Probabilistic set-membership estimation

Luc Jaulin

www.ensieta.fr/jaulin/
ENSIETA, Brest

Munich, September 15, 2009

Second Workshop on Principles and Methods of Statistical Inference with Interval Probability

1 Probabilistic-set approach

Bounded-error estimation

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

where

- $\mathbf{e} \in \mathbb{E} \subset \mathbb{R}^m$ is the error vector,
- $\mathbf{y} \in \mathbb{R}^m$ is the collected data vector,
- $\mathbf{p} \in \mathbb{R}^n$ is the parameter vector to be estimated.

Or equivalently

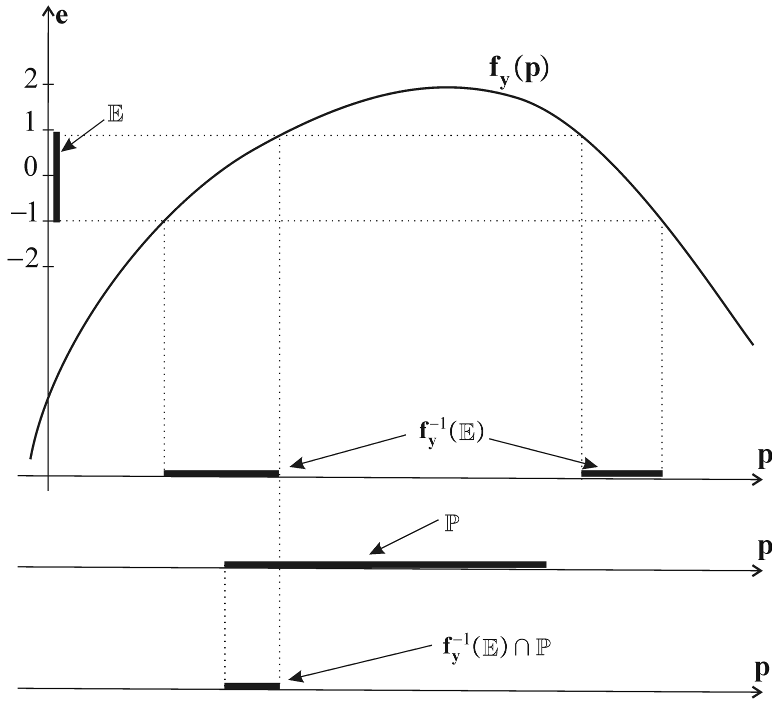
$$\mathbf{e} = \mathbf{f}(\mathbf{y}, \mathbf{p}) = \mathbf{f}_{\mathbf{y}}(\mathbf{p}),$$

where

$$\mathbf{f}_{\mathbf{y}}(\mathbf{p}) = \mathbf{y} - \boldsymbol{\psi}(\mathbf{p}).$$

The *posterior feasible set* for the parameters is

$$\widehat{\mathbb{P}} = \mathbf{f}_{\mathbf{y}}^{-1}(\mathbb{E}) \cap \mathbb{P}.$$



About interval methods

Interval methods provide guaranteed results only if some assumptions (bounds on the errors, constraints, model, ...) are satisfied.

In practice we are not able to give 100% reliable assumptions, but we can associate some probabilities on them.

