Robust centralized and distributed bounded-error parameter estimation

Jean-Benoist Léger, Michel Kieffer

L2S - CNRS - SUPELEC - Univ Paris-Sud

2 february 2012

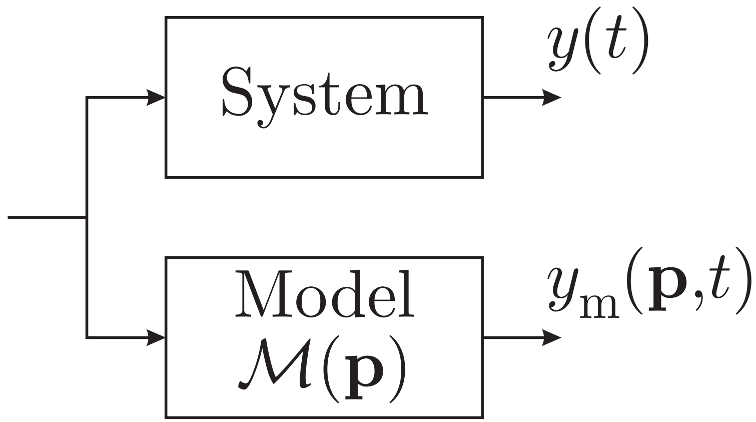
GT MEA - ENSTA ParisTech

First Part

—

Bounded-error estimation

Parameter estimation



$\mathbf{y}$: vector of experimental data

$\mathbf{p}$: vector of unknown, constant parameters

$\mathbf{y}_m(\mathbf{p})$: vector of model output

Parameter estimation :

Determination of $\widehat{\mathbf{p}}$ from $\mathbf{y}$.

Problem formulation

1. Minimisation of a cost function, *e.g.*,

$$\widehat{\mathbf{p}} = \arg\min_{\mathbf{p}} j_{\mathsf{LS}}(\mathbf{p}) = (\mathbf{y} - \mathbf{y}_{\mathsf{m}}(\mathbf{p}))^{\mathsf{T}} (\mathbf{y} - \mathbf{y}_{\mathsf{m}}(\mathbf{p}))$$

   or

$$\widehat{\mathbf{p}} = \arg\max_{\mathbf{p}} f_{P|Y}(\mathbf{p} \mid \mathbf{y})$$

   – Local techniques : Gauss-Newton, Levenberg-Marquardt. . .

   – Random search : simulated annealing, genetic algorithms. . .

   – Global guaranteed techniques : Hansen's algorithm

2. Parameter bounding

Parameter bounding

Experimental data : $y(t_i)$,

$t_i, i = 1 \ldots, N$, known measurement times

$[\varepsilon_i] = [\underline{\varepsilon}_i, \overline{\varepsilon}_i], i = 1, \ldots, N$, known acceptable errors

$\mathbf{p} \in \mathcal{P}_0$ deemed acceptable if for all $i = 1, \ldots, N$,

$$\underline{\varepsilon}_i \leqslant y(t_i) - y_{\text{m}}(\mathbf{p}, t_i) \leqslant \overline{\varepsilon}_i.$$

$\Longrightarrow$ Bounded-error parameter estimation :

Characterize $\mathcal{S} = \{\mathbf{p} \in \mathcal{P}_0 \mid y(t_i) - y_{\text{m}}(\mathbf{p}, t_i) \in [\underline{\varepsilon}_i, \overline{\varepsilon}_i], \ i = 1, \ldots, N\}$

