

Set Inversion via Interval Analysis for Nonlinear Bounded-error Estimation*

LUC JAULIN and ERIC WALTER†

Finding all parameter vectors that are consistent with the data in the sense that the error falls within prior bounds is a problem of set inversion, solved in the general nonlinear case via interval analysis.

Key Words—Bounded errors; global analysis; guaranteed estimates; identification; interval analysis; nonlinear equations; nonlinear estimation; parameter estimation; set theory; set inversion.

Abstract—In the context of bounded-error estimation, one is interested in characterizing the set of all the values of the parameters to be estimated that are consistent with the data in the sense that the errors between the data and model outputs fall within prior bounds. While the problem can be considered as solved when the model output is linear in the parameters, the situation is far less advanced in the general nonlinear case. In this paper, the problem of nonlinear bounded-error estimation is viewed as one of set inversion. An original algorithm is proposed, based upon interval analysis, that makes it possible to characterize the feasible set for the parameters by enclosing it between internal and external unions of boxes. The convergence of the algorithm is proved and the algorithm is applied to two test cases. The results obtained are compared with those provided by signomial analysis.

1. INTRODUCTION TO BOUNDED-ERROR ESTIMATION

THIS PAPER IS concerned with the problem of estimating the unknown parameters of a model from experimental data collected on a system under given experimental conditions (e.g. location of input and output ports, shape of inputs, measurement times). Let $\mathbf{y} \in \mathbb{R}^{n_y}$ be the vector of all these data. It may consist for instance of n_y scalar measurements performed at given times on a single-input–single-output dynamical system, but multi-input–multi-output dynamical systems or static processes could be considered as well. A parametric structure $M(\cdot)$ is assumed for the model of this system, i.e. a set

of models parametrized by a vector $\mathbf{p} \in \mathbb{P} \subset \mathbb{R}^{n_p}$ to be estimated, where \mathbb{P} is the prior feasible set for the parameters. For the experimental conditions used, each model $M(\mathbf{p})$ generates a vector output $\mathbf{y}_m(\mathbf{p})$ homogeneous to the data \mathbf{y} . The dependency of \mathbf{y}_m in the experimental conditions needs not to be made explicit since these are assumed fixed. The vector function \mathbf{y}_m is assumed to be continuous and locally one-to-one (or identifiable). 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 (e.g. Walter, 1990; and the many papers on the subject in Bányász and Keviczky, 1991) it is assumed that $\mathbf{e}_m(\mathbf{p})$ must belong to some prior feasible set $\mathbb{E} \subset \mathbb{R}^{n_y}$ to be admissible, and the problem to be solved is that of finding the set \mathbb{S} of all admissible values of \mathbf{p} corresponding to an admissible error, i.e.

$$\mathbb{S} = \{\mathbf{p} \in \mathbb{P} \mid \mathbf{e}_m(\mathbf{p}) \in \mathbb{E}\}. \quad (2)$$

In what follows, we shall assume that \mathbb{P} and \mathbb{E} can be defined by finite sets of inequality constraints. For any $\mathbf{p} \in \mathbb{S}$, there exists $\mathbf{e} \in \mathbb{E}$ such that

$$\mathbf{y} = \mathbf{y}_m(\mathbf{p}) + \mathbf{e}. \quad (3)$$

Although set-membership estimation can be set in a purely deterministic context, it can also receive a stochastic interpretation. If \mathbb{P} and \mathbb{E} are, respectively, the support of the prior probability density functions (pdf) for \mathbf{p} and for $\mathbf{e}_m(\mathbf{p})$ then, from Bayes' rule, \mathbb{S} is the support of the posterior pdf for \mathbf{p} . If $\mathbb{P} = \mathbb{R}^{n_p}$ then \mathbb{S} is the set of all values of \mathbf{p} such that the likelihood of the data is nonzero.

The estimation problem can always be reformulated so as to include the inequalities

* Received 6 March 1992; revised 14 September 1992; received in final form 15 December 1992. The original version of this paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor P. van den Hof under the direction of Editor T. Söderström.

† Laboratoire des Signaux et Systèmes, CNRS-Ecole Supérieure d'Electricité, Plateau de Moulon, 91192 Gif-sur-Yvette Cedex, France. Tel: (33) 1 69 41 80 40, Fax: (33) 1 69 41 30 60.

defining \mathbb{P} among those defining \mathbb{E} , so that only one prior feasible set needs to be considered. \mathbb{S} can then equivalently be defined as

$$\mathbb{S} = \mathbf{y}_m^{-1}(\mathbf{y} - \mathbb{E}) = \mathbf{y}_m^{-1}(\mathbb{Y}) = \mathbf{e}_m^{-1}(\mathbb{E}), \quad (4)$$

where \mathbf{e}_m^{-1} and \mathbf{y}_m^{-1} are the reciprocal functions (in a set-theoretic sense) of \mathbf{e}_m and \mathbf{y}_m , and where $\mathbb{Y} = \mathbf{y} - \mathbb{E}$ is the prior feasible set for the model outputs. The problem to be solved thus appears as one of set inversion.

The interested reader is referred to the surveys (Kurzanski and Vályi, 1991; Milanese and Vicino, 1991a; Norton, 1987a; Walter and Piet-Lahanier, 1990) for a bibliography on bounded-error estimation. The methods to be used for characterizing \mathbb{S} depend on whether $\mathbf{e}_m(\mathbf{p})$ is affine (linear) in \mathbf{p} or not. In the first case, the problem of guaranteed estimation can be considered as solved when \mathbb{E} is a box. Methods exist to characterize \mathbb{S} exactly and recursively (Broman and Shensa, 1990; Mo and Norton, 1990; Walter and Piet-Lahanier, 1989). Ellipsoids and boxes guaranteed to contain \mathbb{S} can also be computed (Belforte *et al.*, 1990; Fogel and Huang, 1982; Kurzanski and Vályi, 1991; Milanese and Belforte, 1982; Pronzato *et al.*, 1989). In the latter case, when the error is nonlinear in the parameters, guaranteed bounded-error estimation is far less advanced. For some types of output-error models, it has been proven (Clément and Gentil, 1990; Norton, 1987b) that \mathbb{S} is contained in a union of convex polyhedra, which can be characterized exactly or enclosed in a union of ellipsoids or boxes.

Whether or not the error is affine in \mathbf{p} , computing the smallest axis-aligned box containing \mathbb{S} can be performed by solving $2n_p$ problems of mathematical programming. Each of them corresponds to the maximization or minimization of a component of \mathbf{p} subject to the n_e inequality constraints that define \mathbb{P} and \mathbb{E} and thus \mathbb{S} . When the error is affine in \mathbf{p} , this can be performed by any of the methods available for linear programming, such as Dantzig's simplex or Karmarkar's algorithm, provided that \mathbb{P} and \mathbb{E} are polytopes. In the general case, global optimization methods are needed if guaranteed results are to be obtained. Among the many methods available for global optimization (e.g. Dixon and Szego, 1975, 1978; Mockus, 1989; Zhigljavsky, 1991) only deterministic methods (Horst and Tuy, 1990; Ratschek and Rokne, 1988) can be used since stochastic methods converge only in probability. In a large number of problems of practical interest (such as the estimation of the parameters of an ARMA model or of a discrete linear state space model),

signomial programming can be used (Milanese and Vicino, 1991b).

All the methods available so far to give a guaranteed characterization of \mathbb{S} in the non-linear case were limited to providing a simple-shaped set guaranteed to contain it. In this paper, we propose a new method to obtain a more detailed description of \mathbb{S} based on the use of interval analysis for set inversion. This approach is similar to the one currently and independently being developed by Moore (1992).

Section 2 describes two test cases which will be used throughout the paper to illustrate the various notions needed. Section 3 presents interval analysis, the basic tool of the new approach. Section 4 formulates the problem of set inversion and gives some theoretical results. A new algorithm for set inversion via interval analysis (SIVIA) is proposed in Section 5. Using a new distance between compact sets introduced in Section 4, its theoretical properties are studied and the results obtained on the two test cases are described.

2. TEST CASES

Two test cases will be considered. The first one deals with approximating a function on a finite interval. It will show that the technique to be described can be used even when the dimension n_y of the data \mathbf{y} is infinite. The second test case has already been studied by Milanese and Vicino (1991b) who estimated the smallest axis-aligned box containing the corresponding set \mathbb{S} . It will be used to compare the results provided by the signomial and interval-analysis approaches, when both apply.

