



HAL
open science

Guaranteed Nonlinear Parameter Estimation with Additive Gaussian Noise

Jérémy Nicola, Luc Jaulin

► **To cite this version:**

Jérémy Nicola, Luc Jaulin. Guaranteed Nonlinear Parameter Estimation with Additive Gaussian Noise. Olga Kosheleva; Sergey P. Shary; Gang Xiang; Roman Zapatrin. Beyond Traditional Probabilistic Data Processing Techniques: Interval, Fuzzy etc. Methods and Their Applications, 835, Springer, pp.341-357, 2020, Studies in Computational Intelligence, 978-3-030-31041-7 (Ebook); 978-3-030-31040-0 (Hardcover). 10.1007/978-3-030-31041-7_19 . hal-02520046

HAL Id: hal-02520046

<https://hal.science/hal-02520046v1>

Submitted on 3 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Guaranteed nonlinear parameter estimation with additive Gaussian noise

J. Nicola

Ixblue, ENSTA-Bretagne, LabSTICC

jeremy.nicola@gmail.com

L. Jaulin

ENSTA-Bretagne, LabSTIC, UBO

lucjaulin@gmail.com

Abstract

In this paper we propose a new approach for nonlinear parameter estimation under additive Gaussian noise. We provide an algorithm based on interval analysis and set inversion which computes an inner and an outer approximation of a set enclosing the parameter vector with a given probability. The principle of the approach is illustrated by examples related to parameter estimation and range-only localization.

Keywords: Interval analysis, set-estimation, probabilistic estimation, parameter estimation, localization

AMS subject classifications: 65-00

1 Introduction

Parameter set estimation deals with characterizing a set (preferably small) which encloses the parameter vector \mathbf{p} of a parametric model from a finite set of data collected on the system. In a bounded-error context [26][22][31] the measurement errors are assumed to be bounded and computing the feasible set for \mathbf{p} can be described as a set inversion problem [14] for which interval methods [24] are particularly efficient, even when the model is nonlinear. In a probabilistic context, the error is not anymore described by membership intervals, but by probability density functions (pdf) instead. The correspondence between the two approaches has been studied by Vladik Kreinovich [17] [20]. In this context, Vladik proposed showed that the interval estimation problem was intractable [19], even in a linear context when experimental factors are uncertain [18]. He also provided some links with a fuzzy representation of uncertainties [21] and how to deal with outliers [29].

In a Bayesian context, the Bayes rule makes it possible to get the posterior pdf for \mathbf{p} (see, e.g., [8]). The set to computed becomes the *credible set* [2] and corresponds to the minimal volume set, in the parameter space, which contains \mathbf{p} with a given

probability η . This problem cannot be cast into a set inversion problem but existing interval methods can still be used [10]. Unfortunately, the approach is limited to few parameters (typically less than 3) and few measurements (typically less than 10).

Recently, an original approach [3] named *Sign-Perturbed Sums* (SPS) has proposed to construct non-asymptotic confidence regions which is guaranteed to contain the true parameters with a given probability η . This approach has been used for nonlinear models to compute confidence regions [5] which have not a minimal volume (at least in the Gaussian case). Interval analysis has also been considered to deal with the SPS method [16] to compute guaranteed confidence regions. Other methods such as [6] or [11] are also able to compute guaranteed confidence regions using interval analysis, but the computed set is not of minimal volume and it is difficult to evaluate the resulting pessimism.

There exist other approaches that combine bounded-error estimation with probabilistic estimation [1] [25] [32], [9] or use other frameworks such as random sets [23] [33] [28] or fuzzy-sets [7][30], but all these methods do not solve a problem which is expressed only in terms of probabilities only and can thus not be used to compute confidence regions.

This paper considers a problem which can be considered as classic on probabilistic parameter estimation: compute a set which encloses the parameter vector with a fixed probability η . Our main contribution is to be able to solve this problem in a reliable way in the case where the error is Gaussian and the model is nonlinear.

Section 2 recalls the principle of set-inversion for the specific case where the noise is Gaussian and proposes different shape for the set to be inverted. Section 3 recalls the principle of the linear Gaussian estimation that will be used for comparison. Section 4 illustrates the proposed approach on three simple simulated examples and gives a comparison with a classical linear Gaussian estimator. Section 5 concludes the paper.

2 Set Inversion for Nonlinear Gaussian Estimation

This section recalls the principle of set inversion and considers the special case where the set to be inverted is a confidence region of a Gaussian probability density function. Consider the following parameter estimation problem

$$\mathbf{y} = \boldsymbol{\psi}(\mathbf{p}) + \mathbf{e}, \quad (1)$$

where $\boldsymbol{\psi}$ is the model, $\mathbf{y} \in \mathbb{R}^n$ is the vector of all measurements (which is known) and \mathbf{e} is the error vector. Without loss of generality we assume that $\mathbf{e} : \mathcal{N}(\mathbf{0}, \mathbf{I}^n)$.

Remark. As illustrated by Figure 1, a random variable \mathbf{y} following a normal distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Gamma})$ can always be whitened into a random variable \mathbf{x} distributed as $\mathcal{N}(\mathbf{0}, \mathbf{I}^n)$ by the affine transform $\mathbf{x} = \boldsymbol{\Gamma}^{-1/2}(\mathbf{y} - \boldsymbol{\mu})$.

Definition. Define the function $\mathbf{f}(\mathbf{p}) = \mathbf{y} - \boldsymbol{\psi}(\mathbf{p})$ corresponding to the error \mathbf{e} and a set \mathbb{E}_η containing \mathbf{e} with a probability η . The *probabilistic set* associated to \mathbb{E}_η is defined as

$$\hat{\mathbb{P}}_{\mathbb{E}_\eta} = \mathbf{f}^{-1}(\mathbb{E}_\eta). \quad (2)$$

It contains \mathbf{p} with a prior probability of η [11]. As a consequence, a probabilistic set estimation can be viewed as a set inversion problem for which guaranteed interval

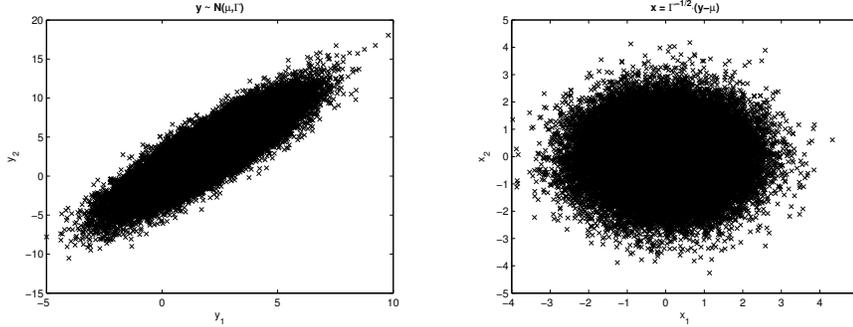


Figure 1: Several realizations of the random vector $\mathbf{y} \sim \mathcal{N}\left(\boldsymbol{\mu} = \begin{pmatrix} 2 & 3 \end{pmatrix}^T, \boldsymbol{\Gamma} = \begin{pmatrix} 3 & 5 \\ 5 & 11 \end{pmatrix}\right)$ and their images $\mathbf{x} = \boldsymbol{\Gamma}^{-1/2} (\mathbf{y} - \boldsymbol{\mu})$

techniques could be used. Now, there exists several manners to choose such a set \mathbb{E}_η . We compare two different types of sets: a sphere (which is a confidence region of minimal volume) and a box, which is a good representation for interval methods.

