



# Computing minimal-volume credible sets using interval analysis; Application to Bayesian estimation

Luc Jaulin

ENSIETA,

2 rue François Verny

29806 Brest Cédex 09

E-mail : luc.jaulin@ensieta.fr

**Abstract:** Given a random vector  $\mathbf{p}$  with a probability density function  $\pi(\mathbf{p})$ , a credible set  $\pi(\mathbf{p})$  with level  $\alpha$  is a set which contains  $\mathbf{p}$  with a probability  $\alpha$ . The problem of characterizing minimal-volume credible sets (which correspond to level sets for  $\pi$ ) is considered. It is only assumed that the expression of  $\pi(\mathbf{p})$  results from a combination of elementary operators ( $+$ ,  $-$ ,  $*$ ,  $/$ ) and elementary functions ( $\sin$ ,  $\cos$ ,  $\text{abs}$ ,  $\dots$ ). This paper provides an algorithm able to compute accurate inner and outer approximations of minimal-volume credible sets, in a guaranteed way. The approach is based on interval analysis and an application to nonlinear parameter estimation, in a Bayesian context, is treated. A windows solver associated with the presented algorithm is made available at <http://www.ensieta.fr/e3i2/Jaulin/bayes.zip>.

## 1. Introduction

Parameter set estimation deals with characterizing a (preferably small) set which encloses the parameter vector  $\mathbf{p}$  of a parametric model from data collected on the system. In the context of bounded-error estimation, the measurement error is assumed to be bounded and the feasible set for the parameter vector  $\mathbf{p}$  can be described as a set inversion problem for which interval methods have been shown to be particularly efficient, even when the model is nonlinear [7]. In a probabilistic context, the error is not anymore described by membership intervals, but by *probability density functions* (pdf). The Bayes rule then makes it possible to obtain the posterior<sup>1</sup> pdf for  $\mathbf{p}$  (see, *e.g.*, [5]). The set to be estimated becomes the *minimal volume credible set* [3] and corresponds to the minimal volume set  $\mathbb{S}_\alpha$ , in the parameter space, which contains  $\mathbf{p}$  with a given probability  $\alpha$ . This set cannot be cast into a set inversion problem and existing interval methods cannot be used without a serious adaptation. To my knowledge, the problem of estimating minimal-volume credible sets in a reliable way has only been solved for some academic examples (see *e.g.* [1]). The goal of this paper is to show how interval methods can be used to solve this problem. Note that the idea of using intervals analysis in a probabilistic context has already been considered by different authors [11, 8]. See also [4] and [9] from which some ideas concerning the representation of pdf have been taken.

---

<sup>1</sup>The prior (resp. posterior) pdf is the pdf for  $\mathbf{p}$  before (resp. after) taking the observations into account.

Section 2 describes in details the problem of characterizing minimal-volume credible sets of a pdf. Some notions on intervals, tilings and staircase functions, needed to understand the methodology, are given in Section 3. These notions will make it possible to solve our problem in Section 4. An application to Bayesian estimation is then considered in Section 5.

## 2. Definition of the problem

For parametric estimation problem, in a Bayesian context, we are generally able to write the posterior pdf for the parameter vector  $\mathbf{p} \in \mathbb{R}^n$  to be estimated into the form  $\pi(\mathbf{p}) = \frac{1}{a}f(\mathbf{p})$ , where  $a$  is the normalization factor and  $f$  is an integrable positive function for which an analytical expression is available (see Section 5 for an illustrative example). When  $a \neq 1$  the function  $f$  is said to be an *unnormalized pdf*. In general  $a$  is not known but, since  $a = \int_{\mathbb{R}^n} f(\mathbf{p})d\mathbf{p}$ , it can be computed approximately. A *credible set* [3] with level  $\alpha$  (or  $\alpha$ -credible set) associated with  $\pi(\mathbf{p})$  is a set  $\mathbb{S}_\alpha$  with probability content  $\int_{\mathbb{S}_\alpha} \pi(\mathbf{p})d\mathbf{p}$  equal to  $\alpha$ . Except for atypical cases that might appear when  $\pi$  has some flat levels, a unique minimal-volume  $\alpha$ -credible set  $\mathbb{S}_\alpha$  exists. This set can be expressed into a set inversion form as  $\mathbb{S}_\alpha = f^{-1}([s_\alpha, +\infty[)$ , where  $s_\alpha \in \mathbb{R}$  is the unknown threshold which should satisfy  $\frac{1}{a} \int_{\mathbb{S}_\alpha} f(\mathbf{p})d\mathbf{p} = \alpha$ . The problem to be considered in this paper is the characterization of  $\mathbb{S}_\alpha$  for any given level  $\alpha \in [0, 1]$ .

**Problem:** Consider a function  $f(\mathbf{p})$  positive for all  $\mathbf{p}$ , such as  $\int_{\mathbb{R}^n} f(\mathbf{p})d\mathbf{p}$  is finite and a real number  $\alpha \in [0, 1]$ . Characterize the set  $\mathbb{S}_\alpha$  defined by

$$\begin{aligned} \text{(i)} \quad & \mathbb{S}_\alpha = f^{-1}([s_\alpha, +\infty[), \\ \text{(ii)} \quad & \int_{\mathbb{S}_\alpha} f(\mathbf{p})d\mathbf{p} = \alpha \int_{\mathbb{R}^n} f(\mathbf{p})d\mathbf{p}. \end{aligned} \tag{2.1}$$

**Example 2.1.** Consider a random variable  $p$ , described by a normal pdf  $\pi(p) = \frac{1}{a}f(p)$  where  $f(p) = \exp\left(-\frac{p^2}{2}\right)$ . The left part of Figure 2.1 represents  $\mathbb{S}_{0.95}$ , its minimal-volume credible set with level 0.95. The grey areas contains 95% of all the area below  $f(p)$ . One cannot find any subset  $\mathbb{S}'$  shorter than  $\mathbb{S}_{0.95}$  such that  $\int_{\mathbb{S}'} f(p)dp \geq 0.95$ . The set  $\mathbb{S}'_{0.95}$  represented on the right part of the figure also contains  $p$  with a probability 0.95 but its length is larger than that of  $\mathbb{S}_{0.95}$ .

