\mathbf{y}$ is defined by

$$\text{incl}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \text{incl}(\mathbf{x}_i, \mathbf{y}_i). \quad (14)$$

Lemma 2: If $\mathbf{x}$ and $\mathbf{y}$ are two boxes of $\mathbb{IR}^n$, then $\text{incl}(\mathbf{x}, \mathbf{y}) = n \Leftrightarrow \mathbf{x} \subset \mathbf{y}$.

Proof: $\text{incl}(\mathbf{x}, \mathbf{y}) = n \Leftrightarrow \forall\, i \in \{1, ..., n\}, \text{incl}(\mathbf{x}_i, \mathbf{y}_i) = 1 \Leftrightarrow \forall\, i, \mathbf{x}_i \subset \mathbf{y}_i \Leftrightarrow \mathbf{x} \subset \mathbf{y}$. ◊

When the box $\mathbf{x}$ reduces to a vector $\mathbf{x}$, Lemma 2 becomes

$$\text{incl}(\mathrm{x}, \mathbb{y}) = n \quad \Leftrightarrow \quad \mathrm{x} \in \mathbb{y}\,. \tag{15}$$

The notions of separation and inclusion degrees are illustrated by Figure 2, where

$$\begin{gathered}
\text{sep}(\mathbb{a}, \mathbb{y}) = 1,\ \text{incl}(\mathbb{a}, \mathbb{y}) = 1,\ \text{incl}(\mathbf{a}, \mathbb{y}) = 1,\\
\text{sep}(\mathbb{b}, \mathbb{y}) = 2,\ \text{incl}(\mathbb{b}, \mathbb{y}) = 0,\ \text{incl}(\mathbf{b}, \mathbb{y}) = 0,\\
\text{sep}(\mathbb{c}, \mathbb{y}) = 0,\ \text{incl}(\mathbb{c}, \mathbb{y}) = 2,\ \text{incl}(\mathbf{c}, \mathbb{y}) = 2,\\
\text{sep}(\mathbb{d}, \mathbb{y}) = 0,\ \text{incl}(\mathbb{d}, \mathbb{y}) = 0,\\
\text{sep}(\mathbb{e}, \mathbb{y}) = 1,\ \text{incl}(\mathbb{e}, \mathbb{y}) = 0.
\end{gathered}$$

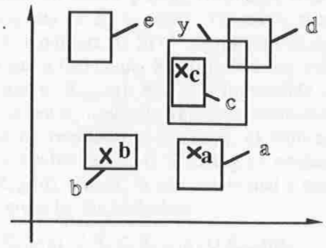

Figure 2: Illustration of the inclusion and separation degrees.

Lemma 3: If $\mathbb{x}$ and $\mathbb{y}$ are scalar intervals and x is a real in $\mathbb{x}$ then $\text{incl}(x, \mathbb{y}) \geq \text{incl}(\mathbb{x}, \mathbb{y})$.

Proof: The value of $\text{incl}(\mathbb{x}, \mathbb{y})$ is either 0 or 1. If $\text{incl}(\mathbb{x}, \mathbb{y}) = 1$ then, $\mathbb{x} \subset \mathbb{y}$ i.e. $\forall\, x \in \mathbb{x}$, $\text{incl}(x, \mathbb{y}) = 1$. If $\text{incl}(\mathbb{x}, \mathbb{y}) = 0$, then $\text{incl}(x, \mathbb{y}) = 0$ or 1. ◊

Lemma 4: If $\mathbb{x}$ and $\mathbb{y}$ are two boxes of $\mathbb{IR}^n$ and x is a vector in $\mathbb{x}$, then $\text{incl}(\mathrm{x}, \mathbb{y}) \geq \text{incl}(\mathbb{x}, \mathbb{y})$.