Let us now recall [27] some results useful to get a set which encloses the normal error \mathbf{e} with a given probability η .

Theorem. The minimal volume confidence region of probability η associated with $\mathbf{e} : \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}^n)$ is the centered n -dimensional sphere \mathbb{S}_η of radius α , where (α, η) are linked by the relation

$$\eta = \int_0^{\alpha^2} \frac{z^{\left(\frac{n}{2}-1\right)} e^{-\frac{z}{2}}}{2^{\frac{n}{2}} \Gamma_e\left(\frac{n}{2}\right)} \cdot dz \tag{3}$$

where Γ_e the Euler function. Recall that for $n \in \mathbb{N}$ the Euler function satisfies

$$\Gamma_e(n) = (n - 1)! \tag{4}$$

Proof. The random variable $z = \mathbf{e}^T \cdot \mathbf{e}$ follows a χ^2 distribution with n degrees of freedom whose probability density function is

$$\pi(z, n) = \frac{z^{\left(\frac{n}{2}-1\right)} \cdot e^{-\frac{z}{2}}}{2^{\frac{n}{2}} \Gamma_e\left(\frac{n}{2}\right)}. \tag{5}$$

The minimal volume confidence region \mathbb{S}_η is the set of all \mathbf{e} such that

$$z = \mathbf{e}^T \cdot \mathbf{e} \leq \alpha^2(\eta) \tag{6}$$

and the probability η to have $\mathbf{e} \in \mathbb{S}_\eta$ is

$$\eta = \int_0^{\alpha^2} \pi(z, n) \cdot dz = \int_0^{\alpha^2} \frac{z^{\left(\frac{n}{2}-1\right)} e^{-\frac{z}{2}}}{2^{\frac{n}{2}} \Gamma_e\left(\frac{n}{2}\right)} \cdot dz. \tag{7}$$

For $n = 1$, $n = 2$ or n large, from the integral in Equation (3), we can have an expression of the radius $\alpha(\eta)$ [2] as recalled in Table 1.

n	$\alpha(\eta)$
1	$\alpha = \sqrt{2} \operatorname{erf}^{-1}(\eta)$
2	$\alpha = \sqrt{-2 \cdot \log(1 - \eta)}$
$n \gg 1$	$\alpha \simeq \sqrt{n + 2\sqrt{n} \cdot \operatorname{erf}^{-1}\left(2\eta + \operatorname{erf}\left(\frac{-\sqrt{n}}{2}\right)\right)}$

Table 1: $\alpha(\eta)$ for $n = 1, 2$ and $n \gg 1$

In our context, the dimension of \mathbf{e} is large and we can consider that the formula corresponding to $n \gg 1$ is correct.

Theorem. With $n \gg 1$, the probability ϕ_η to have $\mathbf{e} : \mathcal{N}(\mathbf{0}, \mathbf{I}^n)$ inside, the box-hull $[\mathbb{S}_\eta]$ of \mathbb{S}_η is

$$\phi_\eta = \operatorname{erf}\left(\sqrt{\sqrt{n} \cdot \operatorname{erf}^{-1}\left(2 \cdot \eta + \operatorname{erf}\left(-\frac{\sqrt{n}}{2}\right)\right) + \frac{n}{2}}\right)^n. \quad (8)$$

Proof. From Table 1 with $n \gg 1$, for a given confidence η , the radius α of \mathbb{S}_η is

$$\alpha = \sqrt{2 \cdot \sqrt{n} \left[\operatorname{erf}^{-1}\left(2\eta + \operatorname{erf}\left(-\frac{\sqrt{n}}{2}\right)\right) \right] + n}. \quad (9)$$

Now, $[\mathbb{S}_\eta]$ is the Cartesian product of n intervals $[e_i]$ of length 2α :

$$[\mathbb{S}_\eta] = [e_1] \times [e_2] \times \dots \times [e_n]. \quad (10)$$

From Table 1 with $n = 1$, we know that the probability to have $e_i \in [e_i]$ is

$$\operatorname{Pr}(e_i \in [e_i]) = \operatorname{erf}\left(\frac{\alpha}{\sqrt{2}}\right). \quad (11)$$

Therefore

$$\operatorname{Pr}(\mathbf{e} \in [\mathbb{S}_\eta]) = \prod_{i=1}^n \operatorname{Pr}(e_i \in [e_i]) = \operatorname{erf}\left(\frac{\alpha}{\sqrt{2}}\right)^n. \quad (12)$$

By combining (9) with (12), we get (8).

Remark. $\forall \eta > 0, \lim_{n \rightarrow +\infty} \operatorname{Pr}(\mathbf{e} \in [\mathbb{S}_\eta]) = 1$. It means that even for low values of η the probability $\operatorname{Pr}(\mathbf{e} \in [\mathbb{S}_\eta])$ increases dramatically fast with the dimension of \mathbf{e} . Therefore when n is large inverting $[\mathbb{S}_\eta]$ yields too much pessimism as illustrated by Figure 2.