**Example 2.2.** Consider a random vector  $\mathbf{p} \in \mathbb{R}^2$  such that  $\pi(\mathbf{p}) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{p_1^2+p_2^2}{2\sigma^2}\right)$ , where  $\sigma = 1$ . The three sets (the disk, the square and the ellipse), represented on Figure 2.2, are credible sets with level 0.9. All sets which contain  $\mathbf{p}$  with a probability equal to 0.9 have an area larger than 14.3, the area of the disk. The disk is thus the minimal-volume 0.9-credible set for  $\mathbf{p}$ .

Notice that the fact that minimum-volume credible sets are level sets for  $f$  is a general property, and is not related to the symmetry of  $f$ , present in the two examples above.

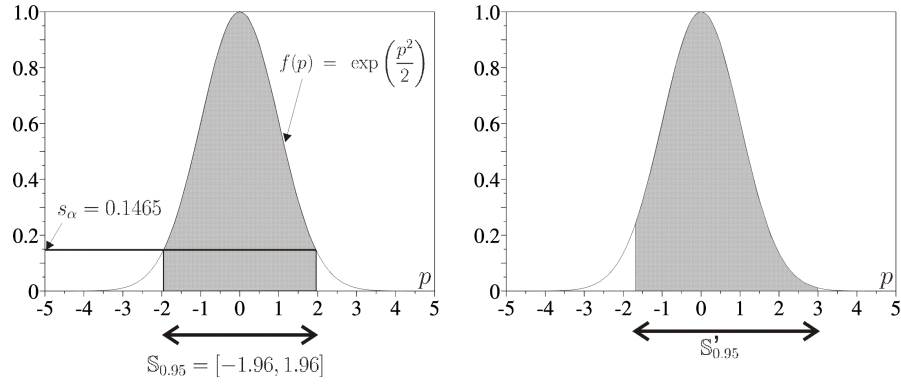


Figure 2.1: Left: minimal-volume credible set  $\mathbb{S}_\alpha$  of level  $\alpha = 0.95$  associated with the standard normal pdf; Right: the set  $\mathbb{S}'_\alpha$  is also a credible region but with a larger volume (or length) than that of  $\mathbb{S}_\alpha$ .

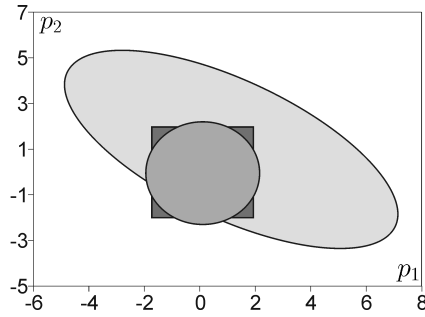


Figure 2.2: The disk is minimal-volume 0.9-credible set for  $\mathbf{p}$

### 3. Intervals

This section recalls some basic notions related to intervals. These notions are needed to understand the algorithm of Section 4 which computes minimal-volume credible sets.

#### 3.1. Intervals of $\mathbb{R}$ and boxes

A *lattice*  $(\mathcal{E}, \leq)$  is a partially ordered set, closed under least upper and greatest lower bounds. The least upper bound (*join*) of  $x$  and  $y$  is written  $x \vee y$ . The greatest lower bound (*meet*) is written  $x \wedge y$ . A lattice  $\mathcal{E}$  is *complete* if for all (finite or not) subsets  $\mathcal{A}$  of  $\mathcal{E}$ , both join and meet of  $\mathcal{A}$  belong to  $\mathcal{E}$ . Define the unary operator  $[\cdot]$  on subsets  $\mathcal{A}$  of a complete lattice  $\mathcal{E}$  as

$$[\mathcal{A}] \stackrel{\text{def}}{=} \{x \in \mathcal{E} \mid \vee \mathcal{A} \leq x \leq \wedge \mathcal{A}\}. \quad (3.1)$$

The operator  $[\cdot]$  is idempotent (i.e.,  $[[\mathcal{A}]] = [\mathcal{A}]$ ). An *interval* of a complete lattice  $\mathcal{E}$  is a subset  $X$  of  $\mathcal{E}$  which satisfies  $[X] = X$ . Note that both  $\emptyset$  and  $\mathcal{E}$  are intervals of  $\mathcal{E}$ .

**Example 3.1.** The set  $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$  is a complete lattice. The intervals of  $\overline{\mathbb{R}}$  are the closed connected subsets of  $\overline{\mathbb{R}}$ . The set of all intervals of  $\overline{\mathbb{R}}$  will be denoted by  $\overline{\mathbb{I}\mathbb{R}}$ .