Proof: From Lemma 3, $\forall\, i \in \{1, ..., n\}$, $\text{incl}(x_i, \mathbb{y}_i) \geq \text{incl}(\mathbb{x}_i, \mathbb{y}_i)$. Therefore

$$\text{incl}(\mathrm{x}, \mathbb{y}) = \sum_{i=1}^{n} \text{incl}(x_i, \mathbb{y}_i) \geq \sum_{i=1}^{n} \text{incl}(\mathbb{x}_i, \mathbb{y}_i) = \text{incl}(\mathbb{x}, \mathbb{y}).\ ◊$$

Lemma 5: If $\mathbb{x}$ and $\mathbb{y}$ are two boxes of $\mathbb{IR}^n$, then $\text{sep}(\mathbb{x}, \mathbb{y}) + \text{incl}(\mathbb{x}, \mathbb{y}) \leq n$.

Proof: Since $\mathbb{x}_i \subset \mathbb{y}_i$ and $\mathbb{x}_i \cap \mathbb{y}_i = \varnothing$ cannot be satisfied together, $\text{sep}(\mathbb{x}_i, \mathbb{y}_i) + \text{incl}(\mathbb{x}_i, \mathbb{y}_i) \leq 1$, $\forall\, i \in \{1, ..., n\}$. Therefore

$$\text{sep}(\mathbb{x}, \mathbb{y}) + \text{incl}(\mathbb{x}, \mathbb{y}) = \sum_{i=1}^{n} (\text{sep}(\mathbb{x}_i, \mathbb{y}_i) + \text{incl}(\mathbb{x}_i, \mathbb{y}_i)) \leq n.\ ◊$$

Lemma 6: If $\mathbb{x}$ and $\mathbb{y}$ are two boxes of $\mathbb{IR}^n$ and x is any vector in $\mathbb{x}$, then $\text{sep}(\mathbb{x}, \mathbb{y}) + \text{incl}(\mathrm{x}, \mathbb{y}) \leq n$.

Proof: Since $x_i \in \mathbb{y}_i$ and $\mathbb{x}_i \cap \mathbb{y}_i = \varnothing$ cannot be satisfied together, $\text{sep}(\mathbb{x}_i, \mathbb{y}_i) + \text{incl}(x_i, \mathbb{y}_i) \leq 1$, $\forall\, i \in \{1, ..., n\}$. Therefore

$$\text{sep}(\mathbb{x}, \mathbb{y}) + \text{incl}(\mathrm{x}, \mathbb{y}) = \sum_{i=1}^{n} (\text{sep}(\mathbb{x}_i, \mathbb{y}_i) + \text{incl}(x_i, \mathbb{y}_i)) \leq n.\ ◊$$

3. GUARANTEED OMNE

The parameter vector $\mathbf{p}$ is consistent with the ith datum if $y_{m,i}(\mathbf{p}) \in \mathbb{y}_i$, so that $\text{incl}(y_{m,i}(\mathbf{p}), \mathbb{y}_i) = 1$. The set $\mathbb{Y}_q$ of model output vectors that are consistent with at least $\dim \mathbb{y} - q$ data points, i.e. such that at most q data points are considered as outliers, is defined by

$$\mathbb{Y}_q = \{\mathbf{y}_m \mid \text{incl}(\mathbf{y}_m, \mathbb{y}) \geq \dim \mathbb{y} - q\}. \tag{16}$$

Note that $\mathbb{Y}_0 = \mathbb{y}$. The set of all feasible parameter vectors that are consistent with at least $\dim \mathbb{y} - q$ data points can then be defined as follows:

$$\mathbb{S}_q = \{\mathbf{p} \in \mathbb{p}_0 \mid \mathbf{y}_m(\mathbf{p}) \in \mathbb{Y}_q\} = \mathbf{y}_m^{-1}(\mathbb{Y}_q) \cap \mathbb{p}_0. \tag{17}$$

The problem of characterizing $\mathbb{S}_q$ has thus been cast into the framework of *set inversion* and can easily be solved with SIVIA, taking into account the fact that outliers are now allowed via the definition of $\mathbb{Y}_q$ and the notions of separation and inclusion degrees. Let $\mathbb{y}_m(.)$ be an inclusion function of $\mathbf{y}_m(.)$, so that for any box $\mathbb{p}$ of the parameter space $\mathbb{y}_m(\mathbb{p})$ is a box guaranteed to contain all values of $\mathbf{y}_m(\mathbf{p})$ for all $\mathbf{p}$ in $\mathbb{p}$. Then the two following propositions hold.