Theorem. The minimal volume box \mathbb{B}_η which encloses $\mathbf{e} : \mathcal{N}(\mathbf{0}, \mathbf{I}^n)$ with a probability η is the centered cube with half-width

$$\alpha = \sqrt{2} \operatorname{erf}^{-1}(\sqrt[n]{\eta}). \quad (13)$$

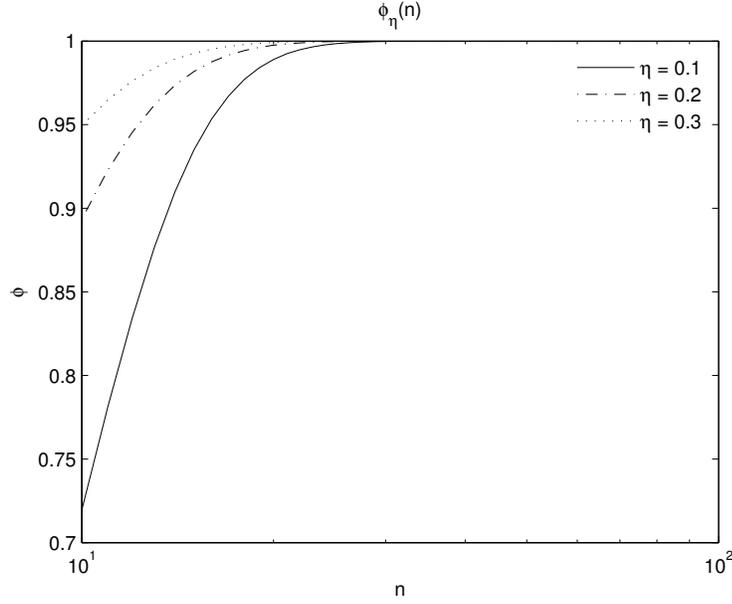


Figure 2: $Pr(\mathbf{e} \in [\mathbb{S}_\eta])$ as a function of n

Proof. The symmetry of the problem implies that \mathbb{B}_η should be centered. Since the e_i are independent, we have:

$$\begin{aligned} \eta &= Pr(\forall i, e_i \in [-\alpha, \alpha]) = \prod_{i=1}^n Pr(e_i \in [-\alpha, \alpha]) \\ &= \prod_{i=1}^n \text{erf}\left(\frac{\alpha}{\sqrt{2}}\right) = \left(\text{erf}\left(\frac{\alpha}{\sqrt{2}}\right)\right)^n \end{aligned} \quad (14)$$

i.e., $\alpha = \sqrt{2} \text{erf}^{-1}(\sqrt[n]{\eta})$.

Theorem. We have

$$\lim_{n \rightarrow \infty} \frac{\text{vol}(\mathbb{S}_\eta)}{\text{vol}(\mathbb{B}_\eta)} = 0. \quad (15)$$

Proof. Since the volume of a n -dimensional sphere \mathbb{S}_η of radius α is

$$V_n = \frac{\pi^{n/2} \alpha^n}{\Gamma_e(n/2 + 1)},$$

we have:

$$\rho_\eta(n) = \frac{\text{vol}(\mathbb{S}_\eta)}{\text{vol}(\mathbb{B}_\eta)} = \frac{\pi^{n/2} \cdot \sqrt{n + 2\sqrt{n} \cdot \text{erf}^{-1}\left(2\eta + \text{erf}\left(\frac{-\sqrt{n}}{2}\right)\right)}^n}{\Gamma_e(n/2 + 1) \cdot (2\sqrt{2} \text{erf}^{-1}(\sqrt[n]{\eta}))^n}$$

The Stirling formula $\Gamma_e(n + 1) = n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$ implies that

$$\rho_\eta(n) \sim \frac{\pi^{n/2} \cdot \sqrt{n + 2\sqrt{n} \cdot \text{erf}^{-1}\left(2\eta + \text{erf}\left(\frac{-\sqrt{n}}{2}\right)\right)}^n}{\sqrt{2\pi \frac{n}{2}} \left(\frac{n}{2e}\right)^{\frac{n}{2}} \cdot (2\sqrt{2} \text{erf}^{-1}(\sqrt[n]{\eta}))^n}.$$

Now, $2\eta + \operatorname{erf}\left(\frac{-\sqrt{n}}{2}\right) \sim 2\eta - 1$. Therefore

$$\rho_\eta(n) \sim \frac{\pi^{n/2} \cdot \sqrt{n + 2\sqrt{n} \cdot \operatorname{erf}^{-1}(2\eta - 1)}^n}{\sqrt{\pi n} \left(\frac{n}{2e}\right)^{\frac{n}{2}} \cdot (2\sqrt{2}\operatorname{erf}^{-1}(\sqrt{\eta}))^n}$$

Since, $n + 2\sqrt{n} \cdot \operatorname{erf}^{-1}(2\eta - 1) \sim n$, we get:

$$\begin{aligned} \rho_\eta(n) &\sim \frac{\pi^{n/2} \cdot \sqrt{n}^n}{\sqrt{\pi n} \left(\frac{n}{2e}\right)^{\frac{n}{2}} \cdot (2\sqrt{2}\operatorname{erf}^{-1}(\sqrt{\eta}))^n} \\ &= \frac{\pi^{n/2} \cdot n^{\frac{n}{2}} \cdot (2e)^{\frac{n}{2}}}{\sqrt{\pi n}^{\frac{n}{2}} \cdot (2\sqrt{2}\operatorname{erf}^{-1}(\sqrt{\eta}))^n} \\ &\sim \frac{(e\pi)^{\frac{n}{2}}}{(2\operatorname{erf}^{-1}(\sqrt{\eta}))^n} = \left(\frac{\sqrt{e\pi}}{2\operatorname{erf}^{-1}(\sqrt{\eta})}\right)^n \end{aligned}$$

which converges to zero.

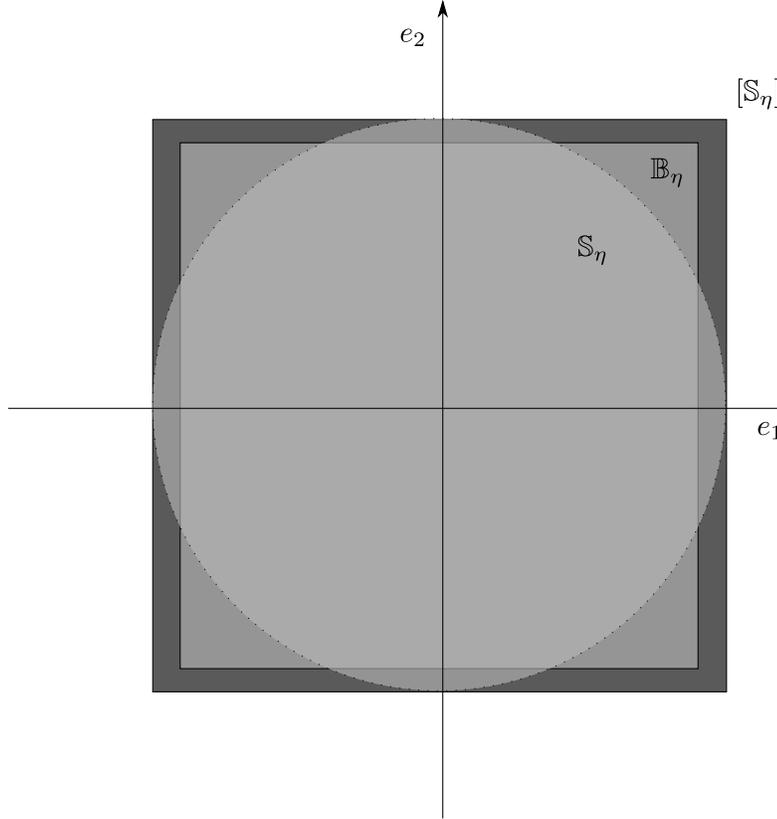


Figure 3: Idealized representation for $\mathbb{S}_\eta, \mathbb{B}_\eta, [\mathbb{S}_\eta]$

Figure 3 illustrates the configuration for the sets $\mathbb{S}_\eta, \mathbb{B}_\eta, [\mathbb{S}_\eta]$ that are used to approximate the error vector \mathbf{e} . Both $\mathbb{S}_\eta, \mathbb{B}_\eta$ contain \mathbf{e} with a probability η .