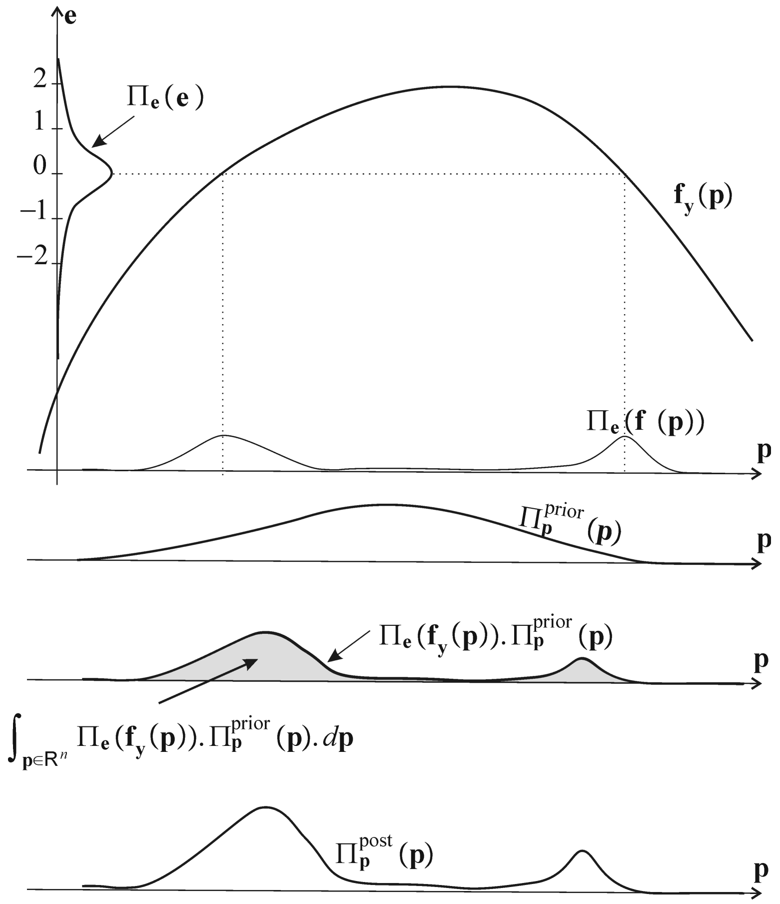
Show setdemo

For parameter estimation, if that the assumptions are satisfied with a probability π, the solution set encloses the true value for the parameter vector with a probability $> \pi$.

In a Bayesian approach, prior pdf $\Pi_{\mathbf{e}}, \Pi_{\mathbf{p}}^{\text{prior}}$ are known for $\mathbf{e}, \mathbf{p}$.

The Bayes rule gives us the posterior pdf for $\mathbf{p}$

$$\Pi_{\mathbf{p}}^{\text{post}}(\mathbf{p}) = \frac{\Pi_{\mathbf{e}}(\mathbf{f_y}(\mathbf{p})).\Pi_{\mathbf{p}}^{\text{prior}}(\mathbf{p})}{\int_{\mathbf{p}\in\mathbb{R}^n} \Pi_{\mathbf{e}}(\mathbf{f_y}(\mathbf{p})).\Pi_{\mathbf{p}}^{\text{prior}}(\mathbf{p}).d\mathbf{p}}.$$



e
2
1
0
−1
−2
Πe(e)
fy(p)
Πe(f (p))
p
Πp prior(p)
p
Πe(fy(p)).Πp prior(p)
p
∫p∈R^n Πe(fy(p)).Πp prior(p).dp
Πp post(p)
p

Probabilistic-set approach. We decompose the error space into two subsets: $\mathbb{E}$ on which we bet $\mathbf{e}$ will belong and $\overline{\mathbb{E}}$.

We set

$$\pi = \mathsf{Pr}\left(\mathbf{e} \in \mathbb{E}\right)$$

The event $\mathbf{e} \in \overline{\mathbb{E}}$ is considered as rare, i.e.,

$$\pi \simeq 1$$

Once $\mathbf{y}$ is collected, we compute

$$\widehat{\mathbb{P}} = \mathbf{f}_{\mathbf{y}}^{-1}(\mathbb{E}) \cap \mathbb{P}.$$

If now $\widehat{\mathbb{P}} \neq \emptyset$, we conclude that $\mathbf{p} \in \widehat{\mathbb{P}}$ with a probability of π.

If $\widehat{\mathbb{P}} = \emptyset$, than we conclude the rare event $\mathbf{e} \in \overline{\mathbb{E}}$ occurred.

Example 1. The model is described by $y = p^2 + e$, *i.e.*,

$$e = y - p^2 = f_y(p)$$

Assume that $\Pi_e : \mathcal{N}(0,1)$. If $\mathbb{E} = [-6, 6]$ then,

$$\Pr\left(e \in \overline{\mathbb{E}}\right) = -\frac{1}{\sqrt{2\pi}} \int_{-6}^{6} \exp\left(-\frac{e^2}{2}\right) de \simeq 1.97 \times 10^{-9}.$$

We now collect $y = 10$. We have

$$
\begin{aligned}
\widehat{\mathbb{P}} &= f_y^{-1}(\mathbb{E}) \cap \mathbb{P} = f_y^{-1}([-6, 6]) \cap [-\infty, \infty] \\
&= \sqrt{10 - [-6, 6]} = \sqrt{[4, 16]} = [-4, -2] \cup [2, 4].
\end{aligned}
$$

with a prior probability of $1 - 1.97 \times 10^{-9}$.