**Example 3.2.** The set  $\overline{\mathbb{R}}^n$  is a complete lattice with respect to the partial order relation given by  $\mathbf{x} \leq \mathbf{y} \Leftrightarrow \forall i \in \{1, \dots, n\}, x_i \leq y_i$ . We have  $\mathbf{x} \wedge \mathbf{y} = (\min(x_1, y_1), \dots, \min(x_n, y_n))$  and  $\mathbf{x} \vee \mathbf{y} = (\max(x_1, y_1), \dots, \max(x_n, y_n))$ . The intervals of  $\overline{\mathbb{R}}^n$  are called boxes. The set of all boxes of  $\overline{\mathbb{R}}^n$  will be denoted by  $\mathbb{I}\overline{\mathbb{R}}^n$ .

### 3.2. Interval subtilings

A tiling  $\mathcal{Q}$  of  $\mathbb{R}^n$  is a set of non overlapping boxes covering  $\mathbb{R}^n$ . A *subtiling* of  $\mathcal{Q}$  is a subset of  $\mathcal{Q}$ . The set of all subtilings of  $\mathcal{Q}$  will be denoted by  $\mathcal{P}(\mathcal{Q})$ . One can show that  $(\mathcal{P}(\mathcal{Q}), \subset)$  is a complete lattice where the join is the union ( $\mathcal{K}_1 \vee \mathcal{K}_2 = \mathcal{K}_1 \cup \mathcal{K}_2$ ) and the meet is the intersection ( $\mathcal{K}_1 \wedge \mathcal{K}_2 = \mathcal{K}_1 \cap \mathcal{K}_2$ ). As a consequence intervals of  $(\mathcal{P}(\mathcal{Q}), \subset)$  can be defined. An interval subtiling of  $\mathcal{Q}$  can be represented by a pair of two subtilings  $[\mathcal{K}^-, \mathcal{K}^+]$  of  $\mathcal{Q}$  such that  $\mathcal{K}^- \subset \mathcal{K}^+$ , as illustrated by Figure 3.1.

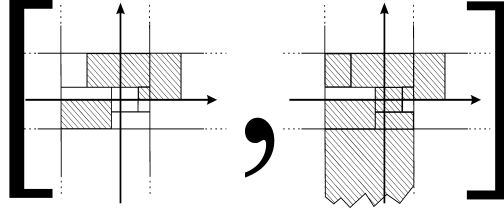


Figure 3.1: An interval subtiling  $[\mathcal{K}^-, \mathcal{K}^+]$

**Example 3.3.** Consider the tiling of  $\mathbb{R}$  defined by  $\mathcal{Q} = \{ ]-\infty, 0], [0, 1], [1, 2], [2, 3], [3, \infty[ \}$ . Three possible interval subtilings of  $\mathcal{Q}$  are

$$[\mathcal{K}_1^-; \mathcal{K}_1^+] = \{ ]-\infty, 0], [1, 2] \}; \{ ]-\infty, 0], [1, 2], [2, 3] \}, \quad (3.2)$$

$$[\mathcal{K}_2^-; \mathcal{K}_2^+] = \{ \{ [1, 2] \} \}; \{ \{ [1, 2] \} \}, \quad (3.3)$$

$$[\mathcal{K}_3^-; \mathcal{K}_3^+] = \{ \{ \}; \{ \} \}. \quad (3.4)$$

The *support*  $\{\mathcal{K}\} \subset \mathbb{R}^n$  of a subtiling  $\mathcal{K} \in \mathcal{P}(\mathcal{Q})$  is the union of all boxes of  $\mathcal{K}$  (see Figure 3.2). A

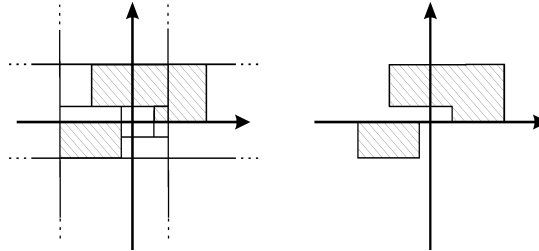


Figure 3.2: A subtiling  $\mathcal{K}$  (left) and its support  $\{\mathcal{K}\}$  (right)

subtiling  $\mathcal{K}$  is said to be *equivalent* to a subset  $\mathbb{S}$  of  $\mathbb{R}^n$  (we shall write  $\mathcal{K} \equiv \mathbb{S}$ ) if  $\{\mathcal{K}\} = \mathbb{S}$ . A set  $\mathbb{S}$  is

said to *belong to* the interval subtiling  $[\mathcal{K}^-, \mathcal{K}^+]$  if all boxes of  $\mathcal{K}^-$  are included in  $\mathbb{S}$  and if all elements of  $\mathbb{S}$  belong to at least one box of  $\mathcal{K}^+$ , *i.e.*,

$$\mathbb{S} \in [\mathcal{K}^-, \mathcal{K}^+] \Leftrightarrow \{\mathcal{K}^-\} \subset \mathbb{S} \subset \{\mathcal{K}^+\}. \quad (3.5)$$

### 3.3. Interval staircase functions

A *staircase function* associated with a tiling  $\mathcal{Q}$  is a function from  $\mathcal{Q}$  to  $\overline{\mathbb{R}}$ . The set of all staircase functions equipped with the natural partial order ( $\phi_1 \leq \phi_2 \Leftrightarrow \forall \mathbf{P} \in \mathcal{Q}, \phi_1(\mathbf{P}) \leq \phi_2(\mathbf{P})$ ) is a complete lattice. Interval staircase functions can thus be defined. An interval staircase function  $\Phi = [\phi^-, \phi^+]$  is a pair of two staircase functions such that  $\phi^- \leq \phi^+$ . A function  $f$  from  $\mathbb{R}^n \rightarrow \mathbb{R}$  is said to *belong to* the interval staircase function  $\Phi$  (we shall write  $f \in \Phi$ ) if

$$\forall \mathbf{P} \in \mathcal{Q}, \forall \mathbf{p} \in \mathbf{P}, f(\mathbf{p}) \in [\phi^-(\mathbf{P}), \phi^+(\mathbf{P})]. \quad (3.6)$$