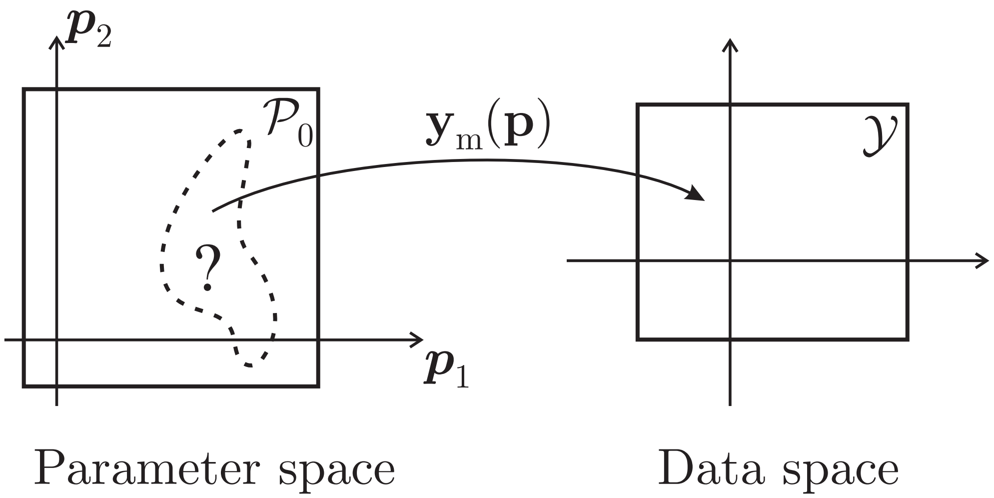
Sivia

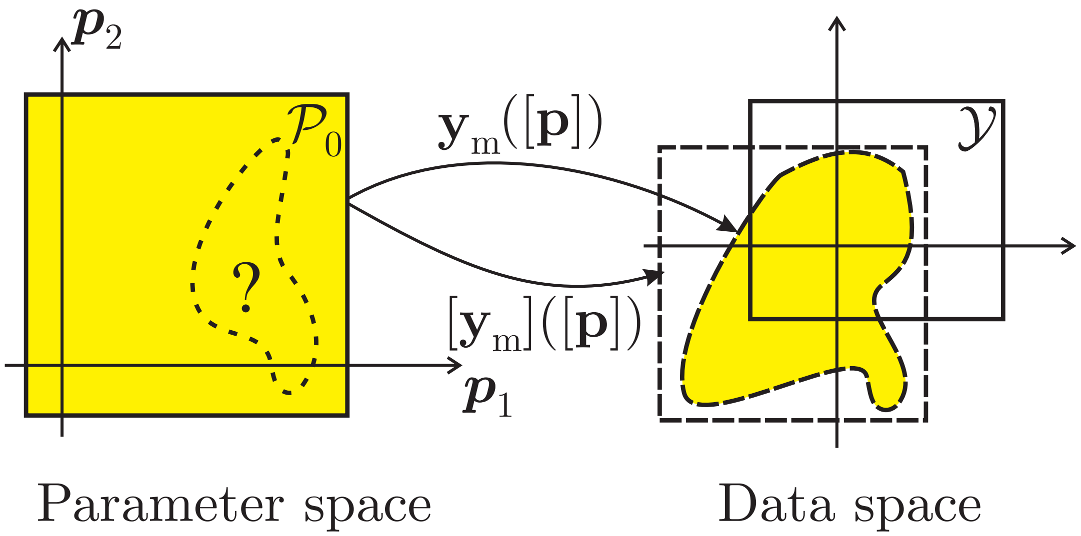
Set to be characterized

$$\mathcal{S} = \{\mathbf{p} \in \mathcal{P}_0 \mid y(t_i) - y_\mathsf{m}(\mathbf{p}, t_i) \in [\underline{\varepsilon}_i, \overline{\varepsilon}_i], i = 1, \ldots, N\}$$
$$= \{\mathbf{p} \in \mathcal{P}_0 \mid \mathbf{y}_\mathsf{m}(\mathbf{p}) \subset \mathcal{Y}\},$$