Let us apply the Bayesian approach, with $\Pi_p^{\text{prior}} : \mathcal{N}(3, 1)$.
The posterior pdf for p is

$$\begin{aligned}
\Pi_p^{\text{post}}(p) &= \frac{\Pi_e(f_y(p)).\Pi_p^{\text{prior}}(p)}{\int_{p\in\mathbb{R}} \Pi_e(f_y(p)).\Pi_p^{\text{prior}}(p)dp} \\
&= \frac{e^{-\frac{\left(10-p^2\right)^2}{2}}.e^{-\frac{(p-3)^2}{2}}}{\int_{-\infty}^{\infty} e^{-\frac{\left(10-p^2\right)^2}{2}}.e^{-\frac{(p-3)^2}{2}}.dp} \\
&\simeq 2.57\ e^{-\frac{p^4-19p^2-6p+109}{2}}.
\end{aligned}$$

Example 2. Now $y = -10$. Since

$$\widehat{\mathbb{P}} = f_y^{-1}(\mathbb{E}) = \emptyset,$$

the probabilistic-set approach concludes to an inconsistency. The Bayesian approach gives

$$\Pi_p^{\text{post}}(p) \simeq 6.930\,5 \times 10^{23}.e^{-\frac{p^4-39p^2-6p+409}{2}}.$$

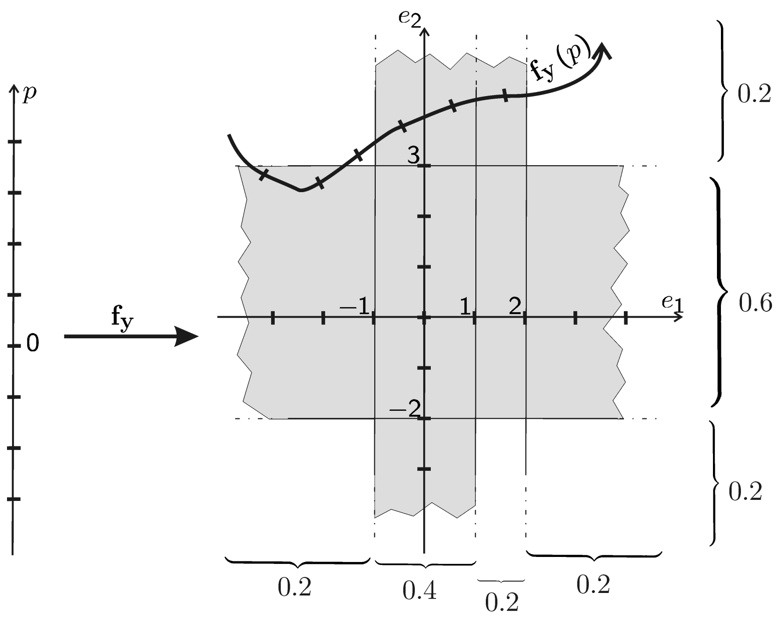
which corresponds to a precise posterior pdf for p around $p = 4.45$.

In practice, the huge factor $(6.930\,5 \times 10^{23})$ is interpreted as an inconsistency.

Example 3. Assume that

$$
\begin{array}{llll}
\Pr(e_1 \leq -1) & = 0.2, & \Pr(e_2 \leq -2) & = 0.2, \\
\Pr(e_1 \in [-1, 1]) & = 0.4, & \Pr(e_2 \in [-2, 3]) & = 0.6, \\
\Pr(e_1 \in [1, 2]) & = 0.2, & \Pr(e_2 \geq 3) & = 0.2, \\
\Pr(e_1 \geq 2) & = 0.2. & &
\end{array}
$$

and that e_1 and e_2 are independent.


p
0
$\mathbf{f_y}$
e_2
$\mathbf{f_y}(p)$
3
−1
1
2
e_1
−2
0.2
0.6
0.2
0.2
0.4
0.2
0.2

The joint pdf for (e_1, e_2) is

$[e_2] \backslash [e_1]$	$[-\infty, -1]$	$[-1, 1]$	$[1, 2]$	$[2, \infty]$
$[3, \infty]$	0.04	$\mathbf{0.08}$	$\mathbf{0.04}$	0.04
$[-2, 3]$	$\mathbf{0.12}$	$\mathbf{0.24}$	$\mathbf{0.12}$	$\mathbf{0.12}$
$[-\infty, -2]$	0.04	$\mathbf{0.08}$	0.04	0.04

Thus

$\Pr(\mathbf{e} \in \mathbb{E}) = 0.08 + 0.04 + 0.12 + 0.24 + 0.12 + 0.12 + 0.08 = 0.8.$

$\widehat{\mathbb{P}} = \mathbf{f}_{\mathbf{y}}^{-1}(\mathbb{E})$ encloses $\mathbf{p}$ with a prior probability of 0.8.