3 Linearization method

To compute a set which encloses the parameter vector \mathbf{e} with a probability η , the previous section proposed to compute the probabilistic set $\hat{\mathbb{P}}_{\mathbb{E}_\eta}$ associated to \mathbb{E}_η , the set which contains \mathbf{e} with a probability η . This set can be expressed as the set inversion problem $\hat{\mathbb{P}}_{\mathbb{E}_\eta} = \mathbf{f}^{-1}(\mathbb{E}_\eta)$ where $\mathbf{f}(\mathbf{p}) = \mathbf{y} - \boldsymbol{\psi}(\mathbf{p})$. For a comparison, we recall the classical the *Maximum Likelihood* approach to estimate such a confidence set by a linearization of the model. Unfortunately, the linearization error cannot be quantified in a reliable way.

The linearization method searches for the parameter vector $\hat{\mathbf{p}}$ which maximizes the *likelihood function*

$$L(y_i | \mathbf{p}) = \prod_i \pi(y_i | \mathbf{p}) \propto \prod_i e^{-(\psi_i(\mathbf{p}) - y_i)^2}. \quad (16)$$

This is equivalent to minimizing

$$\lambda(\mathbf{p}) = -\log L(y_i | \mathbf{p}) = \sum_i (\psi_i(\mathbf{p}) - y_i)^2 \quad (17)$$

which corresponds to a non-linear least-square minimization problem. It seems reasonable to assume that the true value for \mathbf{p} is closed to the minimizer $\hat{\mathbf{p}}$ and that $\lambda(\mathbf{p})$ can be approximated by a second order Taylor development of Equation 17 around $\hat{\mathbf{p}}$. Since the gradient of λ at $\hat{\mathbf{p}}$ is zero, we get

$$\lambda(\mathbf{p}) \sim \lambda(\hat{\mathbf{p}}) + \frac{1}{2} \cdot (\mathbf{p} - \hat{\mathbf{p}})^T \cdot \mathbf{H}_\lambda(\hat{\mathbf{p}}) \cdot (\mathbf{p} - \hat{\mathbf{p}}) \quad (18)$$

where \mathbf{H}_λ is the Hessian matrix of λ . Now $\mathbf{e}^T \cdot \mathbf{e} \leq \alpha(\eta)^2 \Leftrightarrow \lambda(\mathbf{p}) = \sum_i (\psi_i(\mathbf{p}) - y_i)^2 \leq \alpha^2(\eta)$. As a consequence, a confidence ellipsoid which contains \mathbf{p} with a probability η is:

$$\lambda(\hat{\mathbf{p}}) + \frac{1}{2} \cdot (\mathbf{p} - \hat{\mathbf{p}})^T \cdot \mathbf{H}_\lambda(\hat{\mathbf{p}}) \cdot (\mathbf{p} - \hat{\mathbf{p}}) \leq \alpha^2(\eta).$$

Note that $\mathbf{H}_\lambda(\hat{\mathbf{p}})$ corresponds to the observed Fisher information matrix at $\hat{\mathbf{p}}$ [31][2] which is the inverse of the covariance matrix $\boldsymbol{\Sigma}_{\hat{\mathbf{p}}}$ for the estimated maximum likelihood parameter $\hat{\mathbf{p}}$. Note also that the linearization method provides an ellipsoid associated to the probability η but this ellipsoid cannot be considered as reliable: the probability that it contains \mathbf{p} is most of the time far from η .

4 Test-cases

To illustrate our method, we consider here three illustrative test-cases involving parameter estimation under white, additive Gaussian noise.

4.1 Test-case 1

Consider the following model

$$y(t) = p_2 \cdot e^{-p_1 \cdot t} + p_1 \cdot e^{-p_2 \cdot t} + w(t) \quad (19)$$

where $t \in \{0, 0.01, 0.02, \dots, 12\}$ and $w(t)$ is a white centred Gaussian noise with a variance $\sigma^2 = 1$. Figure 4 represents the collected data $y(t)$.

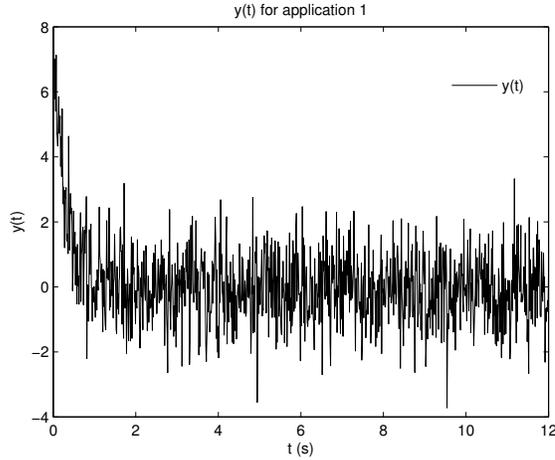
Figure 4: Measurements $y(t)$ for Test-case 1

Figure 5 represents the three sets $\hat{\mathbb{P}}_{0.99}$ obtained by an inversion of $[\mathbb{S}_{0.99}]$, $\mathbb{B}_{0.99}$ and $\mathbb{S}_{0.99}$. This comparison confirms that the box-hull inversion $\hat{\mathbb{P}}_{[\mathbb{S}_{0.99}]}$ is too pessimistic. Figure 6 illustrates a situation where $\hat{\mathbb{P}}_{\mathbb{S}_{0.99}} \not\subset \hat{\mathbb{P}}_{\mathbb{B}_{0.99}}$. From Theorem 2 we could have expected an inclusion. Now, this example is quite atypical: the parametric model is not globally identifiable, *i.e.*, p_1 and p_2 can be interchanged without any effect on the output. Figure 6 also represents the confidence ellipsoid generated by the linear estimator. Due to the non identifiability problem, we have two global minimizers. We have chosen to draw the ellipsoid centred around the minimizer corresponding to the true parameter vector \mathbf{p}^* . Otherwise, the 0.99 ellipsoid would not contains \mathbf{p}^* .