Proposition 1: $\forall\, \mathbb{p} \subset \mathbb{p}_0$, $\text{incl}(\mathbb{y}_m(\mathbb{p}), \mathbb{y}) \geq \dim \mathbb{y} - q \Rightarrow \mathbb{p} \subset \mathbb{S}_q$.

Proof: If $\text{incl}(\mathbb{y}_m(\mathbb{p}), \mathbb{y}) \geq \dim \mathbb{y} - q$ then, from Lemma 4, $\forall\, \mathbf{p} \in \mathbb{p}$, $\text{incl}(\mathbf{y}_m(\mathbf{p}), \mathbb{y}) \geq \dim \mathbb{y} - q$, so $\forall\, \mathbf{p} \in \mathbb{p}$, $\mathbf{y}_m(\mathbf{p}) \in \mathbb{Y}_q \Leftrightarrow \mathbb{p} \subset \mathbf{y}_m^{-1}(\mathbb{Y}_q)$. Since $\mathbb{p} \subset \mathbb{p}_0$, $\mathbb{p} \subset \mathbb{S}_q$. ◊

Proposition 2: $\text{sep}(\mathbb{y}_m(\mathbb{p}), \mathbb{y}) > q \Rightarrow \mathbb{p} \cap \mathbb{S}_q = \varnothing$.

Proof: Assume that $\text{sep}(\mathbb{y}_m(\mathbb{p}), \mathbb{y}) > q$. From Lemma 6, $\forall\, \mathbf{p} \in \mathbb{p}$, $\text{sep}(\mathbb{y}_m(\mathbb{p}), \mathbb{y}) + \text{incl}(\mathbf{y}_m(\mathbf{p}), \mathbb{y}) \leq \dim \mathbb{y}$. Therefore $\forall\, \mathbf{p} \in \mathbb{p}$, $\text{incl}(\mathbf{y}_m(\mathbf{p}), \mathbb{y}) \leq \dim \mathbb{y} - \text{sep}(\mathbb{y}_m(\mathbb{p}), \mathbb{y}) < \dim \mathbb{y} - q$ so $\mathbf{y}_m(\mathbf{p}) \notin \mathbb{Y}_q$. Therefore $\mathbb{p} \cap \mathbf{y}_m^{-1}(\mathbb{Y}_q) = \varnothing$, which implies that $\mathbb{p} \cap \mathbb{S}_q = \varnothing$.

The version of SIVIA to be presented in what follows is parametrized by q, the number of data points considered as outliers, and ε, the required accuracy. For a given value of q, if a box $\mathbb{p} \subset \mathbb{p}_0$ satifies none of the two conditions $\text{incl}(\mathbb{y}_m(\mathbb{p}), \mathbb{y}) \geq \dim \mathbb{y} - q$ and $\text{sep}(\mathbb{y}_m(\mathbb{p}), \mathbb{y}) > q$, $\mathbb{p}$ is said to be *indeterminate*. SIVIA(q, ε) generates two subpavings, namely $\mathbb{K}_{OK}(q)$ that contains all boxes that have been proved to be included in $\mathbb{S}_q$ and $\mathbb{K}_{ind}(q)$ that contains all indeterminate boxes with size smaller than ε.

Algorithm SIVIA(q, ε)

Inputs: q, number of allowed outliers,
ε, required accuracy.

Initialization: stack := ∅, $\mathbb{K}_{OK}(q) := \varnothing$,
$\mathbb{K}_{ind}(q) := \varnothing$, $\mathbb{p} := \mathbb{p}_0$.