Test case 1. Find the set \mathbb{S} of all values of $\mathbf{p} \in \mathbb{P} = [0, 5] \times [0, 5] \subset \mathbb{R}^2$ such that

$$|e_m(\mathbf{p}, t)| = |y(t) - y_m(\mathbf{p}, t)| \leq 1 \quad \forall t \in [0, 1], \quad (5)$$

where

$$y(t) = t^2 + 2t + 1 \quad (6)$$

and

$$y_m(\mathbf{p}, t) = p_1 \exp(p_2 t).$$

The vector \mathbf{p} is feasible if

$$\forall t \in [0, 1], -1 \leq e_m(\mathbf{p}, t) \leq 1, \quad (7)$$

$$\Leftrightarrow \begin{cases} \min_{0 \leq t \leq 1} y(t) - y_m(\mathbf{p}, t) \geq -1, \\ \max_{0 \leq t \leq 1} y(t) - y_m(\mathbf{p}, t) \leq 1. \end{cases} \quad (8)$$

Thus, the posterior feasible set \mathbb{S} is the same as

if the data $\mathbf{y} = (0, 0)^T$ were to be fitted with the model output

$$\mathbf{y}_m(\mathbf{p}) = \left(\min_{0 \leq t \leq 1} \{t^2 + 2t + 1 - p_1 \exp(p_2 t)\}, \right. \\ \left. \max_{0 \leq t \leq 1} \{t^2 + 2t + 1 - p_1 \exp(p_2 t)\} \right)^T, \quad (9)$$

the admissible set for the error being $\mathbb{E} = \{\mathbf{e} \mid -\mathbf{1} \leq \mathbf{e} \leq \mathbf{1}\}$, where $\mathbf{1}$ represents the two-dimensional vector with all components equal to one.

Test case 2. Assume that at times $\mathbf{t} = (0.75, 1.5, 2.25, 3, 6, 9, 13, 17, 21, 25)^T$, the following data have been recorded (one at a time) on a single-output system:

$$\mathbf{y} = (7.39, 4.09, 1.74, 0.097, -2.57, -2.71, \\ -2.07, -1.44, -0.98, -0.66)^T. \quad (10)$$

The scalar output of the model $M(\mathbf{p})$ at a given time t is described by

$$y_m(\mathbf{p}, t) = p_1 \exp(-p_2 t) + p_3 \exp(-p_4 t). \quad (11)$$

A MATLAB-like notation will be used (see the notation section), so that the vector of the outputs of the model $M(\mathbf{p})$ for all measurement times \mathbf{t} will be denoted by

$$\mathbf{y}_m(\mathbf{p}) = p_1 \exp(-p_2 \mathbf{t}) + p_3 \exp(-p_4 \mathbf{t}). \quad (12)$$

Following (Milanese and Vicino, 1991b), we assume that the set of admissible errors \mathbb{E} is given by

$$\mathbb{E} = [\mathbf{e}] = [-\mathbf{e}_{\max}, \mathbf{e}_{\max}] \\ = \{\mathbf{e} \mid -\mathbf{e}_{\max} \leq \mathbf{e} \leq \mathbf{e}_{\max}\}, \quad (13)$$

where

$$\mathbf{e}_{\max} = 0.05 |\mathbf{y}| + 0.1 * \mathbf{1}. \quad (14)$$

$|\mathbf{y}|$ is the n_y -dimensional vector with each component equal to the absolute value of the corresponding component of \mathbf{y} and $\mathbf{1}$ is a vector of ones. The prior feasible set for the parameter is the box

$$\mathbb{P} = [2, 60] \times [0, 1] \times [-30, -1] \times [0, 0.5]. \quad (15)$$

Figure 1 presents the data. The bars indicate the uncertainty associated with each datum. \mathbb{S} is the set of all values of \mathbf{p} such that each of these bars contains the scalar model output associated with the same time.

3. INTERVAL ANALYSIS

Interval analysis has been a very active field in scientific computation for the last 20 years (e.g. Moore, 1979; Neumaier, 1990; Ratschek and Rokne, 1988). There are now commercially

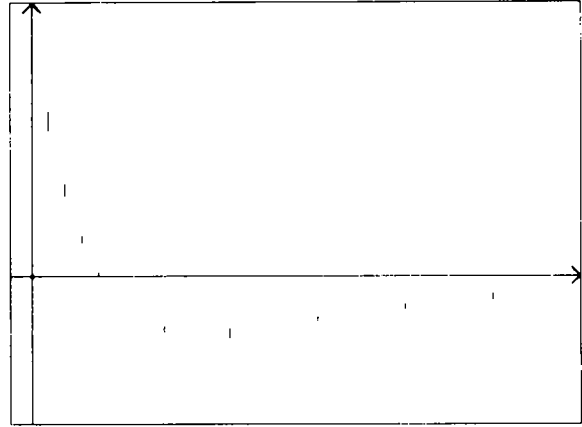


FIG. 1. Data with feasible error bars for Test case 2 in the (t, y) space. The frame corresponds to the domain $[-1, 25] \times [-7, 13]$.

available extensions of FORTRAN and PASCAL that include interval arithmetic among their features (IBM, 1986; Kulisch, 1987). We shall now define the notions of interval analysis that will be used in Section 5 for the description and analysis of SIVIA.

3.1. Boxes

Using boxes in the context of set inversion makes it possible to replace point values of vectors by subsets of the parameter space, thereby allowing a global analysis of infinite sets of points with a finite number of operations. Note that other types of sets based, e.g. on more complex polyhedra or on ellipsoids could be used as well.

Definition 1. An interval $[x]$ of \mathbb{R} (or scalar interval) is a closed, bounded and connected set of real numbers $[x] = [x^-, x^+] = \{x \mid x^- \leq x \leq x^+\}$.

Definition 2. A box $[\mathbf{x}]$ of \mathbb{R}^n (or vector interval) is the cartesian product of n scalar intervals.

The set of all boxes of \mathbb{R}^n will be denoted by \mathbb{IR}^n . Boxes will be specified indifferently in any of the three following ways:

$$[\mathbf{x}] = [x_1^-, x_1^+] \times [x_2^-, x_2^+] \times \cdots \times [x_n^-, x_n^+] \\ = [x_1] \times [x_2] \times \cdots \times [x_n] = [\mathbf{x}^-, \mathbf{x}^+]. \quad (16)$$

Remark 1. Vectors \mathbf{x} of \mathbb{R}^n will also be considered as belonging to \mathbb{IR}^n , with $\mathbf{x}^- = \mathbf{x}^+ = \mathbf{x}$.

Definition 3. The width $w([\mathbf{x}])$ of $[\mathbf{x}] \in \mathbb{IR}^n$, is given by $w([\mathbf{x}]) = \max_{i=1, \dots, n} \{x_i^+ - x_i^-\}$.

Definition 4. The enveloping box $[\mathbb{A}]$ of a bounded subset $\mathbb{A} \subset \mathbb{R}^n$ is the smallest box of \mathbb{IR}^n that contains \mathbb{A} . $[\mathbb{A}] = \bigcap \{[\mathbf{x}] \in \mathbb{IR}^n \mid \mathbb{A} \subset [\mathbf{x}]\}$.

3.2. Minimal inclusion functions

The following definition makes it possible to extend all concepts of vector arithmetic to boxes.

Definition 5. Let \mathbf{f} be a function from \mathbb{R}^n to \mathbb{R}^p . The minimal inclusion function of \mathbf{f} , denoted by $[\mathbf{f}]$, is defined as $[\mathbf{f}]: \mathbb{R}^n \rightarrow \mathbb{R}^p$; $[\mathbf{x}] \rightarrow [\{\mathbf{f}(\mathbf{x}) \mid \mathbf{x} \in [\mathbf{x}]\}]$.

$[\mathbf{f}]([\mathbf{x}])$ is thus the smallest box of \mathbb{R}^p that contains $\mathbf{f}([\mathbf{x}])$, i.e. the enveloping box of $\mathbf{f}([\mathbf{x}])$. It is easy to compute for elementary arithmetic operators and functions.

Example 1. Addition of boxes of \mathbb{R}^2 .

If

$$[\mathbf{x}] = [x_1^-, x_1^+] \times [x_2^-, x_2^+] \\ = [x_1] \times [x_2] = [\mathbf{x}^-, \mathbf{x}^+], \quad (17)$$

$$[\mathbf{y}] = [y_1^-, y_1^+] \times [y_2^-, y_2^+] \\ = [y_1] \times [y_2] = [\mathbf{y}^-, \mathbf{y}^+], \quad (18)$$

and if $+$ is considered as a function from $\mathbb{R}^2 \times \mathbb{R}^2$ to \mathbb{R}^2 then, since $([\mathbf{x}], [\mathbf{y}])$ is a box of \mathbb{R}^4 ,

$$[+](([\mathbf{x}], [\mathbf{y}])) \\ = [x_1^- + y_1^-, x_1^+ + y_1^+] \times [x_2^- + y_2^-, x_2^+ + y_2^+] \\ = [\mathbf{x}^- + \mathbf{y}^-, \mathbf{x}^+ + \mathbf{y}^+]. \quad (19)$$

In what follows, $[+](([\mathbf{x}], [\mathbf{y}]))$ will be denoted by $[\mathbf{x}] + [\mathbf{y}]$, and the same notation will be used for all elementary arithmetic operators.

Example 2. Sum of exponentials.

Consider the function $f: \mathbb{R}^4 \rightarrow \mathbb{R}$; $\mathbf{p} \rightarrow p_1 \exp(p_2) + p_3 \exp(p_4)$. From the monotonicity of the exponential function, we have $\exp([p_i]) = [\exp([p_i])] = [\exp(p_i^-), \exp(p_i^+)]$, and $[f]: \mathbb{R}^4 \rightarrow \mathbb{R}$; $[\mathbf{p}] \rightarrow [p_1] * \exp([p_2]) + [p_3] * \exp([p_4])$. If $p_1 > 0$ and $p_3 > 0$, then $[f](\mathbf{p}^-, \mathbf{p}^+) = [p_1^- \exp(p_2^-) + p_3^- \exp(p_4^-), p_1^+ \exp(p_2^+) + p_3^+ \exp(p_4^+)]$.