An interval staircase function for  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  can easily be obtained by using interval techniques as described in [10]. The *reciprocal image* of an interval  $S \in \mathbb{I}\mathbb{R}$  by the interval staircase function  $\Phi = [\phi^-, \phi^+]$  is the interval subtiling of  $\mathcal{Q}$  defined by

$$\Phi^{-1}(S) \stackrel{\text{def}}{=} [\{\mathbf{P} \in \mathcal{Q} \mid \Phi(\mathbf{P}) \subset S\}, \{\mathbf{P} \in \mathcal{Q} \mid \Phi(\mathbf{P}) \cap S \neq \emptyset\}] \quad (3.7)$$

**Theorem 3.4.** *If  $f \in \Phi$ , then for all intervals  $S \in \mathbb{I}\mathbb{R}$ , we have  $f^{-1}(S) \in \Phi^{-1}(S)$ .*

**Proof:** We have

$$\bigcup \{\mathbf{P} \in \mathcal{Q} \mid \Phi(\mathbf{P}) \cap S \neq \emptyset\} \supset \{\mathbf{p} \in \mathbb{R}^n \mid f(\mathbf{p}) \in S\} \quad (3.8)$$

and

$$\bigcup \{\mathbf{P} \in \mathcal{Q} \mid \Phi(\mathbf{P}) \subset S\} \subset \{\mathbf{p} \in \mathbb{R}^n \mid f(\mathbf{p}) \in S\}. \quad (3.9)$$

From Equations (3.5) and (3.7), we get that  $f^{-1}(S) \in \Phi^{-1}(S)$ .

**Example 3.5.** *Figure 3.3 illustrates Theorem 3.4 for  $S = [16, \infty]$  and  $\mathcal{Q} = \{[i, i + 1], i \in \mathbb{N}\}$ . We have*

$$\begin{aligned} \{\mathbf{P} \in \mathcal{Q} \mid \Phi(\mathbf{P}) \subset S\} &= \{[-1, 0], [0, 1]\} \equiv [-1, 1], \\ \{\mathbf{P} \in \mathcal{Q} \mid \Phi(\mathbf{P}) \cap S \neq \emptyset\} &= \{[-3, -2], [-2, -1], [-1, 0], [0, 1], [1, 2], [2, 3]\} \equiv [-3, 3]. \end{aligned} \quad (3.10)$$

*Thus  $[-1, 1] \subset f^{-1}([16, \infty]) \subset [-3, 3]$ .*

The *integral* of the interval staircase function  $\Phi = [\phi^-, \phi^+]$  over the interval subtiling  $[\mathcal{K}^-, \mathcal{K}^+]$  is defined by

$$\int_{[\mathcal{K}^-, \mathcal{K}^+]} \Phi(\mathbf{p}) d\mathbf{p} = \left[ \sum_{\mathbf{P} \in \mathcal{K}^-} \phi^-(\mathbf{P}) \cdot \text{volume}(\mathbf{P}), \sum_{\mathbf{P} \in \mathcal{K}^+} \phi^+(\mathbf{P}) \cdot \text{volume}(\mathbf{P}) \right]. \quad (3.11)$$

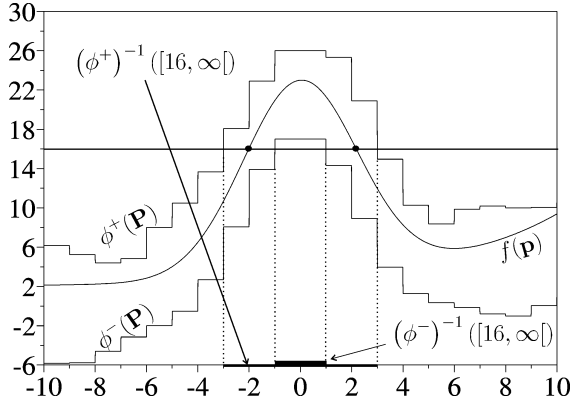


Figure 3.3: An interval staircase function enclosing a function  $f(\mathbf{p})$

**Theorem 3.6.** *If  $f$  is a positive function which belongs to  $\Phi$ , if  $\mathbb{S} \in [\mathcal{K}^-, \mathcal{K}^+]$  is a measurable set then*

$$\int_{\mathbb{S}} f(\mathbf{p}) d\mathbf{p} \in \int_{[\mathcal{K}^-, \mathcal{K}^+]} \Phi(\mathbf{p}) d\mathbf{p}. \quad (3.12)$$

**Proof:** Since  $f$  is positive and  $\mathbb{S} \in [\mathcal{K}^-, \mathcal{K}^+]$ ,  $\int_{\mathbb{S}} f(\mathbf{p}) d\mathbf{p} \in [a] \stackrel{\text{def}}{=} \left[ \int_{\{\mathcal{K}^-\}} f(\mathbf{p}) d\mathbf{p}, \int_{\{\mathcal{K}^+\}} f(\mathbf{p}) d\mathbf{p} \right]$ . Now,

$$[a] = \left[ \sum_{\mathbf{P} \in \mathcal{K}^-} \int_{\mathbf{P}} f(\mathbf{p}) d\mathbf{p}, \sum_{\mathbf{P} \in \mathcal{K}^+} \int_{\mathbf{P}} f(\mathbf{p}) d\mathbf{p} \right] \quad (3.13)$$

is included in  $\left[ \sum_{\mathbf{P} \in \mathcal{K}^-} \int_{\mathbf{P}} \phi^-(\mathbf{p}) d\mathbf{p}, \sum_{\mathbf{P} \in \mathcal{K}^+} \int_{\mathbf{P}} \phi^+(\mathbf{p}) d\mathbf{p} \right]$  which is equal to  $\int_{[\mathcal{K}^-, \mathcal{K}^+]} \Phi(\mathbf{p}) d\mathbf{p}$ .