Iteration:

Step 1 If $\text{incl}(\mathbb{y}_m(\mathbb{p}), \mathbb{y}) \geq \dim \mathbb{y} - q$,
then $\mathbb{K}_{OK}(q) := \mathbb{K}_{OK}(q) \cup \mathbb{p}$. Go to Step 4.

Step 2 If $sep(y_m(p), y) > q$, then go to Step 4.
Step 3 If $w(p) \le \varepsilon$, then $K_{ind}(q) := K_{ind}(q) \cup p$,
else {bisect p along a principal plane.
Stack the two resulting boxes}.
Step 4 If the stack is not empty,
then unstack into p and go to Step 1.
End.
Outputs: $K_{OK}(q)$, $K_{ind}(q)$. ◊

If SIVIA(q, ε) returns empty $K_{OK}(q)$ and $K_{ind}(q)$, then there exists no vector **p** in p_0 consistent with at least dim y − q data i.e. S_q is empty. Therefore, there are at least q + 1 outliers. If SIVIA(q, ε) returns an empty $K_{OK}(q)$ and a non-empty $K_{ind}(q)$, then any vector **p** in S_q belongs to $K_{ind}(q)$, but it is impossible to know whether or not S_q is empty. This indetermination can be removed by reexecuting SIVIA(q, ε) with a smaller accuracy coefficient ε. If SIVIA(q, ε) returns a non-empty $K_{OK}(q)$, then S_q is non-empty and a bracketting of S_q is given by the inclusions

$$K_{OK}(q) \subset S_q \subset K_{OK}(q) \cup K_{ind}(q). \quad (18)$$

The main algorithm GOMNE (Guaranteed Outlier Minimal Number Estimator), to be presented now, uses SIVIA(q, ε) as a subroutine and aims at characterizing the set S'. GOMNE first calls SIVIA(0, ε), with an adaptative ε. If $K_{OK}(0) \cup K_{ind}(0)$ turns out to be empty, there is at least one outlier and GOMNE calls SIVIA(1, ε). The procedure is iterated, increasing q up to q^* such that $K_{OK}(q^*)$ is not empty. A guaranteed characterization for the solution set S' is then given by the inclusions $K_{OK}(q^*) \subset S' \subset K_{OK}(q^*) \cup K_{ind}(q^*)$.

Algorithm GOMNE

Input: ε_0, required accuracy for the characterization of S'.

Initialization: q := 0.

Iteration:
Step 1 $\varepsilon := \varepsilon_0$.
Step 2 Call SIVIA(q, ε)
Step 3 If $K_{OK}(q)$ and $K_{ind}(q)$ are empty,
then set q := q + 1 and go to Step 1.
Step 4 If $K_{OK}(q)$ is not empty,
then $K_{OK}(q) \subset S_q \subset K_{OK}(q) \cup K_{ind}(q)$. End.
Step 5 If $K_{OK}(q)$ is empty and $K_{ind}(q)$ is non-empty,
then set ε := 0.5 ε. Go to Step 2.
End. ◊

If the required accuracy ε_0 is small enough, the condition in Step 5 of GOMNE almost never occur except in atypical situations studied in Jaulin and Walter (1993a).

Test case: To illustrate the behavior of GOMNE, a two-parameter estimation problem is now considered, which makes it possible to draw pictures of the paving obtained. This example has been taken from Jaulin and Walter (1993b) and is a two parameter version of a problem treated by Milanese and Vicino (1991). In these papers, the vector of all available data was given by

$$\mathbf{y} = (7.39, 4.09, 1.74, 0.097, -2.57, -2.71, -2.07, -1.44, -0.98, -0.66)^T, \quad (19)$$

but here, to simulate outliers, we arbitrarily replace two data points by zero and take

$$\mathbf{y} = (7.39, 0, 1.74, 0.097, -2.57, -2.71, -2.07, 0, -0.98, -0.66)^T. \quad (20)$$