4.2 Test-case 2

Consider the following model studied in [15]

$$y(t) = 20 \cdot e^{-p_1 \cdot t} - 8 \cdot e^{-p_2 \cdot t} + w \quad (20)$$

which is similar to the model of Test-case 1 but the model is now identifiable. Again, $w(t)$ is a centred normal noise with a unit variance. We collected 1000 measurements for $y(t)$ at different times $t \in [0, 25]$ as represented on Figure 7.

Figure 8 shows that the inversion $\hat{\mathbb{P}}_{\mathbb{S}_{0.99}}$ of the confidence sphere $\mathbb{S}_{0.99}$ is more precise than the inversion of $[\mathbb{S}_{0.99}]$ and $\mathbb{B}_{0.99}$. The set $\hat{\mathbb{P}}_{\mathbb{S}_{0.99}}$ has two disjoint components at a confidence level $\eta = 0.99$. Figure 9 shows that the linear estimator was able to capture the correct parameters vector.

Remark. Figure 8 shows that the proposed approach suffers from an important pessimism: the border of the computed set is quite thick, and the generated subpaving is not minimal. This is due to the multiple-occurrences in the parameter variables in the expression of the inequalities describing \mathbb{S}_η . Interval methods are sensitive to this type of situation which adds pessimism in the propagation of uncertainties [13]. To limit this phenomena, linear approximations such as the centered or affine forms of

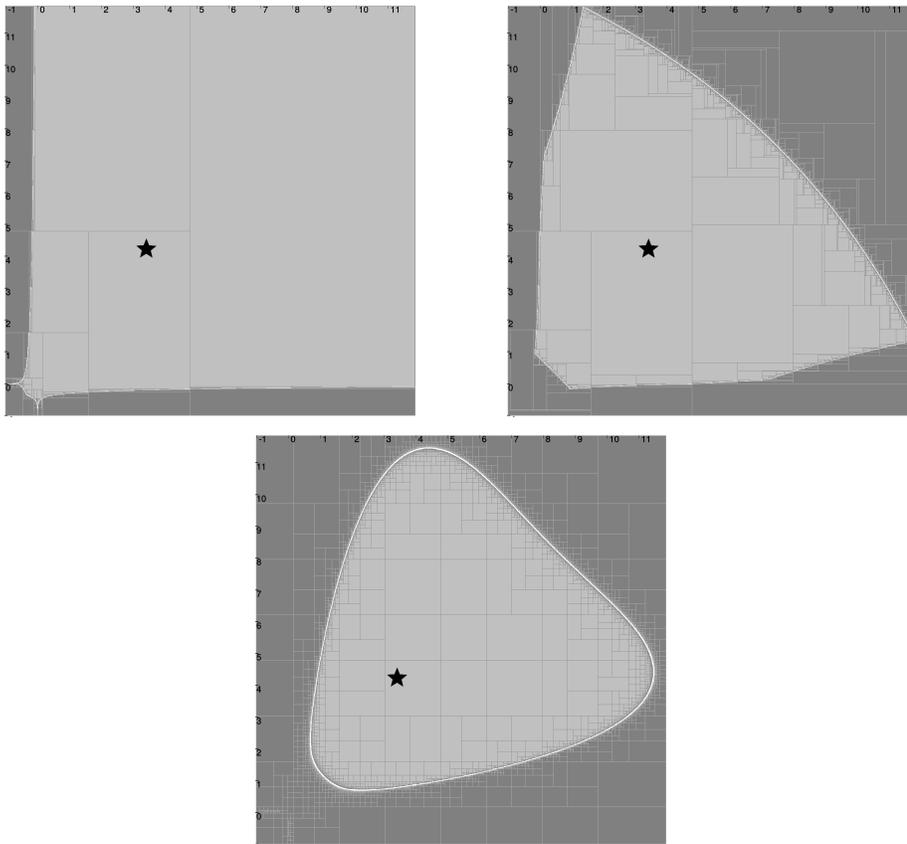


Figure 5: $\hat{\mathbb{P}}_{[S_{0.99}]}$ (top left), $\hat{\mathbb{P}}_{\mathbb{E}_{0.99}}$ (top right), $\hat{\mathbb{P}}_{S_{0.99}}$ (bottom). The black star is the true parameters vector.

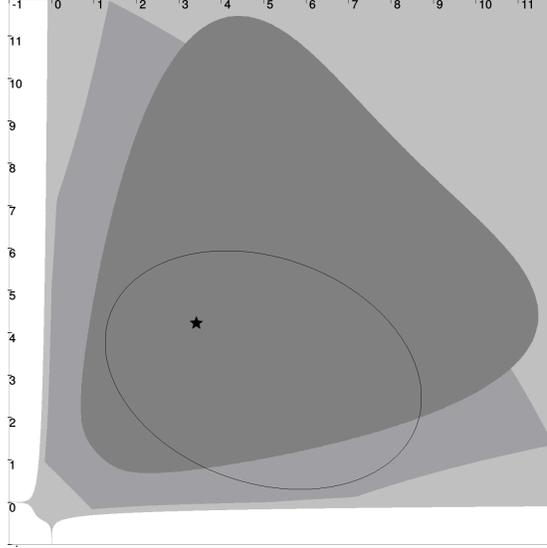


Figure 6: Superposition of $\hat{\mathbb{P}}_{[S_{0.99}]}$ (light gray), $\hat{\mathbb{P}}_{B_{0.99}}$ (gray), $\hat{\mathbb{P}}_{S_{0.99}}$ (dark gray), and the 0.99 confidence ellipse obtained with a linear estimator. The black star is the true parameter vector \mathbf{p}^*

the constraints could be used.

4.3 Test-case 3

In this example, a lost underwater vehicle tries to get its position by gathering range-only measurements to three beacons [4] [12]. The position $\mathbf{x}_j = (x_j \ y_j \ z_j)$ of the j^{th} beacon is precisely known from a previous survey of the area, as well as the altitude z_m of the robot, thanks to a pressure sensor. The three beacons are almost aligned, which causes a bad conditioning. The robot is assumed to be static during the acquisition. For each measurement \tilde{d}_i to the beacon j we have

$$\tilde{d}_{ij} = \sqrt{(x_j - x_m)^2 + (y_j - y_m)^2 + (z_j - z_m)^2} + w \quad (21)$$

where w is a white centred Gaussian noise, whose variance is given by the sensor for each measurement. The signals associated to the three beacons are pictured in Figure 10.