#### 4. Algorithm

In this section, we shall provide a numerical interval algorithm to characterize the minimal-volume  $\alpha$ -credible set of an unnormalized pdf  $f(\mathbf{p})$ . We need first to solve the following equation (see (2.1)):

$$\alpha = h(s_\alpha) \stackrel{\text{def}}{=} \frac{\int_{f^{-1}([s_\alpha, \infty])} f(\mathbf{p}) d\mathbf{p}}{\int_{\mathbb{R}^n} f(\mathbf{p}) d\mathbf{p}} \quad (4.1)$$

where the single unknown is  $s_\alpha$ . Since

$$s_1 \leq s_2 \Rightarrow [s_2, \infty[ \subset [s_1, \infty[ \Rightarrow f^{-1}([s_2, \infty]) \subset f^{-1}([s_1, \infty]) \quad (4.2)$$

$$\Rightarrow \int_{f^{-1}([s_2, \infty])} f(\mathbf{p}) d\mathbf{p} \leq \int_{f^{-1}([s_1, \infty])} f(\mathbf{p}) d\mathbf{p}, \quad (4.3)$$

the function  $h(s)$  is decreasing. Moreover,

$$h(s) \in H(s) \stackrel{\text{def}}{=} \frac{\int_{\Phi^{-1}([s, \infty])} \Phi(\mathbf{p}) d\mathbf{p}}{\int_{\mathbb{R}^n} \Phi(\mathbf{p}) d\mathbf{p}}. \quad (4.4)$$

Note that  $H(s)$  is a function from  $\mathbb{R} \rightarrow \mathbb{I}\mathbb{R}$ .

**Remark 1.** For a given  $s$ , to compute  $H(s)$ , one should first compute the interval subtiling  $[\mathcal{K}^-, \mathcal{K}^+] = \Phi^{-1}([s, \infty[)$ . Then, we have to compute the two intervals  $\int_{\Phi^{-1}([s, \infty[)} \Phi(\mathbf{p}) d\mathbf{p}$  and  $\int_{\mathbb{R}^n} \Phi(\mathbf{p}) d\mathbf{p}$  (see Equation (3.11)). The two resulting intervals are then divided using interval arithmetic rules. This procedure can be interpreted as an adaptation of the Moore's method for interval numerical integration [10].

Since  $h(s)$  is decreasing and since  $h(s) \in H(s)$ , we have the following implications

$$\begin{aligned} (a) \quad \alpha < \text{lb}(H(s^-)) &\Rightarrow s^- < s_\alpha \\ (b) \quad \alpha > \text{ub}(H(s^+)) &\Rightarrow s^+ > s_\alpha \end{aligned} \tag{4.5}$$

where  $lb$  and  $ub$  correspond to the lower and upper bounds, respectively. These two implications are illustrated by Figure 4.1.

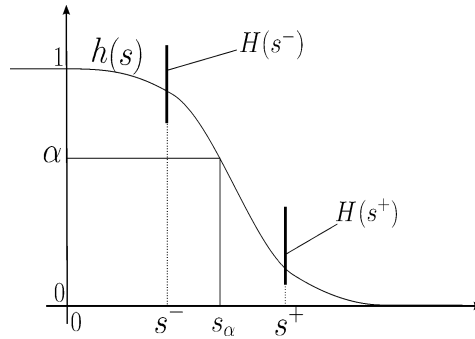


Figure 4.1: Illustration of the two tests used to find an upper and a lower bound for  $s_\alpha$

The principle of the method to compute an enclosure  $[s^-, s^+]$  of  $s_\alpha$  is to decrease the value  $s^-$  initialized at  $+\infty$  until  $\alpha < \text{lb}(H(s^-))$  and to increase the value  $s^+$  initialized at 0 until  $\alpha > \text{ub}(H(s^+))$ .

The procedure for computing  $\mathbb{S}_\alpha$  is as follows:

1. Take a tiling  $\mathcal{Q}$  of  $\mathbb{R}^n$ ;
2. Compute an interval staircase function  $\Phi$  enclosing  $f$ ;
3. Compute the two lists  $\phi^-(\mathcal{Q}) := \{\phi^-(\mathbf{P}), \mathbf{P} \in \mathcal{Q}\}$  and  $\phi^+(\mathcal{Q}) := \{\phi^+(\mathbf{P}), \mathbf{P} \in \mathcal{Q}\}$ ;
4. Compute  $s^-$  the largest element of  $\phi^-(\mathcal{Q})$ ;
5. If  $\alpha \geq \text{lb}(H(s^-))$ , remove  $s^-$  from the list  $\phi^-(\mathcal{Q})$  and go to 4;
6. Denote by  $s^+$  the smallest element of  $\phi^+(\mathcal{Q})$ ;
7. If  $\alpha \leq \text{ub}(H(s^+))$ , remove  $s^+$  from the list  $\phi^+(\mathcal{Q})$  and go to 6;
8.  $[\mathcal{K}_\alpha^-, \mathcal{K}_\alpha^+] := (\Phi - [s^-, s^+])^{-1}([0, \infty[)$ .