These data correspond to ten scalar measurements, taken at times

$$\mathbf{t} = (0.75, 1.5, 2.25, 3, 6, 9, 13, 17, 21, 25)^T. \quad (21)$$

The ith component of $y_m(p)$ is given by

$$y_{m,i}(\mathbf{p}) = 20\exp(-p_1 t_i) - 8\exp(-p_2 t_i). \quad (22)$$

As in (Jaulin and Walter, 1993b), the set of all feasible model outputs is the box defined by

$$\mathbf{y} = [\mathbf{y} - \mathbf{e}_{max}, \mathbf{y} + \mathbf{e}_{max}], \quad (23)$$

with

$$\mathbf{e}_{max} = (4.695, 1., 1.87, 1.0485, 2.285, 2.355, 2.035, 1., 1.49, 1.33)^T. \quad (24)$$

Figure 3 presents the data. The bars indicate the uncertainty associated with each datum. S_q is the set of all the values of **p** such that the model output goes through (10 − q) bars.

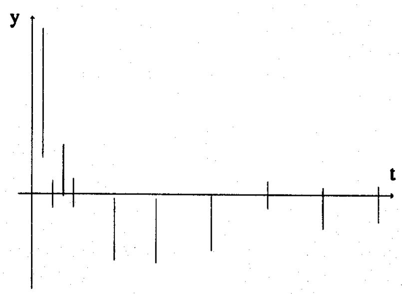

Figure 3: Contaminated experimental data with error bars.

For a required accuracy $\varepsilon_0 = 0.005$, and $p_0 = [-0.1, 1.5] \times [-0.1, 1.5]$, GOMNE calls SIVIA(0, ε_0) which proves that S_0 is empty (Figure 4). It then calls SIVIA(1, ε_0), which succeeds in bracketting S_1 while generating the paving presented in Figure 5. The whole procedure takes less than 13 seconds on a 486 DX4-100 computer.

GOMNE has thus detected that there exists at least one outlier. To protect oneself against one additional undetected outlier, it suffices to run SIVIA with q = 2. This generates the paving presented on Figure 6. Note that $K_{OK}(2) \cup K_{ind}(2)$ is disconnected, which proves that S_2 is. This may be due to the fact that the data considered as outliers are not always the same.

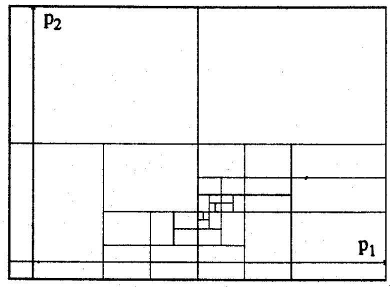

Figure 4: Paving generated from contaminated experimental data, assuming no outlier. All white boxes have been proved not to belong to $\mathbb{K}_{OK}(0)$ and $\mathbb{K}_{ind}(0)$, which turn out to be empty. The frame corresponds to the search domain $[-0.1, 1.5] \times [-0.1, 1.5]$.

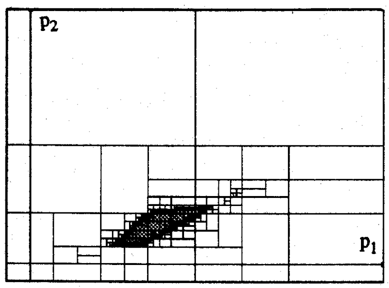

Figure 5: Paving generated from contaminated experimental data, assuming one outlier. $\mathbb{K}_{OK}(1)$ is in light grey and $\mathbb{K}_{ind}(1)$ in black. The frame corresponds to the search domain $[-0.1, 1.5] \times [-0.1, 1.5]$.

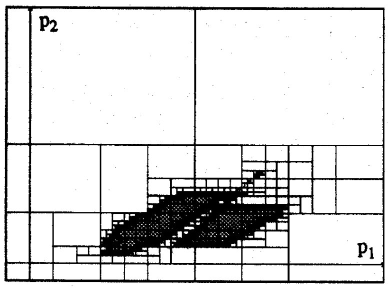

Figure 6: Paving generated from contaminated experimental data, assuming two outliers. $\mathbb{K}_{OK}(2)$ is in light grey and $\mathbb{K}_{ind}(2)$ in black. The frame corresponds to the search domain $[-0.1, 1.5] \times [-0.1, 1.5]$.