From Figure 12, we observe that $\hat{\mathbb{P}}_{S_{0.99}} \subset \hat{\mathbb{P}}_{B_{0.99}} \subset \hat{\mathbb{P}}_{[S_{0.99}]}$, which confirms that the $\hat{\mathbb{P}}_{S_{0.99}}$ is more precise than the two other confidence regions. Figure 13 is the superposition of $\hat{\mathbb{P}}_{S_{0.99}}$, $\hat{\mathbb{P}}_{B_{0.99}}$, $\hat{\mathbb{P}}_{[S_{0.99}]}$ and the 0.99 confidence ellipse (flat and horizontal) of a linear estimator. While the linear estimator gives an estimate that is consistent (it contains the true solution), it is obvious that it doesn't fully capture the underlying banana-shaped probability density function, which is more accurately seized by our nonlinear methods.

Table 2 compares the time it takes to compute $\hat{\mathbb{P}}_{[S_{0.99}]}$, $\hat{\mathbb{P}}_{B_{0.99}}$, $\hat{\mathbb{P}}_{S_{0.99}}$ on a classical laptop for the three test-cases. As it could have been anticipated, it is clear that

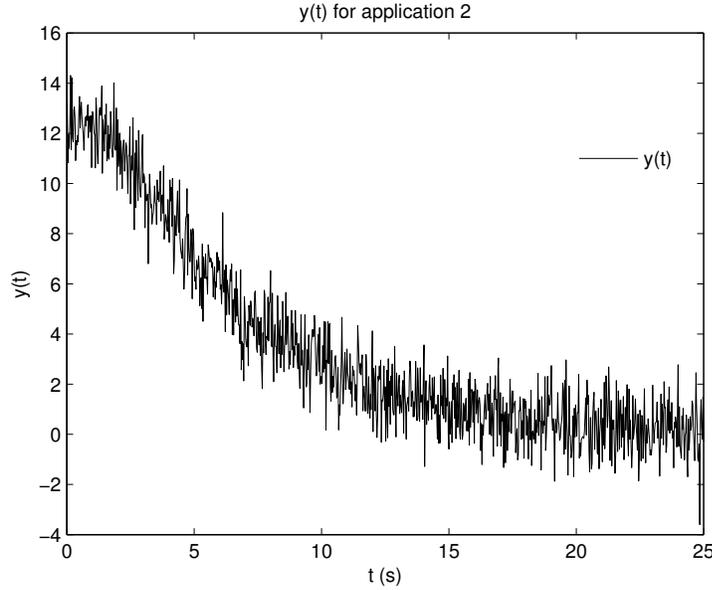


Figure 7: Collected data $y(t)$ for Test-case 2

inverting boxes, which are convenient representations for interval methods, takes much less time than inverting a sphere.

Computation time (s)	Test-case 1	Test-case 2	Test-case 3
$\hat{\mathbb{P}}_{[S_{0.99}]}$	35 sec	1sec	26sec
$\hat{\mathbb{P}}_{\mathbb{B}_{0.99}}$	62sec	6sec	45sec
$\hat{\mathbb{P}}_{S_{0.99}}$	839sec	89sec	510sec

Table 2: Computation times for Test-cases 1, 2 and 3

5 Conclusion

In this paper, we have presented a new approach for parameter estimation of nonlinear models with additive Gaussian noise. The resulting method makes it possible to compute a set which contains the parameter vector with a given probability. The main contribution of this paper is that the results are guaranteed, which is not the case for existing approaches. Indeed, although if existing methods are also able to provide an estimation of such a confidence region of probability η , they perform some linearizations without quantifying the corresponding error. Three simulated test-cases were presented and compared to existing and linear methods.

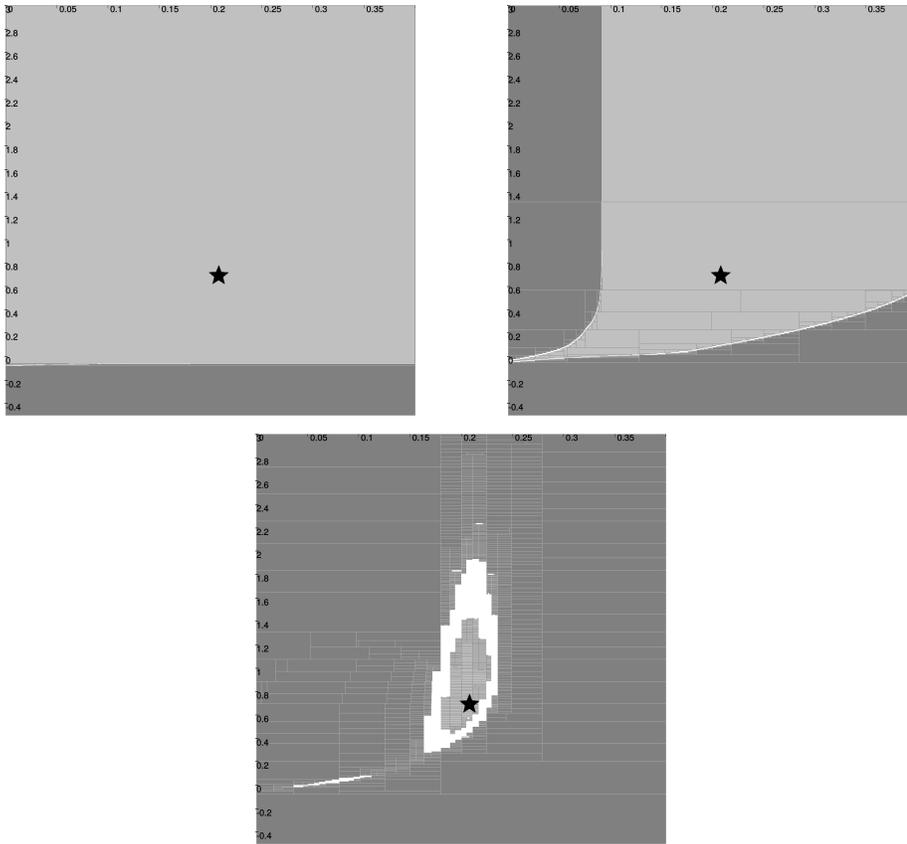


Figure 8: $\hat{\mathbb{P}}_{[\mathcal{S}_{0.99}]}$ (top left), $\hat{\mathbb{P}}_{\mathbb{B}_{0.99}}$ (top right), $\hat{\mathbb{P}}_{\mathcal{S}_{0.99}}$ (bottom). The black star is the true parameter vector \mathbf{p}^*

References

- [1] F. Abdallah, A. Gning, and P. Bonnifait. Box particle filtering for nonlinear state estimation using interval analysis. *Automatica*, 44(3):807–815, 2008.
- [2] J. Berger. *Statistical Decision Theory and Bayesian Analysis*, 2nd edition. Springer-Verlag, New York, NY, 1985.
- [3] M.C. Campi and E. Weyer. Guaranteed non-asymptotic confidence regions in system identification. *Automatica*, 41:1751–1764, 2005.
- [4] V. Creuze. Robots marins et sous-marins ; perception, modélisation, commande. *Techniques de l'ingénieur*, 2014.
- [5] B.C. Csaji, M.C. Campi, and E. Weyer. Non-asymptotic confidence regions for the least-squares estimate. In *the 16th IFAC Symposium on System Identification*, 2012.