Example 3. Sinusoid.

Consider now the function $f: \mathbb{R}^2 \rightarrow \mathbb{R}$; $\mathbf{p} \rightarrow p_1 \sin p_2$. Taking advantage of the fact that the two parameters p_1 and p_2 appear independently in the expression of f , we obtain $[f]: \mathbb{R}^2 \rightarrow \mathbb{R}$; $[\mathbf{p}] \rightarrow [p_1] * [\sin]([p_2])$, where the minimal inclusion function of \sin , denoted by $[\sin] = [\sin^-, \sin^+]$ is defined by:

If $\exists k \in \mathbb{Z}$ such that $2k\pi - \pi/2 \in [p_2]$ then $\sin^-(p_2) = -1$ else $\sin^-(p_2) = \min(\sin p_2^-, \sin p_2^+)$, if $\exists k \in \mathbb{Z}$ such that $2k\pi + \pi/2 \in [p_2]$ then $\sin^+(p_2) = 1$ else $\sin^+(p_2) = \max(\sin p_2^-, \sin p_2^+)$, $[\sin]([p_2]) = [\sin^-(p_2), \sin^+(p_2)]$.

Test case 1. Since $p_1 \geq 0$, the minimal inclusion function for $\mathbf{y}_m(\mathbf{p})$ as defined by (9) is given by

$$[\mathbf{y}_m](\mathbf{p}) \\ = \left[\min_{0 \leq t \leq 1} \{t^2 + 2t + 1 - p_1^+ \exp(p_2^+ t)\}, \right. \\ \left. \min_{0 \leq t \leq 1} \{t^2 + 2t + 1 - p_1^- \exp(p_2^- t)\} \right] \\ \times \left[\max_{0 \leq t \leq 1} \{t^2 + 2t + 1 - p_1^+ \exp(p_2^+ t)\}, \right. \\ \left. \max_{0 \leq t \leq 1} \{t^2 + 2t + 1 - p_1^- \exp(p_2^- t)\} \right]. \quad (20)$$

One evaluation of $[\mathbf{y}_m](\mathbf{p})$ thus amounts to solving four simple one-dimensional optimization problems. It is trivial to prove that each of them can be solved using any local optimization method twice (once initialized at $t=0$, then at $t=1$).

Test case 2. As $p_1 \geq 0$, $p_3 \leq 0$ and $t > 0$, the minimal inclusion function for $\mathbf{y}_m(\mathbf{p})$ as defined by (12) is

$$[\mathbf{y}_m](\mathbf{p}) = [p_1^- \exp(-p_2^+ t) + p_3^- \exp(-p_4^- t), \\ p_1^+ \exp(-p_2^- t) + p_3^+ \exp(-p_4^+ t)]. \quad (21)$$

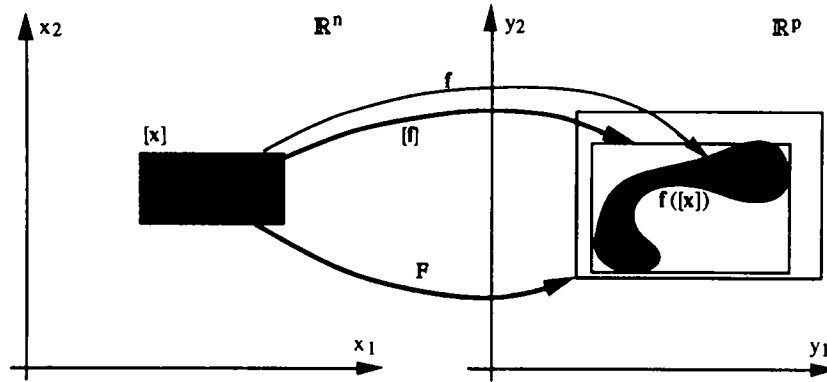
3.3. Inclusion functions

When $[\mathbf{f}]$ cannot be computed, it can be approximated by a (nonminimal) inclusion function \mathbb{F} .

Definition 6. $\mathbb{F}: \mathbb{R}^n \rightarrow \mathbb{R}^p$ is an inclusion function of $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^p$ if (i) $\forall [\mathbf{x}] \in \mathbb{R}^n$, $\mathbf{f}([\mathbf{x}]) \subset \mathbb{F}([\mathbf{x}])$ and (ii) $w([\mathbf{x}]) \rightarrow 0 \Rightarrow w(\mathbb{F}([\mathbf{x}])) \rightarrow 0$.

Remark 2. An inclusion function \mathbb{F} exists if and only if \mathbf{f} is continuous. Contrary to the minimal inclusion function $[\mathbf{f}]$, \mathbb{F} is not unique, and $[\mathbf{f}] \subset \mathbb{F}$. Any intersection of inclusion functions is an inclusion function.

Figure 2 illustrates Definitions 5 and 6. For any function \mathbf{f} obtained by composition of elementary operators such as $+$, $-$, $*$, $/$, \sin , \cos , \exp , \dots , it is easy to obtain an inclusion function by replacing each of these elementary operators by its minimal inclusion function in the formal expression of \mathbf{f} (Moore, 1979). The resulting inclusion function is called *natural interval extension* in the literature. Relaxing Definition 6 by discarding Condition (ii), it is also possible to take into account the effect of rounding in the computation so as to obtain intervals guaranteed to contain the exact

FIG. 2. Minimal inclusion function $[f]$ and inclusion function F of a function f .

mathematical solutions. It must be noted, however, that the algorithms presented in this paper are guaranteed to converge only if (ii) is valid.

Example 4. Consider the function $f(x) = x^2 - x$. A first inclusion function is given by its natural interval extension $F_1([x]) = [x]^2 - [x]$. It is not the minimal inclusion function $[f]$, since for instance

$$\begin{aligned} F_1([-1, 3]) &= [-1, 3]^2 - [-1, 3] \\ &= [0, 9] + [-3, 1] = [-3, 10], \end{aligned} \quad (22)$$

when it is trivial to show that $[f]([-1, 3]) = [-0.25, 6] \subset F_1([-1, 3])$. Another inclusion function can be obtained by factorizing $f(x)$ as $x * (x - 1)$ and by writing $F_2([x]) = [x] * ([x] - 1)$, so that

$$\begin{aligned} F_2([-1, 3]) &= [-1, 3] * ([-1, 3] - 1) \\ &= [-1, 3] * [-2, 2] = [-6, 6]. \end{aligned} \quad (23)$$

A better inclusion function is given by $F_3 = F_1 \cap F_2$, such that $F_3([-1, 3]) = [-3, 6]$.

Thus, depending on the expression used for $f(x)$, substitution of minimal inclusion functions for the elementary operators yields different inclusion functions. How to choose the expression of $f(x)$ so as to obtain the smallest possible inclusion function apparently remains an open question.

3.4. Subpavings

Exploration algorithms aim at covering the space of interest. In the context of interval analysis, covering is performed with sets of boxes, which corresponds to the following notions.

Definition 7. A subpaving of \mathbb{R}^n is a set of non-overlapping boxes of \mathbb{R}^n with nonzero width.

Definition 8. 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} .

When no ambiguity can arise, the notation \mathbb{K} will also be used for the subset of \mathbb{R}^n generated by the union of all boxes of the subpaving \mathbb{K} .

Definition 9. The accumulation set of a subpaving \mathbb{K} is the limit of the subset of \mathbb{R}^n formed by the union of all boxes of \mathbb{K} with width lower than ε when ε tends to zero.

Remark 3. Since subpavings only contain boxes with nonzero width, the accumulation set of a finite subpaving is necessarily void.

4. SET INVERSION AND SET ENCLOSURE

We shall now formulate the problem of computing the inverse of a set by enclosing it between subpavings. We shall also introduce a new distance and the associated notion of continuity to be used for the analysis of the convergence of the new algorithm. The reader only interested in the algorithm can skip this section and proceed directly to Section 5. In what follows, all sets considered will be supposed regular enough for their boundaries to be well defined.

Set inversion problem. Let f be a continuous function from \mathbb{R}^n to \mathbb{R}^p . Let \mathbb{Y} belong to $C(\mathbb{R}^p)$, the set of all compact subsets of \mathbb{R}^p . The set inversion problem is that of characterizing the set \mathbb{X} such that $\mathbb{X} = f^{-1}(\mathbb{Y})$. A possible approach consists of enclosing \mathbb{X} between finite subpavings \mathbb{K}_{in} and \mathbb{K}_{out} in the sense that $\mathbb{K}_{in} \subset \mathbb{X} \subset \mathbb{K}_{out}$. The notions of set inversion and enclosure of compact sets between subpavings are illustrated by Fig. 3. Recall that $f(\mathbb{X}) = f \circ f^{-1}(\mathbb{Y}) = \mathbb{Y} \cap f(\mathbb{R}^n) \subset \mathbb{Y}$, with $f(\mathbb{X}) = \mathbb{Y}$ if f is a mapping of \mathbb{R}^n onto \mathbb{R}^p .