For comparison, the initial data vector with no outliers has also been processed by SIVIA. The paving obtained is presented in Figure 7. One can easily check that the union of $\mathbb{K}_{OK}(0)$ and $\mathbb{K}_{ind}(0)$ obtained with the regular data is included in the union of $\mathbb{K}_{OK}(2)$ and $\mathbb{K}_{ind}(2)$ obtained with the contaminated data.

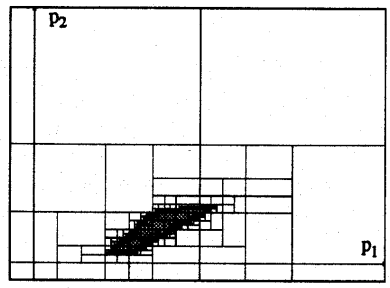

Figure 7: Paving generated from regular data, assuming no outlier. $\mathbb{K}_{OK}(0)$ is in light grey and $\mathbb{K}_{ind}(0)$ in black. The frame corresponds to the search domain $[-0.1, 1.5] \times [-0.1, 1.5]$.

4. CHARACTERIZATION OF ISOCRITERIA

It may be interesting to characterize $\mathbf{S}_q$ simultaneously for several values of q, for instance to detect whether the shape of $\mathbf{S}_q$ changes drastically with the number of data points rejected, which may help to detect outliers. This problem can be solved via the characterization of isocriteria, i.e. sets of values of **p** such that j(**p**) is constant. This is the purpose of the algorithm ISOON (Iso Outlier Number) to be presented now. It relies on the two following conditions, which result from Propositions 1 and 2:

$$\forall\, \mathbf{p} \subset \mathbf{p}_0,\ \mathrm{incl}(\mathbf{y}_m(\mathbf{p}), \mathbf{y}) = \dim \mathbf{y} - q \Rightarrow \mathbf{p} \subset \mathbf{S}_q\,, \quad (25)$$

$$\mathrm{sep}(\mathbf{y}_m(\mathbf{p}), \mathbf{y}) > q - 1 \Rightarrow \mathbf{p} \cap \mathbf{S}_{q-1} = \varnothing\,. \quad (26)$$

Any box **p** that simultaneously satisfies the conditions of (25) and (26) corresponds to values of the parameters that are consistent with exactly q data points. It is thus guaranteed to belong to the isocriterion at level q. Again, the principle of ISOON is to partition the prior box of interest into a set of boxes that either are too small to deserve further consideration or have been proved to belong to or to have an empty intersection with either of the isocriteria of interest. If **p** denotes the current box, the algorithm searches for a value of q among those of interest for which (25) and (26) are both satisfied. If such a value is not found, the current box **p** is bisected.

Let $\mathbb{K}_{iso}(q)$ be the subpaving that contains all the boxes that are proved to be included in $\mathbf{S}_q$ and that have an empty intersection with $\mathbf{S}_{q-1}$, and let $\mathbb{K}_{ind}$ be the subpaving that contains all indeterminate boxes too small to be bisected. In what follows, the required accuracy is denoted by ε and q_{max} is the maximum number of potential outliers to be considered.

Algorithm ISOON

Initialization: stack := ∅; $\mathbb{K}_{ind}$:= ∅; $\mathbb{p} := \mathbb{p}_0$;
for q := 0 to q_{max} do $\mathbb{K}_{iso}(q) := \varnothing$.