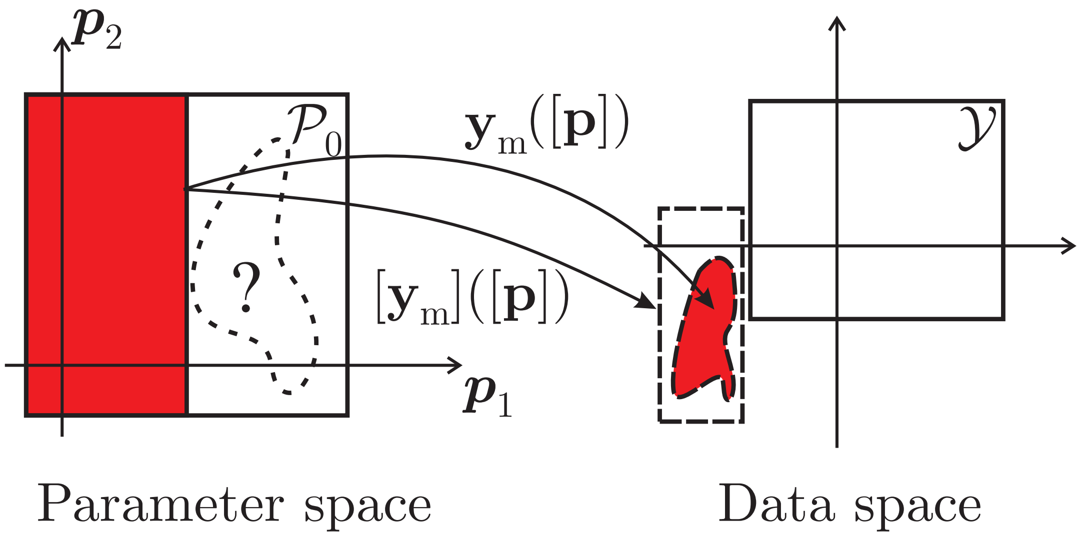
with

$$\mathcal{Y} = [y(t_1) - \overline{\varepsilon}_1, y(t_1) - \underline{\varepsilon}_1] \times \cdots \times [y(t_N) - \overline{\varepsilon}_N, y(t_N) - \underline{\varepsilon}_N]$$