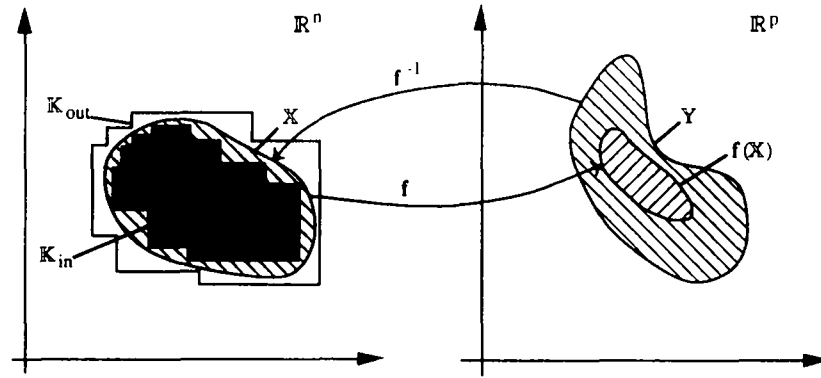


FIG. 3. Set inversion problem and enclosure of the solution compact.

Table 1 gives the interpretation in the context of bounded-error estimation of the set inversion problem as illustrated by Fig. 3.

4.1. Distances

Various distances between compacts will be needed to prove the convergence of the algorithm presented in Section 5.

4.1.1. *Hausdorff distance.* The separation between two subsets A and B of \mathbb{R}^n is given by

$$L_\infty(A, B) = \inf_{a \in A, b \in B} L_\infty(a, b), \quad (24)$$

where $L_\infty(a, b)$ is the distance between a and b induced by the L_∞ -norm. The proximity of A to B will be characterized by

$$h_\infty^0(A, B) = \inf \{r \in \mathbb{R}^+ \mid A \subset B + rU\}, \quad (25)$$

where U is the unit sphere in (\mathbb{R}^n, L_∞) . Note that $h_\infty^0(\cdot, \cdot)$ is not a symmetric operator, and that the proximity may be infinite if A is unbounded. We have

$$h_\infty^0(A, B) = \sup_{a \in A} h_\infty^0(a, B)$$

(here, a is considered as a singleton), (26)

$$A \subset B \Rightarrow h_\infty^0(A, B) = 0, \quad (27)$$

$$h_\infty^0(A, \partial B) = \max \{h_\infty^0(A, B), h_\infty^0(A, \bar{B})\}, \quad (28)$$

$$A \subset B \Rightarrow h_\infty^0(A, C) \leq h_\infty^0(B, C), \quad (29)$$

where \bar{B} is the complement of B in \mathbb{R}^n and ∂B is the boundary of B .

TABLE 1. CORRESPONDENCE BETWEEN THE SET INVERSION PROBLEM AS SUMMARIZED IN FIG. 3 AND BOUNDED-ERROR ESTIMATION

Fig. 3	Bounded-error estimation
\mathbb{R}^n	Parameter space \mathbb{R}^n
\mathbb{R}^p	Data (or error) space \mathbb{R}^p
f	Model output y_m (or error e_m)
Y	Feasible set for the model output Y (or error ϵ)
X	Posterior feasible set for the parameters S
$f(X)$	Set of all model outputs (or errors) associated with S

Proof. We just give a proof for (28), the others being trivial. From (26), $h_\infty^0(A, \partial B) = \sup_{a \in A} h_\infty^0(a, \partial B)$. Now

$$h_\infty^0(a, \partial B) = \max \{h_\infty^0(a, B), h_\infty^0(a, \bar{B})\}.$$

Thus

$$\begin{aligned} h_\infty^0(A, \partial B) &= \sup_{a \in A} \max \{h_\infty^0(a, B), h_\infty^0(a, \bar{B})\} \\ &= \max \left\{ \sup_{a \in A} h_\infty^0(a, B), \sup_{a \in A} h_\infty^0(a, \bar{B}) \right\} \\ &\stackrel{(26)}{=} \max \{h_\infty^0(A, B), h_\infty^0(A, \bar{B})\}. \end{aligned}$$

Definition 10. The Hausdorff distance (e.g. Berger, 1979) between two subsets A and B of \mathbb{R}^n is given by $h_\infty(A, B) = \max \{h_\infty^0(A, B), h_\infty^0(B, A)\}$.

The operator h_∞ is a distance for $C(\mathbb{R}^n)$, i.e. for any A, B and C belonging to $C(\mathbb{R}^n)$, it satisfies (i) $h_\infty(A, B) = 0 \Rightarrow A = B$, (ii) $h_\infty(A, B) = h_\infty(B, A)$ and (iii) $h_\infty(A, C) \leq h_\infty(A, B) + h_\infty(B, C)$. We have

$$h_\infty(A, B) = \inf \{r \mid A \subset B + rU \text{ and } B \subset A + rU\}, \quad (30)$$

$$B \subset A \Rightarrow h_\infty^0(A, B) = h_\infty(A, B), \quad (31)$$

$$h_\infty^0(A, B) = h_\infty(B, A \cup B), \quad (32)$$

$$h_\infty^0(A, \partial B) = \max \{h_\infty(B, A \cup B), h_\infty(\bar{B}, A \cup \bar{B})\}, \quad (33)$$

$$h_\infty(A, B) \geq h_\infty(A \cup C, B \cup C). \quad (34)$$

The proofs for (30), (31), (32) and (34) are trivial, (33) follows directly from (28) and (32).

If the compact obtained by adding to a compact A a single point far from it is h_∞ -far from A , the compact obtained by moving the boundary of A slightly and drilling a finite

number of small holes in the result is h_∞ -close to A . This illustrates the coarseness of the characterization of differences between compacts provided by the Hausdorff distance. A finer characterization will be needed to define the convergence conditions of Theorem 1 in Section 5.2.1, which motivates the introduction of new distances.

4.1.2. Complementary Hausdorff distance.

Definition 11. The complementary Hausdorff distance between two subsets A and B of \mathbb{R}^n is given by $\tilde{h}_\infty(A, B) = h_\infty(\tilde{A}, \tilde{B})$.

Proposition 1. The operator \tilde{h}_∞ is a semi-distance on $C(\mathbb{R}^n)$.

Proof. (i) The operator \tilde{h}_∞ is not a distance since $\tilde{h}_\infty(A, B) = 0$ whenever A and B are singletons. (ii) The symmetry of \tilde{h}_∞ results from Definition 11. Let us now check the triangular inequality (iii). For any A, B and C belonging to $C(\mathbb{R}^n)$, $\tilde{h}_\infty(A, B) = h_\infty(\tilde{A}, \tilde{B}) = h_\infty(\text{clo}(\tilde{A}), \text{clo}(\tilde{B})) = h_\infty(\text{clo}(\tilde{A}) \cap Z, \text{clo}(\tilde{B}) \cap Z)$ where Z is a very large compact sphere containing A, B and C . The triangular inequality satisfied by h_∞ when operating on compacts implies: $h_\infty(\text{clo}(\tilde{A}) \cap Z, \text{clo}(\tilde{B}) \cap Z) \leq h_\infty(\text{clo}(\tilde{A}) \cap Z, \text{clo}(\tilde{C}) \cap Z) + h_\infty(\text{clo}(\tilde{C}) \cap Z, \text{clo}(\tilde{B}) \cap Z) = h_\infty(\tilde{A}, \tilde{C}) + h_\infty(\tilde{C}, \tilde{B}) = \tilde{h}_\infty(A, C) + \tilde{h}_\infty(C, B)$.

As h_∞ , the complementary Hausdorff distance \tilde{h}_∞ fails to give a fine characterization of the difference between compacts as illustrated by the following example. If the compact obtained by drilling a single small hole in a compact A is not \tilde{h}_∞ -close to A , the compact obtained by moving the boundary of A slightly and adding a finite number of vectors far from A to the result is \tilde{h}_∞ -close to A .

4.1.3. Generalized Hausdorff distance. Combining h_∞ and \tilde{h}_∞ , it is possible to obtain a new distance that avoids the defects of each of them.

Definition 12. The generalized Hausdorff distance between two subsets A and B of \mathbb{R}^n is given by $m_\infty(A, B) = \max(h_\infty(A, B), \tilde{h}_\infty(A, B))$.

Proposition 2. The operator m_∞ is a distance on $C(\mathbb{R}^n)$.

Proof. (i) If $m_\infty(A, B) = 0$ then $h_\infty(A, B) = 0$ thus $A = B$. (ii) m_∞ is obviously symmetric. (iii) For any A, B and C belonging to $C(\mathbb{R}^n)$, the triangular inequalities satisfied by h_∞ and \tilde{h}_∞ imply that $m_\infty(A, B) = \max(h_\infty(A, B), \tilde{h}_\infty(A, B)) \leq \max(h_\infty(A, C) + h_\infty(C, B), \tilde{h}_\infty(A, C) + \tilde{h}_\infty(C, B)) \leq \max(h_\infty(A, C), \tilde{h}_\infty(A, C)) + \max(h_\infty(C, B), \tilde{h}_\infty(C, B)) = m_\infty(A, C) + m_\infty(C, B)$, so that the triangular inequality is satisfied.