**Comments:** At Step 1 a finite tiling  $\mathcal{Q}$  is chosen in order to have small boxes where  $f$  varies significantly and big boxes where the density is almost zero. The interval staircase function  $\Phi$ , computed at Step 2, is obtained using interval arithmetic and is then stored in a list whose elements are pairs  $(\mathbf{P}_i, \Phi_i)$  where  $\mathbf{P}_i$  is the  $i$ th box and  $\Phi_i$  corresponds to  $\Phi(\mathbf{P}_i)$ . At Step 3, we define the lists  $\phi^-(\mathcal{Q})$  and  $\phi^+(\mathcal{Q})$  which correspond to all values of  $s$  where  $\text{lb}(H(s))$  and  $\text{ub}(H(s))$  can change, respectively. Thus, the values for  $s^-$  and  $s^+$  will all be taken in  $\phi^-(\mathcal{Q})$  and  $\phi^+(\mathcal{Q})$ . Steps 4 and 5 compute a lower bound  $s^-$  for  $s_\alpha$ . Steps 6 and 7 compute an upper bound  $s^+$  for  $s_\alpha$ . Recall that  $H(s^-)$  and  $H(s^+)$  are computed using the procedure described on Remark 1. When the algorithm reaches Step 8, we have  $\text{ub}(H(s^+)) < \alpha < \text{lb}(H(s^-))$  and thus  $s_\alpha \in [s^-, s^+]$ . At Step 8, when we compute  $\Phi - [s^-, s^+]$ , we used abusively of an arithmetic over interval staircase functions (such an arithmetic has not been defined in this paper). In fact,  $\Phi - [s^-, s^+]$  denotes the interval staircase function associated with the function  $f - s_\alpha$ . Indeed,  $\mathbb{S}_\alpha = f^{-1}([s_\alpha, \infty[) = (f - s_\alpha)^{-1}([0, \infty[)$  and thus from Theorem 3.4,  $\mathbb{S}_\alpha \in [\mathcal{K}_\alpha^-, \mathcal{K}_\alpha^+]$ .

**Theorem 4.1.** *After completion of this algorithm, we have  $\mathbb{S}_\alpha \in [\mathcal{K}_\alpha^-, \mathcal{K}_\alpha^+]$  and  $s_\alpha \in [s^-, s^+]$ .*

A solver implementing this algorithm can be freely downloaded with its C++ code [6].

## 5. Application to Bayesian estimation

Consider now a parametric model for a system for which  $m$  output data have been collected (see [14] for notions related to parameter estimation). All these data have been stored within a vector  $\mathbf{y}$ . The output vector  $\mathbf{y}$  is assumed to be related to the parameter vector by  $\mathbf{y} = \psi(\mathbf{p}) + \mathbf{n}$ , where  $\mathbf{n}$  is the noise vector, and  $\psi$  is the model function. In a probabilistic approach, we generally assume that prior pdf  $\pi_{\text{prior}}(\mathbf{p})$  for  $\mathbf{p}$  and  $\pi_n(\mathbf{n})$  for  $\mathbf{n}$  are known. The Bayes rule (see *e.g.* [5], [14]) states that the posterior pdf for  $\mathbf{p}$  is given by

$$\pi_{\text{post}}(\mathbf{p}) = \frac{\pi_n(\mathbf{y} - \psi(\mathbf{p})) \cdot \pi_{\text{prior}}(\mathbf{p})}{\int_{\mathbf{p} \in \mathbb{R}^n} \pi_n(\mathbf{y} - \psi(\mathbf{p})) \cdot \pi_{\text{prior}}(\mathbf{p}) d\mathbf{p}}. \quad (5.1)$$

For most applications, we have an analytical expression for  $\pi_n(\mathbf{n})$  and  $\pi_{\text{prior}}(\mathbf{p})$ , but not<sup>2</sup> for the normalization factor  $a$  given by the denominator of (5.1). Therefore, characterizing the minimal-volume credible set of  $\pi_{\text{post}}(\mathbf{p})$  amounts to solving (2.1) where

$$f(\mathbf{p}) = \pi_n(\mathbf{y} - \psi(\mathbf{p})) \cdot \pi_{\text{prior}}(\mathbf{p}). \quad (5.2)$$

As an illustration, consider the model described by  $y(t) = p_1 \sin(p_2 t) + n(t)$  where  $n(t)$  is a white random signal. For each  $t$ , its pdf is normal, *i.e.*, it satisfies

$$\pi_n(n) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{n^2}{2\sigma^2}\right) \quad (5.3)$$

---

<sup>2</sup>except for some special cases such as when  $\phi$  is linear and when  $\pi_n(\mathbf{n})$  and  $\pi_{\text{prior}}(\mathbf{p})$  are Gaussian