2 Robust regression

Consider the error model

$$\mathbf{e} = \mathbf{f}_{\mathbf{y}}(\mathbf{p}).$$

y_i is an *inlier* if $e_i \in [e_i]$ and an *outlier* otherwise. We assume that

$$\forall i, \ \mathsf{Pr}(e_i \in [e_i]) = \pi$$

and that all e_i's are independent.

Equivalently,

$$\begin{cases} f_1(\mathbf{y}, \mathbf{p}) \in [e_1] & \text{with a probability } \pi \\ \vdots & \vdots \\ f_m(\mathbf{y}, \mathbf{p}) \in [e_m] & \text{with a probability } \pi \end{cases}$$

The number k of inliers follows a binomial distribution

$$\frac{m!}{k!\,(m-k)!}\pi^k.\,(1-\pi)^{m-k}\,.$$

The probability of having strictly more than q outliers is thus

$$\gamma\left(q, m, \pi\right) \overset{\text{def}}{=} \sum_{k=0}^{m-q-1} \frac{m!}{k!\left(m-k\right)!} \pi^{k}.\left(1-\pi\right)^{m-k}.$$

Example. For instance, if $m = 1000$, $q = 900, \pi = 0.2$, we get $\gamma\left(q, m, \pi\right) = 7.04 \times 10^{-16}$. Thus having more than 900 outliers can be seen as a rare event.

Denote by $\mathbb{E}$ the set of all $\mathbf{e} \in \mathbb{R}^m$ such that the number of outliers is smaller (or equal) than q.

$\widehat{\mathbb{P}} = \mathbf{f}_{\mathbf{y}}^{-1}(\mathbb{E})$ will contain the parameter vector with a prior probability of $1 - \gamma(q, m, \pi)$.

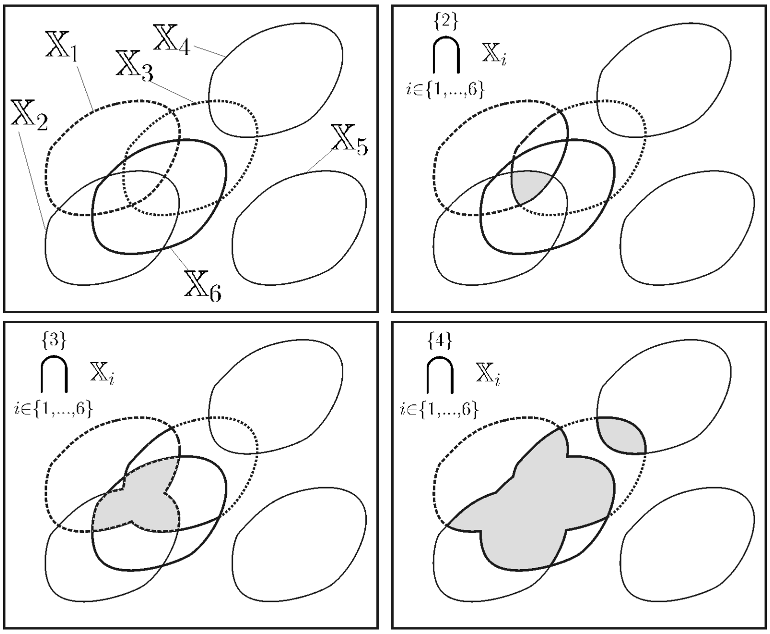


Illustration the q-relaxed intersection

Demo of Jan Sliwka

3 Test case

Generation of data. $m = 500$ data are generated as follows

$$y_i = p_1 \sin(p_2 t_i) + e_i, \text{ with a probability 0.2.}$$
$$y_i = r_1 \exp(r_2 t_i) + e_i, \text{ with a probability 0.2.}$$
$$y_i = n_i$$

where $t_i = 0.02*(i+1)$, $i \in \{1, 500\}$, $e_i : \mathcal{U}([-0.1, 0.1])$ and $n_i : \mathcal{N}(2, 3)$.

We took $\mathbf{p}^* = (2, 2)^\top$ and $\mathbf{r}^* = (4, -0.4)^\top$.