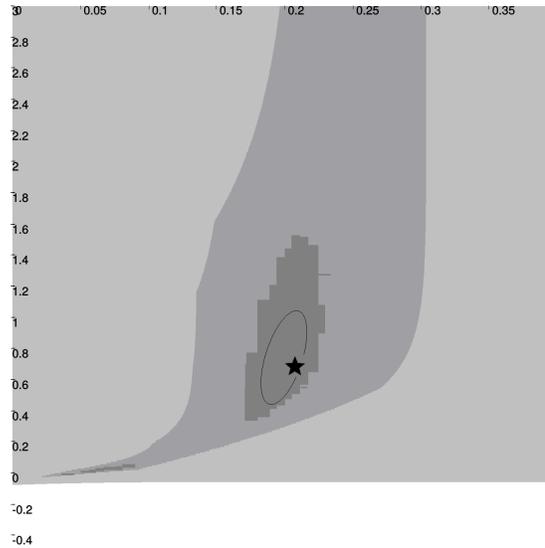


Figure 9: Superposition of $\hat{\mathbb{P}}_{[S_{0.99}]}$ (light gray), $\hat{\mathbb{P}}_{\mathbb{B}_{0.99}}$ (gray), $\hat{\mathbb{P}}_{S_{0.99}}$ (dark gray), and the 0.99 confidence ellipse obtained with a linear estimator.

- [6] V. Drevelle and P. Bonnifait. Localization confidence domains via set inversion on short-term trajectory. *IEEE Transactions on Robotics*, 2013.
- [7] D. Dubois and H. Prade. *Fuzzy Sets and Systems-Theory and Applications*. Academic Press, New York, NY, 1980.
- [8] P. Eykhoff. *System Identification, Parameter and State Estimation*. John Wiley, London, 1974.
- [9] A. Gning, B. Ristic, L. Mihaylova, and F. Abdallah. An Introduction to Box Particle Filtering. *IEEE Signal Processing Magazine*, 30(1):166–171, 2013.
- [10] L. Jaulin. Computing minimal-volume credible sets using interval analysis; application to bayesian estimation. *IEEE Trans. on Signal Processing*, 54(9):3632–3636, 2006.
- [11] L. Jaulin. Probabilistic set-membership approach for robust regression. *Journal of Statistical Theory and Practice*, 4(1), 2010.
- [12] L. Jaulin. *Mobile Robotics*. ISTE editions, 2015.
- [13] 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.
- [14] L. Jaulin and E. Walter. Guaranteed nonlinear parameter estimation via interval computations. In *Conference on Numerical Analysis with Automatic Result Verification, (Lafayette)*, 1993.

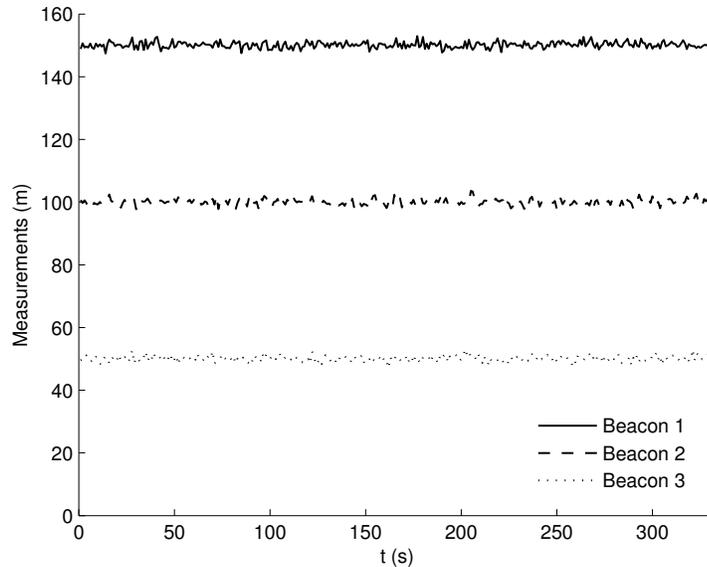


Figure 10: Range signals received from the three beacons

- [15] L. Jaulin and E. Walter. *Bounding Approaches to System Identification*, chapter Guaranteed nonlinear set estimation via interval analysis, pages 363–382. Plenum, 1996.
- [16] M. Kieffer and E. Walter. Guaranteed Characterization of Exact non-Asymptotic Confidence Regions as Defined by LSCR and SPS. *Automatica*, 50(2):507–512, 2014.
- [17] V. Kreinovich, G.P. Dimuro, and A. Carlos 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.
- [18] V. Kreinovich, A.V. Lakeyev, and S.I. Noskov. Approximate Linear Algebra is Intractable. *Linear Algebra and its Applications*, 232:45–54, 1996.
- [19] V. Kreinovich, A.V. Lakeyev, J. Rohn, and P.T. Kahl. Computational complexity and feasibility of data processing and interval computations. *Reliable Computing*, 4(4):405–409, 1997.
- [20] V. Kreinovich and S. Shary. Interval methods for data fitting under uncertainty: A probabilistic treatment. *Reliable Computing*, 23:105–140, 2016.
- [21] L. Longpré, C. Servin, and V. Kreinovich. Quantum computation techniques for gauging reliability of interval and fuzzy data. *International Journal of General Systems*, 40(1):99–109, 2011.

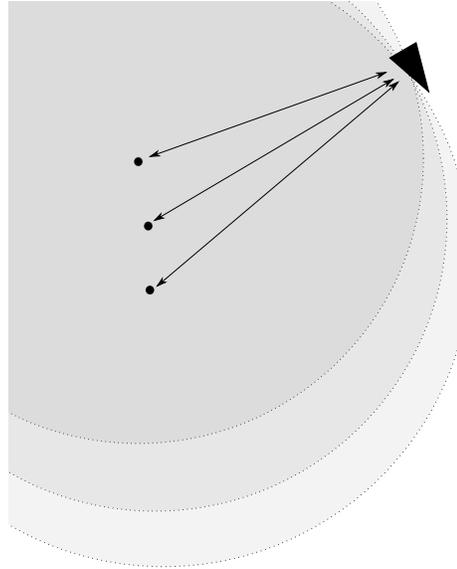


Figure 11: An underwater robot stays fixed on the seafloor and gathers range measurements from 3 beacons whose positions are known, in order to estimate its position