Iteration:

Step 1 For q := 0 to q_{max} do {
If $\text{incl}(\mathbf{y}_m(\mathbb{p}), \mathbf{y}) = \dim \mathbf{y} - q$,
then {$\mathbb{K}_{iso}(q) := \mathbb{K}_{iso}(q) \cup \mathbb{p}$, go to Step 3}.
If $\text{sep}(\mathbf{y}_m(\mathbb{p}), \mathbf{y}) = q$, then go to Step 2.
}

Step 2 If $w(\mathbb{p}) \leq \varepsilon$, then $\mathbb{K}_{ind} = \mathbb{K}_{ind} \cup \mathbb{p}$,
else {bisect $\mathbb{p}$ along a principal plane and stack the two resulting boxes}.

Step 3 If the stack is not empty,
then unstack into $\mathbb{p}$ and go to Step 1.

End. ◊

If, for any $q > 0$, the current box $\mathbb{p}$ satisfies the first condition of the loop in Step 1, then from Proposition 1, $\mathbb{p} \subset \mathbb{S}_q$. Moreover, $\text{sep}(\mathbf{y}_m(\mathbb{p}), \mathbf{y}) > q - 1$, otherwise, the second condition of Step 1 would have been satisfied during a previous iteration. Therefore, $\mathbb{p} \cap \mathbb{S}_{q-1} = \varnothing$.

Test case: For the bounded-error estimation problem presented above, with an accuracy $\varepsilon = 0.005$ and $q_{max} = 3$, ISOON is completed in less than 3 minutes and the paving generated is presented in Figure 8.

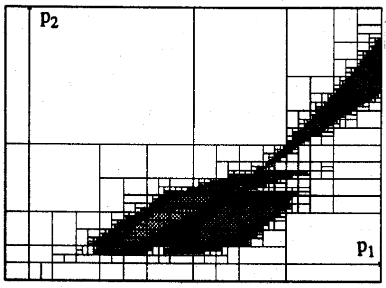

Figure 8: Paving generated by ISOON from contaminated experimental data. $\mathbb{K}_{iso}(1)$ is in light grey, $\mathbb{K}_{iso}(2)$ in medium grey and $\mathbb{K}_{iso}(3)$ in dark grey. The frame corresponds to the search domain $[-0.1, 1.5] \times [-0.1, 1.5]$.

5. CONCLUSIONS

The implementations of OMNE available so far could not guarantee their results, because of the random nature of the search. The new algorithm GOMNE now makes it possible to characterize the set of all parameter vectors that are consistent with the largest possible number of data in a guaranteed way. The data points considered as outliers can be indicated. They may vary when the parameters describe this set.

To protect oneself against undetected outliers, one may choose to decrease the number of data to be consistent with. The study of the evolution of the estimated set with the number of rejected data is facilitated by the new algorithm ISOON.

6. REFERENCES

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

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

Lahanier H., E. Walter and R. Gomeni (1987). OMNE: a new robust membership set estimator for the parameters of nonlinear models, *J. Pharmacokin. Biopharm.* **15**, 203-219.

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

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

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

Norton, J.P. (Ed.) (1994). Bounded-Error Estimation 1, Special issue of *Int. J. Adapt. Control Signal Process.*, **8**, 1-118.

Norton, J.P. (Ed.) (1995). Bounded-Error Estimation 2, Special issue of *Int. J. Adapt. Control Signal Process.*, **9**, 1-132.

Piet-Lahanier, H., and E. Walter (1990). Characterization of non-connected parameter uncertainty regions, *Math. Comput. Simulation*, **32**, 553-560.

Pronzato, L., and E. Walter (1996). Robustness to outliers of bounded-error estimators and consequences on experiment design. In Milanese, M., J. Norton, H. Piet-Lahanier and E. Walter (Eds.) (1996). *Bounding Approaches to System Identification*, Plenum Press, New York, pp. 199-212.

Ratschek, D., and J. Rokne (1988). *Computer Methods for the Range of Functions*, Ellis-Horwood—Wiley, New York.

Walter, E. (Ed.) (1990). Parameter Identifications with Error Bound, special issue of *Math. Comput. Simulation*, **32**, 447-607.

Walter, E., and H. Piet-Lahanier (1988). Estimation of the parameter uncertainty resulting from bounded-error data, *Math. Biosci.*, **92**, 55-74.