Estimation. We know that

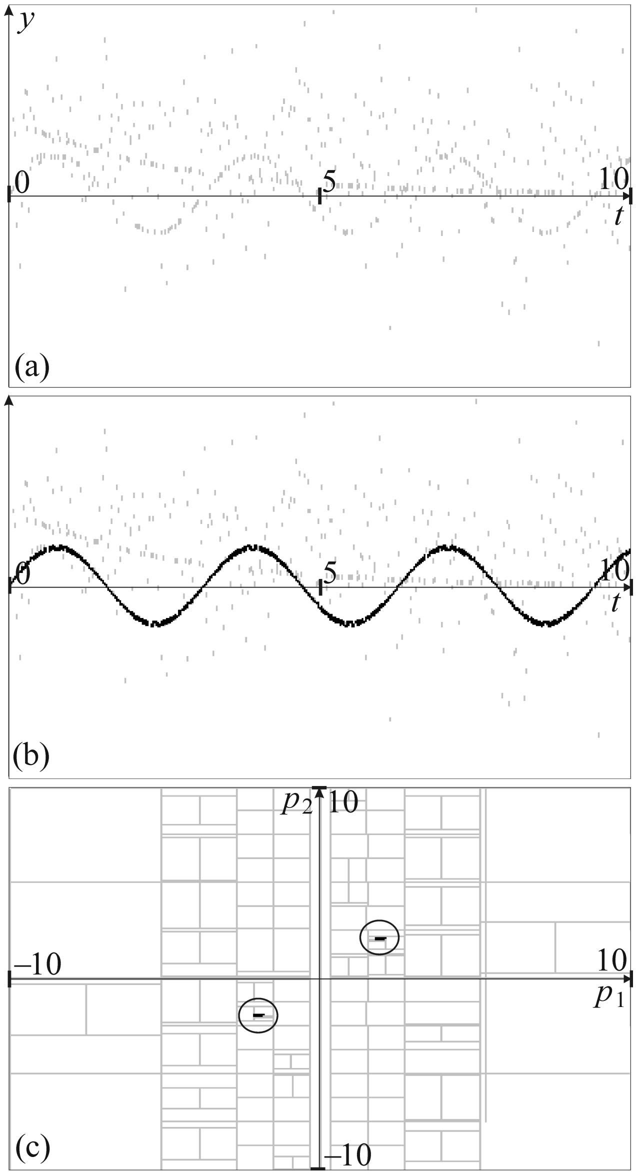
$$y_i = p_1 \sin(p_2 t_i) + e_i, \text{with a probability 0.2.}$$

and that we have no idea of what happen otherwise.

We want

$$\Pr\left(\mathbf{p}^* \in \widehat{\mathbb{P}}\right) \geq 0.95$$

Since $\gamma(414, 500, 0.2) = 0.0468$ and $\gamma(413, 500, 0.2) = 0.12$, we should assume $q = 414$ outliers.


y
0
5
10
t
(a)
0
5
10
t
(b)
p_2
10
−10
−10
10
p_1
(c)

4 State estimation

$$\begin{cases} \mathbf{x}(k+1) & = & \mathbf{f}_k(\mathbf{x}(k), \mathbf{n}\,(k)) \\ \mathbf{y}(k) & = & \mathbf{g}_k(\mathbf{x}(k)), \end{cases}$$

with $\mathbf{n}\,(k) \in \mathbb{N}\,(k)$ and $\mathbf{y}\,(k) \in \mathbb{Y}\,(k)$.

Without outliers

$$\mathbb{X}(k+\mathbf{1}) = \mathbf{f}_k\left(\mathbb{X}(k)\cap \mathbf{g}_k^{-1}\left(\mathbb{Y}(k)\right),\quad \mathbb{N}\left(k\right)\right).$$