- [22] M. Milanese, J. Norton, H. Piet-Lahanier, and E. Walter, editors. *Bounding Approaches to System Identification*. Plenum Press, New York, NY, 1996.
- [23] I. Molchanov. *The Theory of Random Sets*. Springer, New York, 2005.
- [24] R. E. Moore. *Methods and Applications of Interval Analysis*. SIAM, Philadelphia, PA, 1979.
- [25] R. Neuland, J. Nicola, R. Maffei, L. Jaulin, E. Prestes, and M. Kolberg. Hybridization of monte carlo and set-membership methods for the global localization of underwater robots. In *IROS 2014*, pages 199–204, 2014.
- [26] J. P. Norton, editor. Special Issue on Bounded-Error Estimation: Issue 1. 1994. *International Journal of Adaptive Control and Signal Processing* **8**(1):1–118.
- [27] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, New York, 1984.
- [28] B. Ristic. Bayesian estimation with imprecise likelihoods: Random set approach. *Signal Processing Letters*, 18(7), 2011.
- [29] J. Sliwka, L. Jaulin, M. Ceberio, and V. Kreinovich. Processing interval sensor data in the presence of outliers, with potential applications to localizing underwater robots. In *IEEE SMC*, Anchorage, Alaska, 2011.

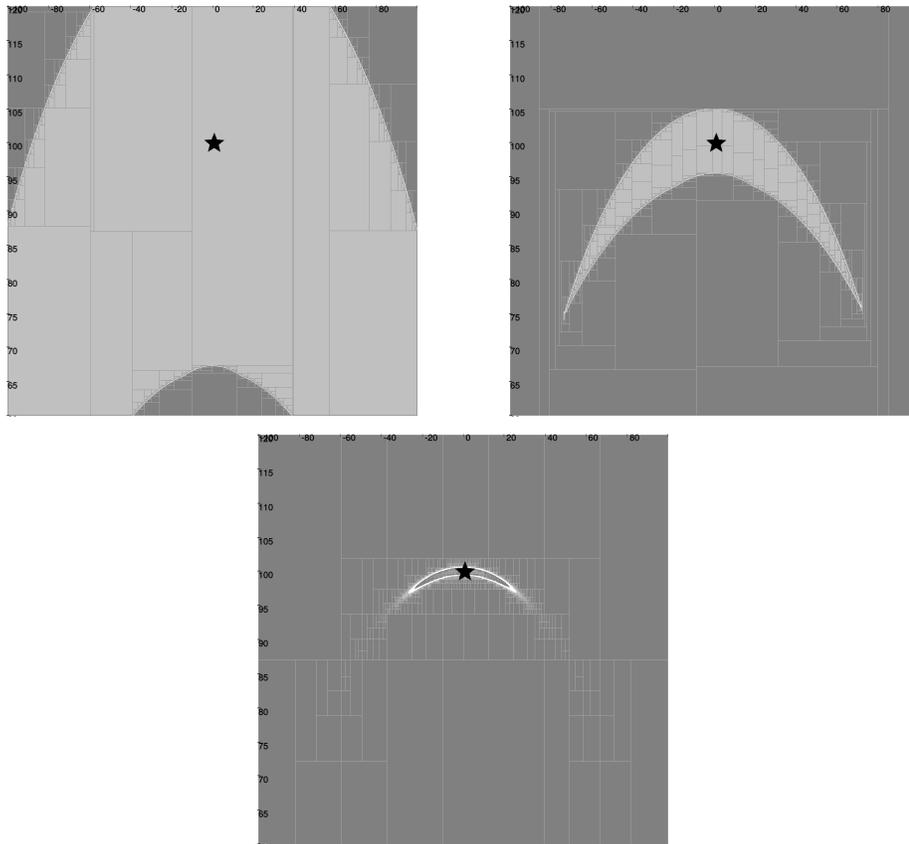


Figure 12: $\hat{\mathbb{P}}_{\mathbb{S}_{0.99}}$ (top left), $\hat{\mathbb{P}}_{\mathbb{B}_{0.99}}$ (top right), $\hat{\mathbb{P}}_{\mathbb{S}_{0.99}}$ (bottom). The black star represents \mathbf{p}^*

- [30] O. Strauss. Quasi-continuous histograms. *Fuzzy Sets and Systems*, 160(17):2442–2465, 2009.
- [31] E. Walter and L. Pronzato. *Identification of Parametric Models from Experimental Data*. Springer-Verlag, London, UK, 1997.
- [32] J. Xiong, C. Jauberthie, L. Travé-Massuyes, and F. Le Gall. Fault detection using interval kalman filtering enhanced by constraint propagation. In *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*, pages 490–495. IEEE, 2013.
- [33] J.T. Yao, Y.Y. Yao, V. Kreinovich, P. Pinheiro da Silva, S.A. Starks, G. Xi-ang, and H. T. Nguyen. Towards more adequate representation of uncertainty: From intervals to set intervals, with the possible addition of probabilities and certainty degrees. In *Proceedings of the IEEE World Congress on Computational Intelligence WCCI'2008*, pages 983–990, Hong Kong, China, 2008.

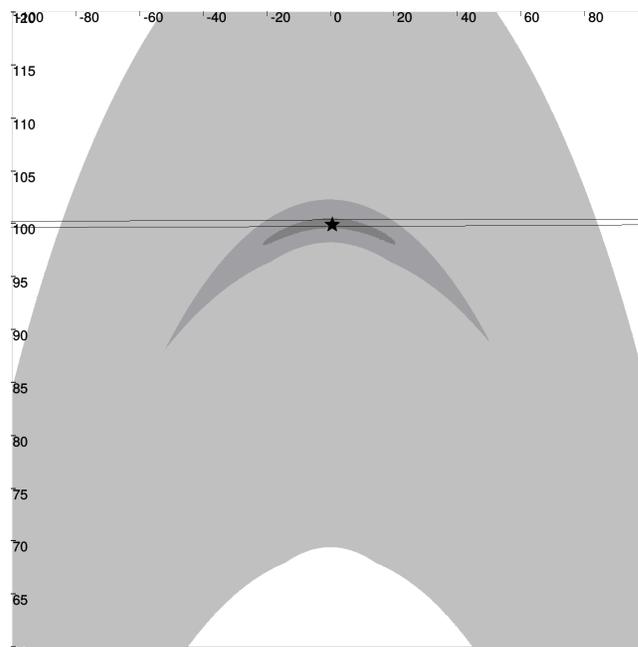


Figure 13: Superposition of $\hat{\mathbb{P}}_{[S_{0.99}]}$ (light gray), $\hat{\mathbb{P}}_{B_{0.99}}$ (gray), $\hat{\mathbb{P}}_{S_{0.99}}$ (dark gray), and the 0.99 confidence ellipse of the linear estimator (black). The black star represents \mathbf{p}^* .