4.2. Compact enclosure between two subpavings

Let $(C(\mathbb{R}^n), \subset, m_\infty)$ be the set of all compacts in \mathbb{R}^n equipped with the partial ordering \subset and the distance m_∞ . The set of finite subpavings is dense from outside in $(C(\mathbb{R}^n), \subset, m_\infty)$, i.e. we can find an external subpaving K_{out} as m_∞ -close to every compact as desired. This, however, does not necessarily hold true for an internal subpaving. Consider, for instance, a segment of a line in \mathbb{R}^2 . It can be approximated as closely as desired by a subpaving of \mathbb{R}^2 from the outside but not from the inside. To avoid this type of problem, we will sometimes restrict consideration to the (large) class of compacts defined as follows.

Definition 13. The compact A is full if $\text{clo}(\text{int}(A)) = A$.

The set of all full compact subsets of \mathbb{R}^n will be denoted by $C_f(\mathbb{R}^n)$. For any $A \in C_f(\mathbb{R}^n)$, there exist sequences of subpavings $K_{\text{in}}(k)$ such that

$$\begin{cases} K_{\text{in}}(k) \subset A \\ m_\infty(K_{\text{in}}(k), A) \rightarrow 0. \end{cases} \quad (35)$$

Proposition 3. \tilde{h}_∞ is a distance on $C_f(\mathbb{R}^n)$.

Proof. (i) For any A and B belonging to $C_f(\mathbb{R}^n)$, $\tilde{h}_\infty(A, B) = h_\infty(\tilde{A}, \tilde{B}) = h_\infty(\text{clo}(\tilde{A}), \text{clo}(\tilde{B})) = h_\infty(\text{clo}(\tilde{A}) \cap Z, \text{clo}(\tilde{B}) \cap Z)$, where Z is a very large compact sphere containing A and B . Therefore $\tilde{h}_\infty(A, B) = 0 \Rightarrow h_\infty(\text{clo}(\tilde{A}) \cap Z, \text{clo}(\tilde{B}) \cap Z) = 0$, which implies, as h_∞ is a distance in $C(\mathbb{R}^n)$, $\text{clo}(\tilde{A}) \cap Z = \text{clo}(\tilde{B}) \cap Z \Leftrightarrow \text{clo}(\tilde{A}) = \text{clo}(\tilde{B}) \Leftrightarrow \text{int}(A) = \text{int}(B) \Rightarrow \text{clo}(\text{int}(A)) = \text{clo}(\text{int}(B))$, so that $A = B$ from Definition 13. The two other properties (ii) and (iii) of a distance result from Proposition 1.

All finite subpavings of \mathbb{R}^n belong to $C_f(\mathbb{R}^n)$. The set of finite subpavings is dense from inside and outside in $(C_f(\mathbb{R}^n), \subset, m_\infty)$. Thus for any $X \in C_f(\mathbb{R}^n)$, it is possible to find finite subpavings K_{in} and K_{out} such that $K_{\text{in}} \subset X \subset K_{\text{out}}$ and that the subset $\mathcal{V}(X)$ of $C(\mathbb{R}^n)$ consisting of all compacts X' such that $K_{\text{in}} \subset X' \subset K_{\text{out}}$ has a diameter $m_\infty(K_{\text{in}}, K_{\text{out}})$ as small as desired. $\mathcal{V}(X)$ is therefore a neighborhood of X , so that X is enclosed between K_{in} and K_{out} .

Enclosure of a characteristic $Z(X)$. Let Z be an increasing function from $(C(\mathbb{R}^n), \subset)$ to the

partially-ordered metric set (Z, \ll) . $Z(\mathbb{X})$ may for example be its volume $\text{vol}(\mathbb{X})$, its enveloping box $[\mathbb{X}]$, or any function of \mathbb{X} resulting from the maximization of a convex criterion on \mathbb{X} . All these characteristics are very easy to compute for subpavings. If $\mathbb{K}_{\text{in}} \subset \mathbb{X} \subset \mathbb{K}_{\text{out}}$, then $Z(\mathbb{K}_{\text{in}}) \ll Z(\mathbb{X}) \ll Z(\mathbb{K}_{\text{out}})$. If Z is continuous around \mathbb{X} , then $Z(\mathbb{K}_{\text{in}}) \xrightarrow{m_\infty} Z(\mathbb{X})$ and $Z(\mathbb{K}_{\text{out}}) \xrightarrow{m_\infty} Z(\mathbb{X})$ when $m_\infty(\mathbb{K}_{\text{in}}, \mathbb{K}_{\text{out}}) \rightarrow 0$.

Next section proposes an algorithm that encloses the solution $\mathbb{X} = \mathbf{f}^{-1}(\mathbb{Y})$ of a set-inversion problem between two finite subpavings bracketing a neighborhood of \mathbb{X} in $(C(\mathbb{R}^n), m_\infty)$ with a diameter as small as desired, provided that \mathbf{f}^{-1} is h_∞ and \tilde{h}_∞ continuous around \mathbb{Y} and \mathbb{X} is full. This algorithm will therefore make it possible to bracket any monotonic characteristic $Z(\mathbb{X})$ continuous around \mathbb{X} as precisely as desired.

5. ALGORITHM FOR SET INVERSION

Set inversion (i.e. finding $\mathbb{X} = \mathbf{f}^{-1}(\mathbb{Y})$ given \mathbf{f} and \mathbb{Y}) will be addressed in a general setting, before specializing the result to parameter estimation. We shall assume that \mathbb{X} is bounded and included in a prior box $[\mathbf{x}](0)$, used as the initial search domain.

5.1. Set inversion via interval analysis

SIVIA (Set Inverter Via Interval Analysis) applies to any function \mathbf{f} for which an inclusion function \mathbb{F} can be computed. Note that this class is not restricted to explicit functions, since inclusion functions exist for solutions of differential equations. We shall say that a box $[\mathbf{x}]$ of \mathbb{R}^n is feasible if $[\mathbf{x}] \subset \mathbb{X}$ and unfeasible if $[\mathbf{x}] \cap \mathbb{X} = \emptyset$,

else $[\mathbf{x}]$ is ambiguous. Interval analysis gives us two conditions, illustrated by Fig. 4, for deciding the feasibility of a box $[\mathbf{x}]$. If $\mathbb{F}([\mathbf{x}]) \subset \mathbb{Y}$ then $[\mathbf{x}] \subset \mathbb{X}$, so that $[\mathbf{x}]$ is feasible. If $\mathbb{F}([\mathbf{x}]) \cap \mathbb{Y} = \emptyset$ then $[\mathbf{x}] \cap \mathbb{X} = \emptyset$, so that $[\mathbf{x}]$ is unfeasible. In all other cases, the box $[\mathbf{x}]$ will be said to be indeterminate. Note that an indeterminate box is not necessarily ambiguous, but could be feasible or unfeasible as well.

SIVIA makes an extensive use of a stack of boxes. A stack is a dynamical structure on which only three operations are possible. One may stack, i.e. put an element on top of the stack, unstack, i.e. remove the element located on top of the stack or test the stack for emptiness. We shall call principal plane of a box $[\mathbf{x}]$ a symmetry plane of this box that is orthogonal to an axis i of maximal length, i.e. $i \in \{j \mid w([\mathbf{x}]) = w([x_j])\}$. In what follows, \mathbb{K}_{in} and \mathbb{K}_i will, respectively denote the subpavings of all feasible and indeterminate boxes (Fig. 5); $[\mathbf{x}](k)$ will be the box considered at iteration k , and ε_r will denote the accuracy required for the paving. Upon completion of the algorithm, all indeterminate boxes will have a width lower than or equal to ε_r .

The basic structure of SIVIA can be described as follows:

Program inputs

Inclusion function: \mathbb{F}

Set to be inverted: \mathbb{Y}

Prior feasible box: $[\mathbf{x}](0)$

Required accuracy for the paving: ε_r

Initialization

$k = 0$, $\text{stack} = \emptyset$, $\mathbb{K}_{\text{in}} = \emptyset$, $\mathbb{K}_i = \emptyset$.

Iteration k

Step 1: If $\mathbb{F}([\mathbf{x}](k)) \subset \mathbb{Y}$, then $\mathbb{K}_{\text{in}} = \mathbb{K}_{\text{in}} \cup [\mathbf{x}](k)$. Go to Step 4.

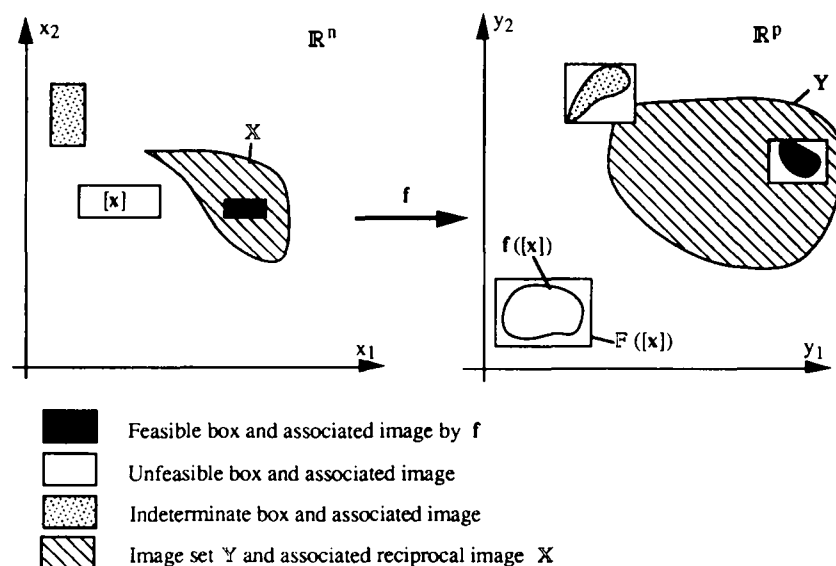


FIG. 4. Feasibility of boxes.