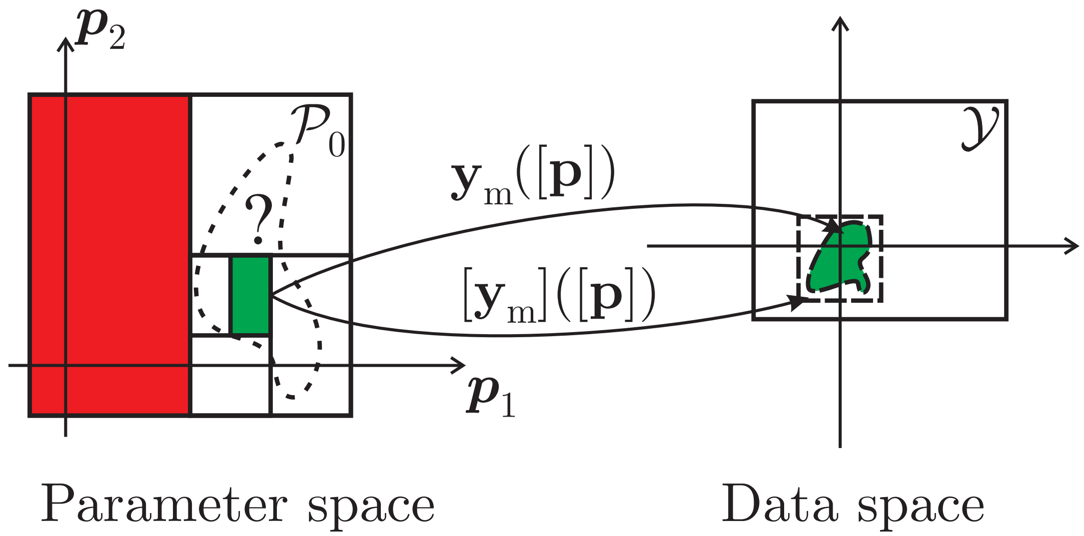




Yellow box is undetermined



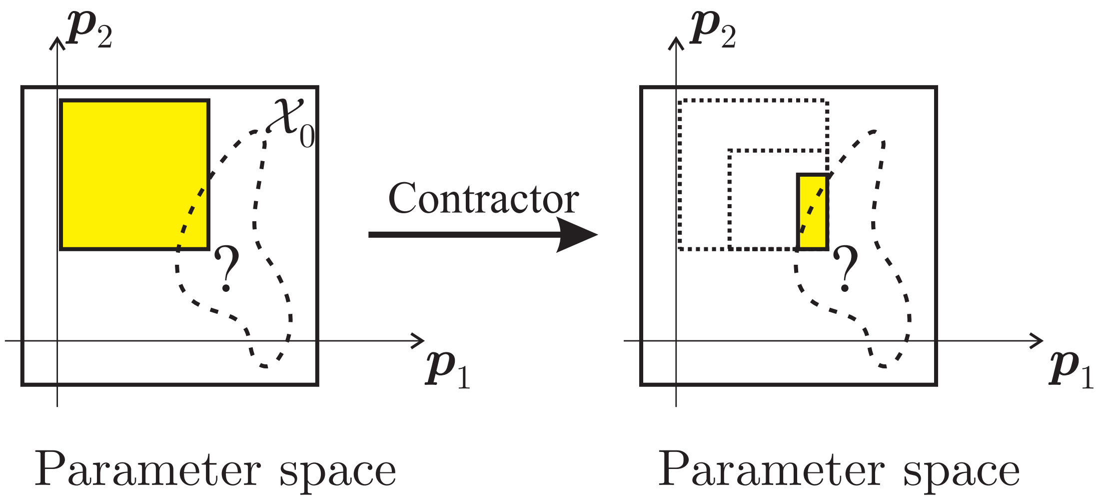
Red box proven to be outside $\mathcal{S}$



Green box proven to be included in $\mathcal{S}$

Sivia with contractors

Reduce the size of undetermined boxes without any bissection



Contractors (Jaulin *et al*, 2001) based on

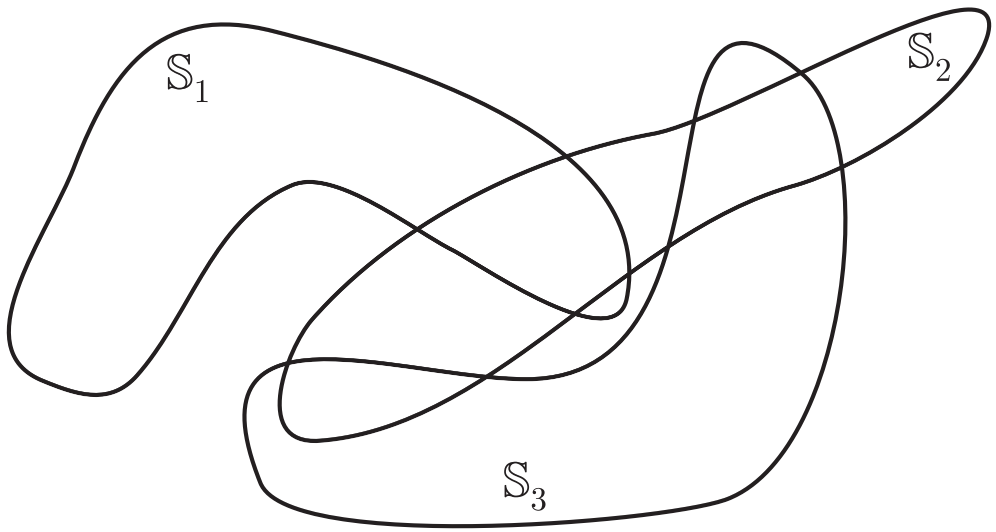
– interval constraint propagation

– linear programming

– parallel linearization

– ...