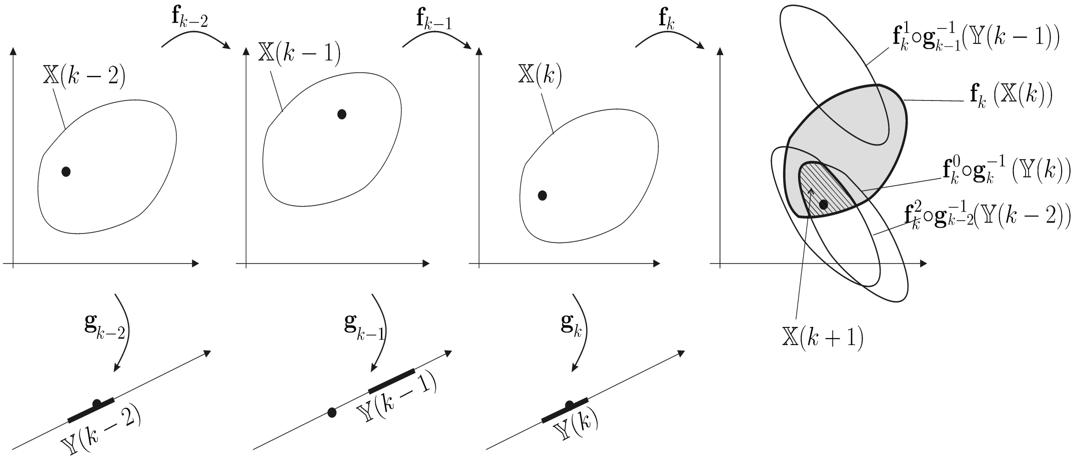
Define

$$\begin{cases} \mathbf{f}_{k:k}(\mathbb{X}) & \overset{\text{def}}{=} \mathbb{X} \\ \mathbf{f}_{k_1:k_2+1}(\mathbb{X}) & \overset{\text{def}}{=} \mathbf{f}_{k_2}(\mathbf{f}_{k_1:k_2}(\mathbb{X}), \mathbb{N}(k_2)), \ k_1 \leq k_2. \end{cases}$$

The set $\mathbf{f}_{k_1:k_2}(\mathbb{X})$ represents the set of all $\mathbf{x}(k_2)$, consistent with $\mathbf{x}(k_1) \in \mathbb{X}$.

Consider the set state estimator

$$\begin{cases} \mathbb{X}(k) = \mathbf{f}_{0:k}(\mathbb{X}(0)) \quad \text{if } k < m, \text{ (initialization step)} \\ \mathbb{X}(k) = \mathbf{f}_{k-m:k}(\mathbb{X}(k-m)) \cap \\ \qquad \bigcap_{i \in \{1,\ldots,m\}}^{\{q\}} \mathbf{f}_{k-i:k} \circ \mathbf{g}_{k-i}^{-1}(\mathbb{Y}(k-i)) \quad \text{if } k \geq m \end{cases}$$



$\mathbf{f}_{k-2}$
$\mathbf{f}_{k-1}$
$\mathbf{f}_k$
$\mathbb{X}(k-2)$
$\mathbb{X}(k-1)$
$\mathbb{X}(k)$
$\mathbf{f}_k^1 \circ \mathbf{g}_{k-1}^{-1}(\mathbb{Y}(k-1))$
$\mathbf{f}_k\,(\mathbb{X}(k))$
$\mathbf{f}_k^0 \circ \mathbf{g}_k^{-1}\,(\mathbb{Y}(k))$
$\mathbf{f}_k^2 \circ \mathbf{g}_{k-2}^{-1}(\mathbb{Y}(k-2))$
$\mathbb{X}(k+1)$
$\mathbf{g}_{k-2}$
$\mathbf{g}_{k-1}$
$\mathbf{g}_k$
$\mathbb{Y}(k-2)$
$\mathbb{Y}(k-1)$
$\mathbb{Y}(k)$

We assume that all errors are time independent.

If (i) within any time window of length m we have less than q outliers and that (ii) $\mathbb{X}(0)$ contains $\mathbf{x}(0)$, then $\mathbb{X}(k)$ encloses $\mathbf{x}(k)$.

What is the probability of this assumption ?

Theorem. Consider the sequence of sets $\mathbb{X}(0), \mathbb{X}(1), \ldots$ built by the set observer. We have

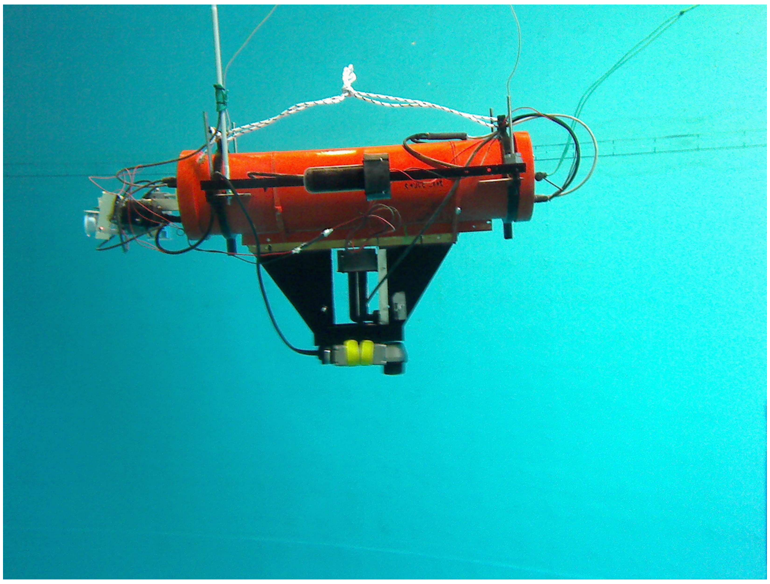
$$\Pr\left(\mathbf{x}\left(k\right) \in \mathbb{X}(k)\right) \geq \alpha \; * \; \Pr\left(\mathbf{x}\left(k-1\right) \in \mathbb{X}(k-1)\right)$$