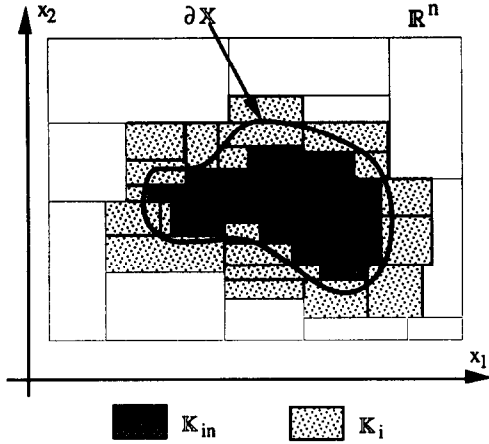


FIG. 5. Enclosure of \mathbb{X} between two subpavings ($\mathbb{K}_{\text{out}} = \mathbb{K}_{\text{in}} \cup \mathbb{K}_i$).

- Step 2: If $\mathbb{F}([\mathbf{x}](k)) \cap \mathbb{Y} = \emptyset$, then go to Step 4.
 Step 3: If $w([\mathbf{x}](k)) \leq \varepsilon_r$, then $\mathbb{K}_i = \mathbb{K}_i \cup [\mathbf{x}](k)$, else bisect $[\mathbf{x}](k)$ along a principal plane and stack the two resulting boxes.
 Step 4: If the stack is not empty, then unstack into $[\mathbf{x}](k+1)$, increment k and go to Step 1.

End.

When one is only interested in specific properties of \mathbb{X} , special care can be taken to avoid memorizing \mathbb{K}_{in} and \mathbb{K}_i by use of suitable exhaustive summaries of the properties of \mathbb{K}_{in} and \mathbb{K}_i of interest (Jaulin and Walter, 1993).

5.2. Properties of SIVIA

SIVIA is a finite algorithm, which terminates in less than $\{w([\mathbf{x}](0))/\varepsilon_r + 1\}^n$ iterations. It provides the subpavings \mathbb{K}_{in} and \mathbb{K}_i (the dependency of these subpavings in ε_r will be omitted for notational simplicity).

5.2.1. *Convergence.* Let us prove that the enclosure of \mathbb{X} generated by SIVIA as

$$\mathbb{K}_{\text{in}} \subset \mathbb{X} \subset \mathbb{K}_{\text{out}} \triangleq \mathbb{K}_{\text{in}} \cup \mathbb{K}_i, \quad (36)$$

defines a neighborhood of \mathbb{X} with a diameter that converges to zero when ε_r tends to zero.

Lemma 1. $\lim_{\varepsilon_r \rightarrow 0} h_\infty^0(\mathbb{f}(\mathbb{K}_i), \partial\mathbb{Y}) = 0$.

Proof. If $[\mathbf{x}] \in \mathbb{K}_i$, then $w([\mathbf{x}]) \leq \varepsilon_r$, and the inclusion function \mathbb{F} (to be provided as a program input) satisfies $w(\mathbb{F}([\mathbf{x}])) \rightarrow 0$ when $\varepsilon_r \rightarrow 0$. Now, $\mathbb{F}([\mathbf{x}])$ is neither inside nor outside \mathbb{Y} and thus intersects its boundary $\partial\mathbb{Y}$. Consequently, $\forall [\mathbf{x}] \in \mathbb{K}_i$, $h_\infty^0(\mathbb{f}([\mathbf{x}]), \partial\mathbb{Y}) \rightarrow 0 \Rightarrow h_\infty^0(\mathbb{f}(\mathbb{K}_i), \partial\mathbb{Y}) \rightarrow 0$.

Theorem 1. If \mathbb{f}^{-1} is h_∞ - and \bar{h}_∞ -continuous

around \mathbb{Y} then \mathbb{K}_{in} , \mathbb{K}_{out} and \mathbb{K}_i satisfy

- (i) $\mathbb{K}_i \rightrightarrows \partial\mathbb{X}$,
- (ii) $\mathbb{K}_{\text{out}} \rightrightarrows \mathbb{X}$,
- (iii) $\mathbb{K}_{\text{in}} \overset{\circ}{\rightrightarrows} \mathbb{X}$ (if \mathbb{X} is full),

when ε_r tends to zero, where $\overset{\circ}{\rightrightarrows}$ and \rightrightarrows , respectively, mean the h_∞ -convergence from within and without.

Proof.

Part (i). From (33), $h_\infty^0(\mathbb{f}(\mathbb{K}_i), \partial\mathbb{Y}) = \max \{h_\infty(\mathbb{Y}, \mathbb{f}(\mathbb{K}_i) \cup \mathbb{Y}), h_\infty(\bar{\mathbb{Y}}, \mathbb{f}(\mathbb{K}_i) \cup \bar{\mathbb{Y}})\} = \max \{h_\infty(\mathbb{Y}, \mathbb{f}(\mathbb{K}_i) \cup \mathbb{Y}), \bar{h}_\infty(\mathbb{Y}, \mathbb{f}(\mathbb{K}_i) \cup \bar{\mathbb{Y}})\}$. Lemma 1 therefore implies

$$\begin{cases} h_\infty(\mathbb{Y}, \mathbb{f}(\mathbb{K}_i) \cup \mathbb{Y}) \rightarrow 0, \\ \bar{h}_\infty(\mathbb{Y}, \mathbb{f}(\mathbb{K}_i) \cup \bar{\mathbb{Y}}) \rightarrow 0, \end{cases} \quad (37)$$

when ε_r tends to 0. Since \mathbb{f}^{-1} is h_∞ and \bar{h}_∞ continuous around \mathbb{Y} , we also have

$$\begin{cases} h_\infty(\mathbb{f}^{-1}(\mathbb{Y}), \mathbb{f}^{-1}(\mathbb{f}(\mathbb{K}_i) \cup \mathbb{Y})) \rightarrow 0, \\ \bar{h}_\infty(\mathbb{f}^{-1}(\mathbb{Y}), \mathbb{f}^{-1}(\mathbb{f}(\mathbb{K}_i) \cup \bar{\mathbb{Y}})) \rightarrow 0. \end{cases} \quad (38)$$

Noting Definition 11 and replacing $\mathbb{f}^{-1}(\mathbb{Y})$ by \mathbb{X} , we obtain:

$$\begin{cases} h_\infty(\mathbb{X}, \mathbb{f}^{-1} \circ \mathbb{f}(\mathbb{K}_i) \cup \mathbb{X}) \rightarrow 0, \\ h_\infty(\bar{\mathbb{X}}, \mathbb{f}^{-1} \circ \mathbb{f}(\mathbb{K}_i) \cup \bar{\mathbb{X}}) \rightarrow 0. \end{cases} \quad (39)$$

Equations (33) and (39) imply that $h_\infty^0(\mathbb{f}^{-1} \circ \mathbb{f}(\mathbb{K}_i), \partial\mathbb{X}) \rightarrow 0$. Since $\mathbb{K}_i \subset \mathbb{f}^{-1} \circ \mathbb{f}(\mathbb{K}_i)$, (29) then implies that $h_\infty^0(\mathbb{K}_i, \partial\mathbb{X}) \rightarrow 0$. Using (31) and the fact that $\partial\mathbb{X} \subset \mathbb{K}_i$, we finally get $h_\infty^0(\mathbb{K}_i, \partial\mathbb{X}) = h_\infty(\mathbb{K}_i, \partial\mathbb{X}) \rightarrow 0$, which gives (i).

Part (ii). From (34), we have $h_\infty(\mathbb{K}_i, \partial\mathbb{X}) \geq h_\infty(\mathbb{K}_i \cup \mathbb{X}, \partial\mathbb{X} \cup \mathbb{X}) \stackrel{(36)}{=} h_\infty(\mathbb{K}_{\text{out}}, \mathbb{X}) \rightarrow 0$.

Part (iii). Let $\varepsilon > 0$ be an infinitely small real number. Since \mathbb{X} is full, there exists a finite subpaving $\mathbb{K}_1 \subset \mathbb{X}$ such that

$$\begin{cases} h_\infty(\mathbb{K}_1, \mathbb{X}) < 2\varepsilon, \\ L_\infty(\mathbb{K}_1, \partial\mathbb{X}) > \varepsilon. \end{cases} \quad (40)$$

Now, for ε_r sufficiently small, $h_\infty(\mathbb{K}_i, \partial\mathbb{X}) < \varepsilon$, thus $\mathbb{K}_1 \cap \mathbb{K}_i = \emptyset$, i.e. $\mathbb{K}_1 \subset \mathbb{K}_{\text{in}}$. When $\varepsilon \rightarrow 0$, $\mathbb{K}_1 \overset{\circ}{\rightrightarrows} \mathbb{X}$ and then $\mathbb{K}_{\text{in}} \overset{\circ}{\rightrightarrows} \mathbb{X}$.