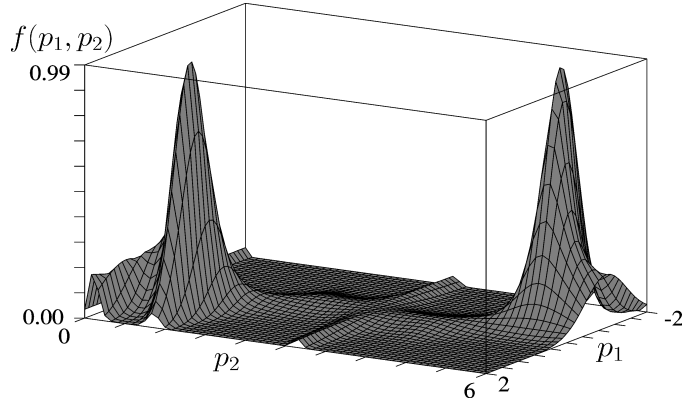


Figure 5.1: Posterior unnormalized pdf for  $\mathbf{p}$

where the standard deviation  $\sigma = \frac{1}{2}$ . Assume that at sampling times  $\mathbf{t} = (1, 2, 3)$ , the three following measurements have been collected  $\mathbf{y} = (0.8, 1.0, 0.2)^\top$ . Therefore

$$\underbrace{\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}}_{\mathbf{y}} = \underbrace{\begin{pmatrix} p_1 \sin(p_2) \\ p_1 \sin(2p_2) \\ p_1 \sin(3p_2) \end{pmatrix}}_{\psi(\mathbf{p})} + \underbrace{\begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}}_{\mathbf{n}} \quad (5.4)$$

where  $n_1 = n(1)$ ,  $n_2 = n(2)$  and  $n_3 = n(3)$ . Since  $n(t)$  is assumed to be white, the vector  $\mathbf{n} = (n_1, n_2, n_3)^\top$  is a white random vector. As a consequence,

$$\pi_n(\mathbf{n}) = \prod_{k=1}^3 \pi_n(n_k) = \prod_{k=1}^3 \sqrt{\frac{2}{\pi}} \exp(-2n_k^2). \quad (5.5)$$

Assume that the prior pdf for  $\mathbf{p}$  is  $\pi_{\text{prior}}(\mathbf{p}) = \frac{1}{24} \Pi_{[-2,2]}(p_1) \cdot \Pi_{[0,6]}(p_2)$ , where  $\Pi_{[a,b]}(p)$  takes the value 1 for  $p \in [a, b]$  and zero elsewhere. This choice for  $\pi_{\text{prior}}(\mathbf{p})$  can be interpreted as a prior knowledge that  $\mathbf{p}$  belongs to the box  $[-2, 2] \times [0, 6]$ . From Equation (5.2), we get a posterior unnormalized pdf for  $\mathbf{p}$ :

$$f(\mathbf{p}) = \left( \prod_{k=1}^3 \exp(-2(y_k - p_1 \sin(kp_2))^2) \right) \cdot \Pi_{[-2,2]}(p_1) \cdot \Pi_{[0,6]}(p_2). \quad (5.6)$$

This function is depicted on Figure 5.1. The algorithm presented on Section 4 made it possible to obtain, via the solver BAYES [6], for 6 different values of  $\alpha$ , a characterization of the minimal-volume credible sets represented on Figure 5.2. If  $\mathcal{K}_\alpha^-$  denotes the subtiling containing all dark grey boxes, and if  $\mathcal{K}_\alpha^+$  denotes the subtiling containing both dark grey and white boxes, then, the interval subtiling  $[\mathcal{K}_\alpha^-, \mathcal{K}_\alpha^+]$  is proved to contain  $\mathbb{S}_\alpha$  (*i.e.*,  $\mathbb{S}_\alpha \in [\mathcal{K}_\alpha^-, \mathcal{K}_\alpha^+]$ ). Note that  $\mathcal{K}_0^- = \emptyset$ , whereas  $\mathcal{K}_1^+ = [-2, 2] \times [0, 6]$ . This is consistent with the fact that  $\mathbb{S}_0 = \emptyset$  and  $\mathbb{S}_1 = [-2, 2] \times [0, 6]$ . The computing time to get all pictures of Figure 5.2 is about 100 seconds on a Pentium 3. The interval function  $H(s)$  represented on Figure 5.3 has been computed to get Figure 5.2 and shows the relation between the threshold  $s$  and the corresponding level  $\alpha$ . The poor quality of the enclosure, mainly due to the dependency effect [10] occurring in Equation (5.6), could be increased at the cost of higher computing time.

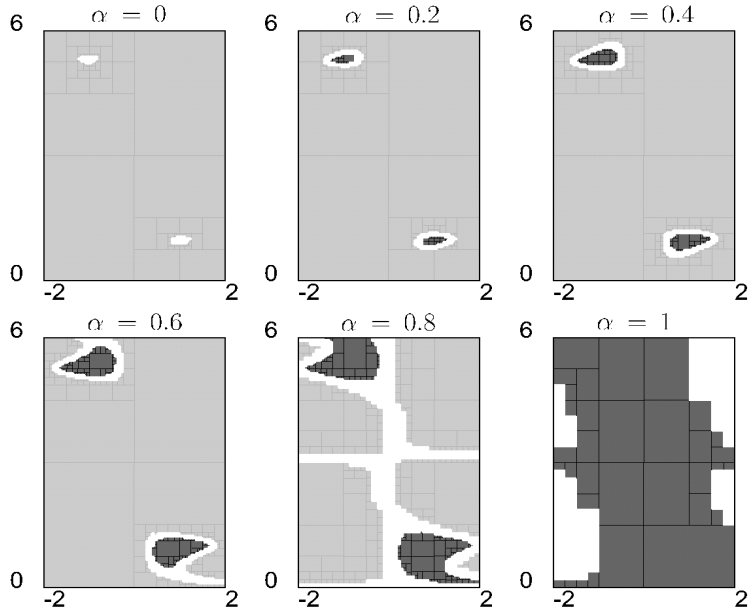


Figure 5.2: Representation of  $\mathbb{S}_\alpha$  for  $\alpha \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$ ; the dark grey zone is proved to be included in  $\mathbb{S}_\alpha$ ; the light grey zone is proved to be outside  $\mathbb{S}_\alpha$