where

$$\alpha = \sqrt[m]{\sum_{i=m-q}^{m} \frac{m!\; \pi_y^i.\left(1-\pi_y\right)^{m-i}}{i!\left(m-i\right)!}}$$

with an equality if $\mathbb{N}\left(k\right)$ are singletons.

5 Application to localization



Sauc'isse robot inside a swimming pool

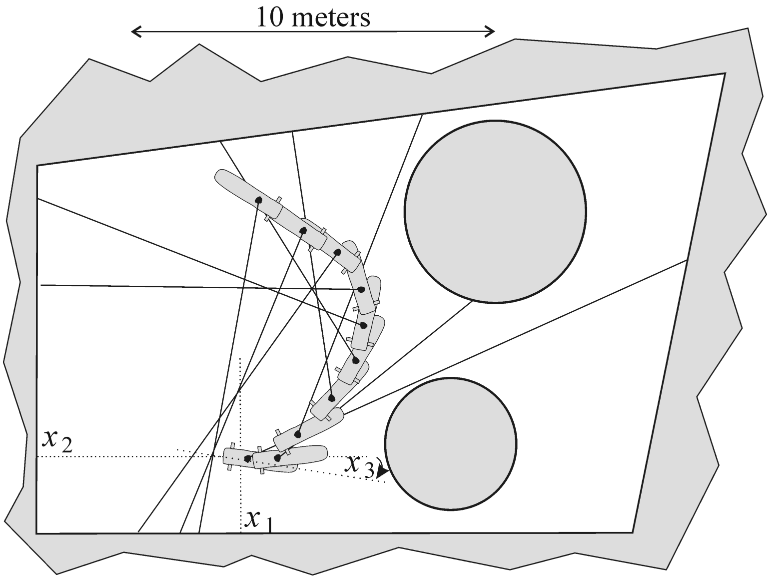
The robot evolution is described by

$$\begin{cases} \dot{x}_1 & = & x_4 \cos x_3 \\ \dot{x}_2 & = & x_4 \sin x_3 \\ \dot{x}_3 & = & u_2 - u_1 \\ \dot{x}_4 & = & u_1 + u_2 - x_4, \end{cases}$$

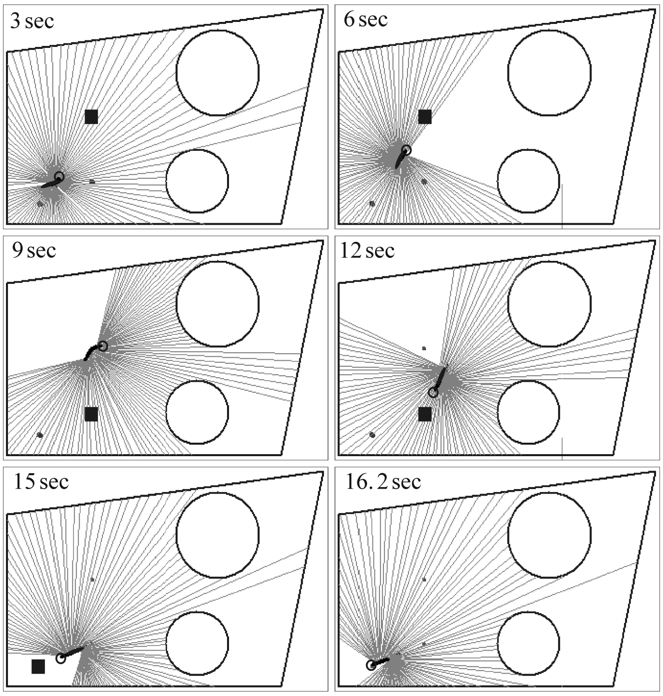
where x_1, x_2 are the coordinates of the robot center, x_3 is its orientation and x_4 is its speed. The inputs u_1 and u_2 are the accelerations provided by the propellers.