5.2.2. *Computing time.* The computing time increases exponentially with n (Jaulin and Walter, 1993). This is the main limitation of SIVIA.

5.2.3. *Memory used.* When one is only interested in computing a characteristic of \mathbb{X} such as $[\mathbb{X}]$ or $\text{vol}(\mathbb{X})$, only the stack takes a significant place in memory (Jaulin and Walter, 1993). This place is extraordinarily small, as it

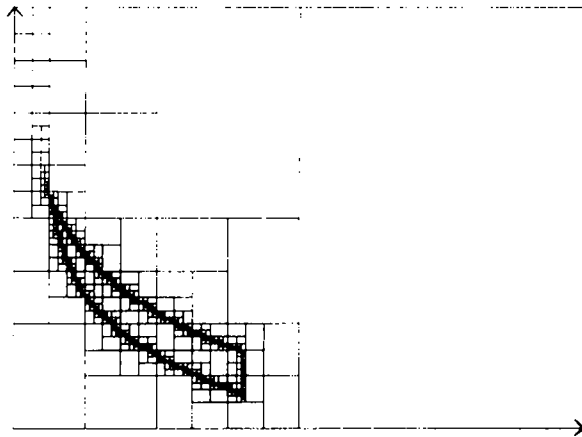


FIG. 6. Paving generated by SIVIA for Test case 1 in the (p_1, p_2) space. The frame corresponds to the search domain $([0, 5])^2$.

can be proved that

$$\#stack < n \cdot \text{int}(\log_2(w([\mathbf{x}](0))) - \log_2(\varepsilon_r) + 1). \quad (41)$$

For instance, if $n = 100$, $w([\mathbf{x}](0)) = 10^4$ and $\varepsilon_r = 10^{-10}$, (41) implies that $\#stack < 4600$.

6. RESULTS OBTAINED ON THE TEST CASES

The estimation of the parameters of the two test cases described in Section 2 will now be performed using SIVIA. The reader is referred to Table 1 for notation.

Test case 1. For a required accuracy $\varepsilon_r = 0.01$, in about 31 sec on a Compaq 386/33, SIVIA produces the paving presented in Fig. 6 while keeping in memory no more than 12 boxes ((41) predicts a number smaller than 18).

The feasible set for the parameters is guaranteed to satisfy

$$[0.342, 1.992] \times [0.420, 2.646] \subset [\mathcal{S}] \\ \subset [0.303, 2.002] \times [0.400, 2.813], \quad (42)$$

$$0.76 \leq \text{vol}(\mathcal{S}) \leq 0.84. \quad (43)$$

The center of $[\mathcal{S}]$, i.e. the Tchebyshev center of

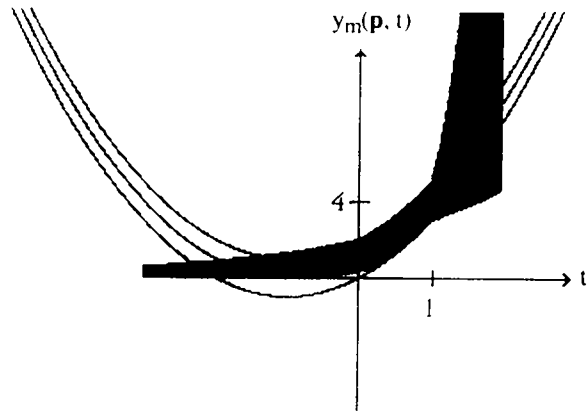


FIG. 7. Superposition of all feasible exponentials in Test case 1.

\mathcal{S} , is a classical point estimator in the bounded-error context. In Test case 1, this estimator gives an estimate which is not feasible as evidenced by Fig. 6. Figure 7 presents the superposition of all exponentials that correspond to a parameter vector belonging to \mathcal{S} .

Test case 2. We choose scales for the parameters such that the prior feasible box \mathbb{P} becomes a cube with side 1. This just changes the bisection policy used during Step 3. Table 2 indicates the performances of SIVIA for various required accuracies ε_r . An important information provided by SIVIA, which could not be presented in this paper for obvious reasons, is the detailed description of all the boxes of the subpavings \mathcal{K}_i and \mathcal{K}_{in} . This information is a much more detailed description of \mathcal{S} than $[\mathcal{S}]$.

Using a signomial approach, Milanese and Vicino (1991b) find a very good estimate of $[\mathcal{S}] = [17.2, 26.9] \times [0.3, 0.49] \times [-16.1, -5.4] \times [0.077, 0.136]$ in about 10 minutes on a VAX 8800 computer. The volume of the set of uncertainty about the location of the parameters drops from $\text{vol}(\mathbb{P}) = 841$ to $\text{vol}([\mathcal{S}]) = 1.16$.

Depending on the quality criterion considered, one or the other approach turns out to give

TABLE 2. PERFORMANCES OF SIVIA ON TEST CASE 2 FOR VARIOUS REQUIRED ACCURACIES ε_r . THE TIMES INDICATED ARE FOR AN IBM-COMPATIBLE COMPAQ 386/33 PERSONAL COMPUTER

ε_r	Time	Iterations	$\text{vol}(\mathcal{K}_{out})$	$\text{vol}(\mathcal{K}_{in})$	#Stack	# \mathcal{K}_i	# \mathcal{K}_{in}
1	0.27 sec	1	841	0	0	1	0
2^{-1}	0.49 sec	31	368	0	4	7	0
2^{-2}	1 sec	167	135	0	6	41	0
2^{-3}	6 sec	765	24	0	9	116	0
2^{-4}	18 sec	2479	3.9	0	11	304	0
2^{-5}	44 sec	6217	0.5	0	14	3×10^3	0
2^{-6}	2 mn	14×10^3	8×10^{-2}	0	16	7×10^3	0
2^{-7}	5 mn	37×10^3	16×10^{-3}	0	19	5×10^3	0
2^{-8}	15 mn	12×10^4	52×10^{-4}	0	23	3×10^4	0
2^{-9}	1 h	66×10^4	26×10^{-4}	26×10^{-5}	27	19×10^4	13×10^3
2^{-10}	10 h	46×10^5	17×10^{-4}	6×10^{-4}	30	15×10^5	22×10^4

TABLE 3. COMPARISON OF THE DESCRIPTIONS $[S]$ AND K_{out} OF S

$[S]$	K_{out}
(1) Accurate individual parameter uncertainty intervals.	(1) Pessimistic individual parameter uncertainty intervals.
(2) Very pessimistic approximation of S .	(2) More detailed description of S .
(3) The models belonging to $M([S])$ may have completely different behaviors.	(3) The models belonging to $M(K_{out})$ have similar behaviors.
(4) Result expressed very concisely.	(4) Result suitable for exploitation on a computer.

better results. If one is interested in computing precise individual parameter uncertainty intervals, then the signomial approach is more efficient. In terms of the volume of the domain guaranteed to contain the parameters, the interval analysis approach gives a better result in 44 sec than the signomial approach in 10 minutes on computers with similar power. In two minutes, SIVIA reduces the volume of the domain guaranteed to contain the parameters by a factor of 10^4 . The resulting $\text{vol}(K_{out})$ is 10 times smaller than $\text{vol}([S])$.

Remark 4. In many practical problems, the set of admissible errors \mathbb{E} is only known approximately. It is therefore important to address the problem of the sensitivity of the estimates obtained to a variation ΔY of Y . From Lemma 1, SIVIA will generate a set K_{out} such that the proximity (25) of $y_m(K_{out})$ to both Y and $Y + \Delta Y$ is small. This means that any model $M(p)$ with p belonging to K_{out} has a behavior consistent (or almost consistent) with the data. On the other hand, p may be very far from S and nevertheless such that the error $e_m(p)$ is close enough to \mathbb{E} for p to belong to K_{out} . Note that there may be some p in $[S]$ that are such that $y_m(p)$ is completely inconsistent with the data.

Table 3 summarizes the properties of two descriptions of S , namely $[S]$ as provided, e.g. by the signomial approach and K_{out} as provided by SIVIA.

7. CONCLUSIONS

Estimating the parameters of a model in the context of bounded errors can be formulated as a problem of set inversion. If this problem can be considered as solved when the model output depends linearly on the parameters to be estimated, the situation is far less advanced in the general nonlinear case considered in this paper. The tools provided by interval analysis appear as very promising, because they make it possible to obtain guaranteed global results, contrary to most methods available so far.

The set inverter via interval analysis proposed here is capable of very quickly eliminating large portions of the parameter space before con-

centrating on the indeterminate region. Theoretical results have been given on its complexity—in terms of memory and computing time—and on its convergence. The required memory remains extremely limited, even when the number of parameters becomes quite large. As could be expected, the number of boxes (which is proportional to the computing time) increases quickly when the number of parameters increases or when more accuracy is required. SIVIA therefore cannot be used with a high accuracy when the number of parameters is too large. On the other hand, used with a low accuracy, it may very quickly eliminate a large portion of the space to be explored even with a rather large number of parameters. That may be very interesting as an initial procedure before using more local approaches.