**Remark 2.** *The dependency effect happens when some variables occur more than once in an interval expression. This effect introduces some pessimism during the interval evaluation. Take for instance  $f(x) = (x - 1)^2$  and  $g(x) = x^2 - 2x + 1$ . It is clear that for all  $x \in \mathbb{R}$ ,  $f(x) = g(x)$ . This is not the case when intervals are involved. For  $X = [1, 2]$ , an interval evaluation of  $f$  and  $g$  gives*

$$f(X) = (X - 1)^2 = ([1, 2] - 1)^2 = [0, 1], \quad (5.7)$$

and

$$\begin{aligned} g(X) &= X^2 - 2X + 1 = [1, 2]^2 - 2 \cdot [1, 2] + 1 \\ &= [1, 4] + [-4, -2] + 1 = [-2, 3]. \end{aligned}$$

Both intervals contain the set  $\{f(x), x \in X\} = [0, 1]$ , but the second interval  $[-2, 3]$  is too large. This overestimation is due to the fact that  $x$  occurs twice in the expression of  $g$ .

## 6. Conclusion

In this paper, we have presented a new interval-based algorithm which made it possible to characterize the minimal-volume credible sets of an unnormalized pdf, in a guaranteed way. A reliable visualization of minimum-volume credible sets is especially useful when the number of data is small, so that the posterior density of the model parameters is far from being normal. A free solver associated to the presented approach has also been made available [6]. An illustrative example related to Bayesian estimation has been given. This example has shown the main limitation of our algorithm: even a

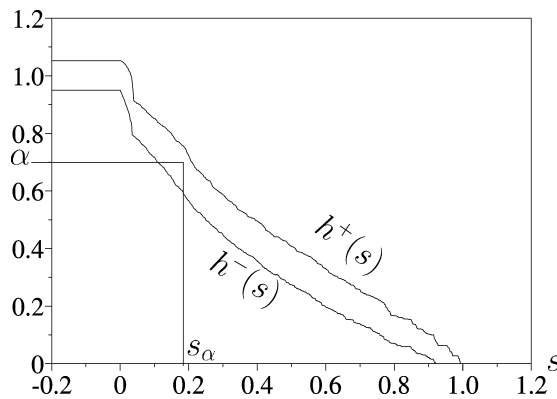


Figure 5.3: Interval function  $H(s)$  representing the relation between the threshold  $s$  and the corresponding level  $\alpha$

few numbers of parameters and for a rough accuracy, the computing time is relatively high (here 100 seconds). For more parameters, the method should be improved in order to increase its efficiency. This could be done using more accurate inclusion functions and by resorting to constraint propagation methods (see, *e.g.*, [2, 13, 12]).

## References

- [1] M. Bayarri and J. Berger. The interplay between bayesian and frequentist analysis. *Statistical Science*, 19:58–80, 2004.
- [2] F. Benhamou and W. Older. Applying interval arithmetic to real, integer and Boolean constraints. *Journal of Logic Programming*, pages 1–24, 1997.
- [3] J. Berger. *Statistical Decision Theory and Bayesian Analysis*, 2nd edition. Springer-Verlag, New York, NY, 1985.
- [4] D. Berleant, L. Xie, and J. Zhang. Statool: a tool for distribution envelope determination (denv), an interval-based algorithm for arithmetic on random variables. *Reliable Computing*, 9(2):91–108, 2003.
- [5] P. Eykhoff. *System Identification, Parameter and State Estimation*. John Wiley, London, 1974.
- [6] L. Jaulin. *BAYES, a Windows solver to characterize minimal-volume credible sets, available at [www.ensieta.fr/jaulin/bayes.zip](http://www.ensieta.fr/jaulin/bayes.zip)*. E3I2, ENSIETA, 2004.
- [7] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter. *Applied Interval Analysis, with Examples in Parameter and State Estimation, Robust Control and Robotics*. Springer-Verlag, London, 2001.

- [8] V. Kreinovich, G. Dimuro, and A. C. da Rocha Costa. Probabilities, intervals, what next ? extension of interval computations to situations with partial information about probabilities. In *10th IMEKO TC7 International symposium*, 2004.
- [9] W. Li and J. Hyman. Computer arithmetic for probability distribution variables. *Reliable Engineering and System Safety*, 85:191–209, 2004.
- [10] R. E. Moore. *Interval Analysis*. Prentice-Hall, Englewood Cliffs, NJ, 1966.
- [11] P. Nivlet. *Prise en compte des incertitudes dans l'interprétation réservoir des données géophysiques ; application à la segmentation*. PhD dissertation, Institut Français du Pétrole, France, 2001.
- [12] M. van Emden. Algorithmic power from declarative use of redundant constraints. *Constraints*, 4(4):363–381, 1999.
- [13] P. van Hentenryck, Y. Deville, and L. Michel. *Numerica: A Modeling Language for Global Optimization*. MIT Press, Boston, MA, 1997.
- [14] E. Walter and L. Pronzato. *Identification of Parametric Models from Experimental Data*. Springer-Verlag, London, UK, 1997.