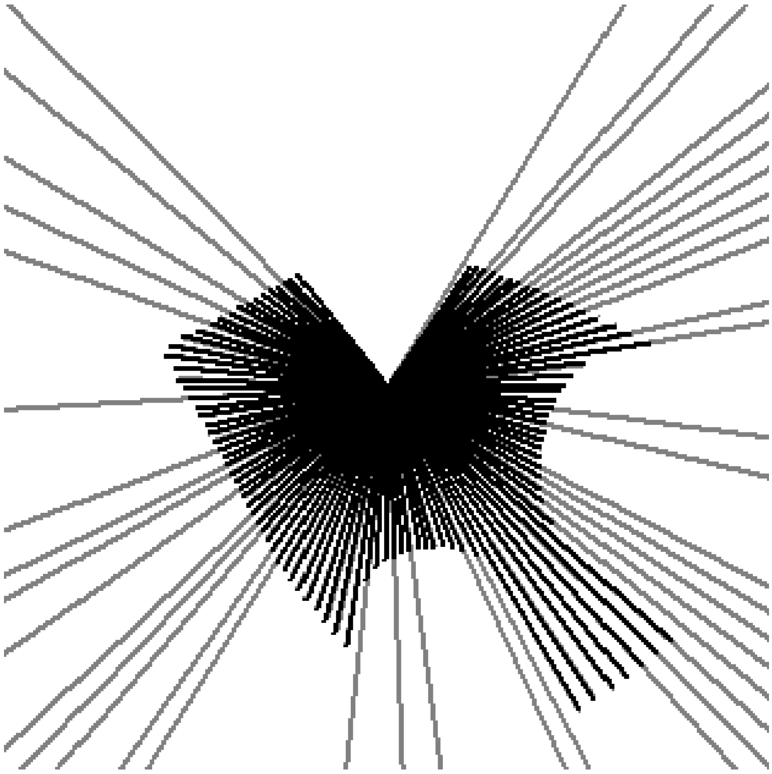
The system can be discretized by $\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k)$, where,

$$\mathbf{f}_k \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 + \delta . x_4 . \cos(x_3) \\ x_2 + \delta . x_4 . \sin(x_3) \\ x_3 + \delta . (u_2(k) - u_1(k)) \\ x_4 + \delta . (u_1(k) + u_2(k) - x_4) \end{pmatrix}$$

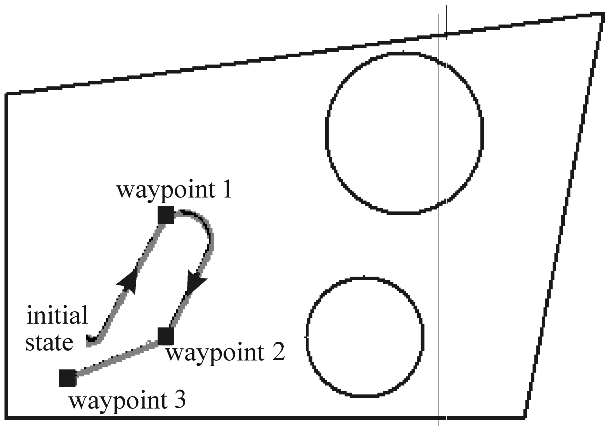


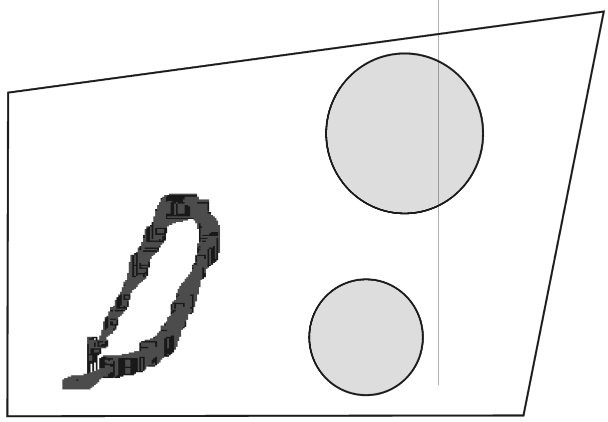
Underwater robot moving inside a pool


3 sec
6 sec
9 sec
12 sec
15 sec
16.2 sec



Emission diagram at time $t = 9\,\mathsf{sec}$


waypoint 1
initial
state
waypoint 2
waypoint 3



t(sec)	Pr $(\mathbf{x} \in \mathbb{X})$	Outliers
3.0	$\geq$ 0.965	58
6.0	$\geq$ 0.932	50
9.0	$\geq$ 0.899	42
12.0	$\geq$ 0.869	51
15.0	$\geq$ 0.838	51
16.2	$\geq$ 0.827	49