To the best of our knowledge, the only other method capable of guaranteeing global results that has been considered in the context of nonlinear estimation from bounded-error data is the signomial approach advocated by Milanese and Vicino. Signomial analysis, when applicable, seems to provide accurate descriptions of the smallest axis-aligned box enclosing the posterior feasible set for the parameters S more quickly than SIVIA. On the other hand SIVIA applies to a larger class of problems of set inversion, and characterizes S in a much more detailed way.

Acknowledgements—The authors wish to thank Professor Kurzhanski and Emmanuel Delaleau for helpful comments during the preparation of this paper.

REFERENCES

- Bányász, Cs. and L. Keviczky (Eds) (1991). *Prep. 9th IFAC/IFORS Symposium on Identification and System Parameter Estimation, 1,2*, IFAC, Budapest.
- Belforte, G., B. Bona and V. Cerone (1990). Parameter estimation algorithms for set membership description of uncertainty. *Automatica*, **26**, 887–898.
- Berger, M. (1979). *Espace Euclidien, Triangles, Cercles et sphères; Géométrie 2*, Cedic/Fernand Nathan, Paris, pp. 127–129.
- Broman, V. and M. J. Shensa (1990). A compact algorithm for the intersection and approximation of n -dimensional polytopes. *Mathematics and Computers in Simulation*, **32**, 469–480.
- Clément, T. and S. Gentil (1990). Recursive membership set estimation for output-error models. *Mathematics and Computers in Simulation*, **32**, 505–513.
- Dixon, L. C. W. and G. P. Szego (Eds) (1975). *Towards Global Optimization*. North-Holland, Amsterdam.

- Dixon, L. C. W. and G. P. Szego (Eds) (1978). *Towards Global Optimization 2*. North-Holland, Amsterdam.
- Fogel, E. and Y. F. Huang (1982). On the value of information in system identification—bounded noise case. *Automatica*, **18**, 229–238.
- Horst, R. and H. Tuy. (1990). *Global Optimization, Deterministic Approaches*. Springer-Verlag, Berlin.
- IBM (1986). High-accuracy arithmetic subroutine library (ACRITH). Program description and user's guide, SC 33-6164-02, 3rd ed.
- Jaulin, L. and E. Walter (1993). Guaranteed nonlinear parameter estimation from bounded-error data via interval analysis. *Mathematics & Computers in Simulation* **35**, 123–127.
- Kulisch, U. (Ed.) (1987). *PASCAL-SC: A PASCAL Extension for Scientific Computation, Information Manual and Floppy Disks*. Teubner, Stuttgart.
- Kurzhanski, A. B. and I. Vályi (1991). Guaranteed state estimation for dynamic systems: beyond the overviews. *Prep. 9th IFAC/IFORS Symposium on Identification and System Parameter Estimation*, Budapest, pp. 1033–1037.
- Milanese, M. and G. Belforte (1982). Estimation theory and uncertainty intervals evaluation in presence of unknown-but-bounded errors. Linear families of models and estimators. *IEEE Trans. Aut. Control*, **AC-27**, 408–414.
- Milanese, M. and A. Vicino (1991a). Estimation theory for dynamic systems with unknown but bounded uncertainty: an overview. *Prep. 9th IFAC/IFORS Symposium on Identification and System Parameter Estimation*, Budapest, pp. 859–867.
- Milanese, M. and A. Vicino (1991b). Estimation theory for nonlinear models and set membership uncertainty. *Automatica*, **27**, 403–408.
- Mo, S. H. and J. P. Norton (1990). Fast and robust algorithm to compute exact polytope parameter bounds. *Mathematics and Computers in Simulation*, **32**, 481–493.
- Mockus, J. (1989). *Bayesian Approach to Global Optimization*. Kluwer, Dordrecht.
- Moore, R. E. (1979). *Methods and Applications of Interval Analysis*. SIAM, PA.
- Moore, R. E. (1992). Parameter sets for bounded-error data. *Mathematics and Computers in Simulation*, **34**, 113–119.
- Neumaier, A. (1990). *Interval Methods for Systems of Equations*. Cambridge University Press, Cambridge.
- Norton, J. P. (1987a). Identification and application of bounded-parameter models. *Automatica*, **23**, 497–507.
- Norton, J. P. (1987b). Identification of parameter bounds for ARMAX models from records with bounded noise. *Int. J. Control*, **45**, 375–390.
- Pronzato, L., E. Walter and H. Piet-Lahanier (1989). Mathematical equivalence of two ellipsoidal algorithms for bounded-error estimation. *Proc. 28th IEEE Conf. on Decision and Control*, Tampa, pp. 1952–1955.
- Ratschek, A. and J. Rokne (1988). *New Computer Methods for Global Optimization*. John Wiley, New York.
- Walter, E. and H. Piet-Lahanier (1989). Exact recursive polyhedral description of the feasible parameter set for bounded error models. *IEEE Trans. Aut. Control*, **AC-34**, 911–915.
- Walter, E. and H. Piet-Lahanier (1990). Estimation of parameters bounds from bounded-error data: a survey. *Mathematics and Computers in Simulation*, **32**, 449–468.
- Walter, E. (Ed.) (1990). Special issue on parameter identifications with error bound. *Mathematics and Computers in Simulation*, **32**, 447–607.

Zhigljavsky, A. A. (1991). *Theory of Global Random Search*. Kluwer, Dordrecht.

APPENDIX

Notation

Brackets [] are set apart for interval analysis and therefore never used as substitutes for parentheses. The symbols – and +, used as exponents for a quantity, respectively mean the lowest and the largest possible value for this quantity.

Vectors \mathbf{v} and vector functions \mathbf{f} are printed in bold. Vector equations and inequalities are to be understood componentwise. Usual scalar functions such as exp, sin, . . . , when applied to vectors are also to be understood componentwise. For instance, if $\mathbf{u} = (0, \pi/2, \pi)^T$ and \mathbf{v} is the three-dimensional vector satisfying $v_i = \sin(u_i)$ $1 \leq i \leq 3$, then $\mathbf{v} = \sin(\mathbf{u}) = (0, 1, 0)^T$.

- $\text{clo}(\mathbf{A})$: closure of \mathbf{A} .
- $C(\mathbb{R}^n)$: set of all compacts of \mathbb{R}^n .
- $C_f(\mathbb{R}^n)$: set of all full compacts of \mathbb{R}^n (Definition 13).
- $\partial\mathbf{A}$: boundary of \mathbf{A} .
- $\mathbf{e}_m(\mathbf{p})$: error between the model output and the data (1).
- \mathbf{E} : set of admissible errors.
- ε_r : required accuracy for the paving to be obtained.
- $h_\infty(\mathbf{A}, \mathbf{B})$: Hausdorff distance (Definition 10).
- h_∞^- : complementary Hausdorff distance (Definition 11).
- $h_\infty^0(\mathbf{A}, \mathbf{B})$: proximity of \mathbf{B} to \mathbf{A} (25).
- $\text{int}(a)$: integer part of the real a .
- $\text{int}(\mathbf{A})$: interior of the set \mathbf{A} .
- \mathbb{IR}^n : set of the boxes of \mathbb{R}^n .
- \mathbf{K}_i : indeterminate subpaving.
- $\mathbf{K}_{in}, \mathbf{K}_{out}$: subpavings enclosing the set to be characterized (36).
- $L_\infty(\mathbf{a}, \mathbf{b})$: distance between \mathbf{a} and \mathbf{b} , induced by the L_∞ -norm.
- $L_\infty(\mathbf{A}, \mathbf{B})$: separation between \mathbf{A} and \mathbf{B} (24).
- $m_\infty(\mathbf{A}, \mathbf{B})$: generalized Hausdorff distance (Definition 12).
- n_p : dimension of \mathbf{p} .
- n_y : dimension of \mathbf{y} , i.e. number of measurements.
- \mathbf{p} : vector of the parameters of the model to be estimated from the data.
- \mathbf{P} : prior feasible set for the parameters.
- \mathbf{S} : posterior feasible set for the parameters (2), (4).
- \mathbf{U} : unit sphere in (\mathbb{R}^n, L_∞) .
- $\text{vol}(\mathbf{A})$: volume of the compact set \mathbf{A} .
- $w([\mathbf{p}])$: width of a box (Definition 3).
- \mathbf{y} : vector of all data available on the system.
- $\mathbf{y}_m(\mathbf{p})$: vector of all model outputs.
- \mathbf{Y} : measurement set; $\mathbf{Y} = \mathbf{y} - \mathbf{E}$.
- $[\mathbf{x}]$: box with the same dimension as \mathbf{x} (Definitions 1 and 2).
- $[\mathbf{A}]$: enveloping box of \mathbf{A} (Definition 4).
- $[\mathbf{f}]$: minimal inclusion function of \mathbf{f} (Definition 5).
- \mathbf{f} : inclusion function of \mathbf{f} (Definition 6).
- \mathbf{f}^{-1} : reciprocal function of \mathbf{f} , defined by $\mathbf{f}^{-1}(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{f}(\mathbf{x}) \in \mathbf{A}\}$.
- \times : Cartesian product of sets.
- $\#\mathbf{A}$: cardinal of a finite set \mathbf{A} .
- $\bar{\mathbf{A}}$: complement of \mathbf{A} .
- \circ : composition operator.