Robust parameter estimation



Interpretation of empty solution set

$$\mathbb{S} = \bigcap_{\ell=1\ldots N} \mathbb{S}_\ell = \emptyset.$$

Hypotheses on model or noise not satisfied

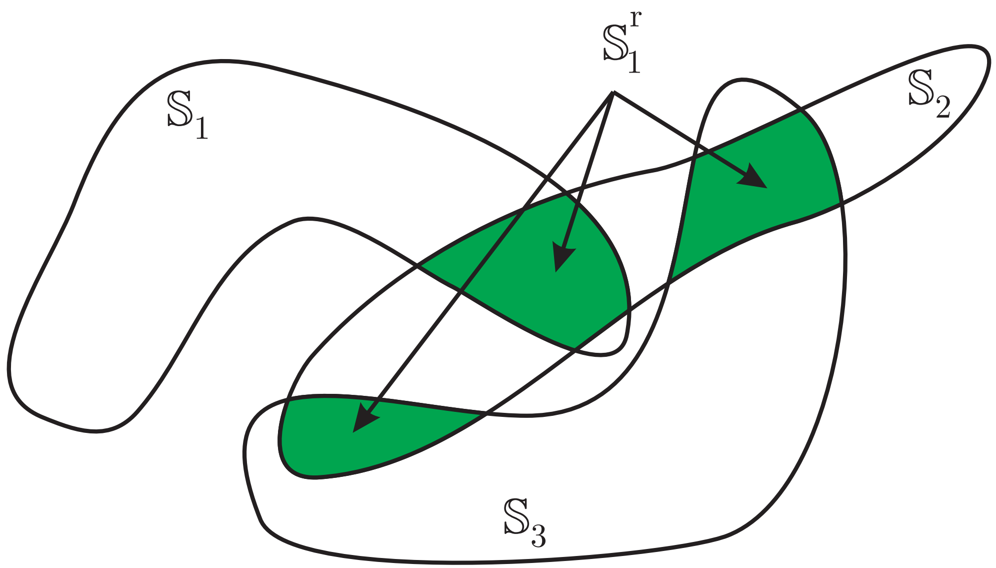
(easy detection)

Situation very frequently encountered when considering actual measurements

$\Downarrow$

Robust parameter estimation techniques are necessary

Estimator robust to q outliers



New solution set, q outliers are tolerated

$$\mathbb{S}_q^{\mathrm{r}} = \bigcup_{1 \leqslant \ell_1 < \cdots < \ell_q \leqslant N} \; \bigcap_{\ell \neq \ell_1, \ldots, \ell \neq \ell_q} \mathbb{S}_\ell .$$

$$\Leftrightarrow$$

Union of intersections of $N - q$ sets among N

Interval analysis [JKDW01] allows to get

$$\underline{\mathbb{S}}_q^{\mathrm{r}} \subset \mathbb{S}_q^{\mathrm{r}} \subset \overline{\mathbb{S}}_q^{\mathrm{r}},$$

without combinatorial techniques.

Consider

$$t_\ell(\mathbf{p}) = \begin{cases} 1 & \text{if } y_\ell^{\mathsf{m}}(\mathbf{p}) \in [y_\ell] \\ 0 & \text{else,} \end{cases}$$

and

$$t(\mathbf{p}) = \sum_{\ell=1}^{N} t_\ell(\mathbf{p}).$$

Then

$$\begin{aligned} \mathbb{S}_q^{\mathsf{r}} &= \bigcup_{1\leqslant \ell_1 < \cdots < \ell_q \leqslant N} \ \bigcap_{\ell\neq\ell_1,\ldots,\ell\neq\ell_q} \mathbb{S}_\ell \\ &= \{\mathbf{p}\in\mathbb{P} \mid t(\mathbf{p}) \geqslant N - q\} \end{aligned}$$

and

– there is no combinatorial any more,

– it is not necessary to choose a priori the outliers.

Inclusion functions

To apply SIVIA, for example, an inclusion function for $t_\ell(\mathbf{p})$ is needed...

$$[t_\ell]([\mathbf{p}]) = \begin{cases} 1 & \text{if } [y_\ell^{\mathsf{m}}]([\mathbf{p}]) \subset [y_\ell] \\ 0 & \text{if } [y_\ell^{\mathsf{m}}]([\mathbf{p}]) \cap [y_\ell] = \emptyset \\ [0,1] & \text{else} \end{cases}$$

... which itself requires an inclusion function for $y_\ell^{\mathsf{m}}(\mathbf{p})$.

Guaranteed robust parameter estimation :

no $\mathbf{p}$ consistent with more than $N - q$ data and bounds is missed

If $\mathbf{p}^*$ exists, and if there are not more than q outliers

$$\Downarrow$$

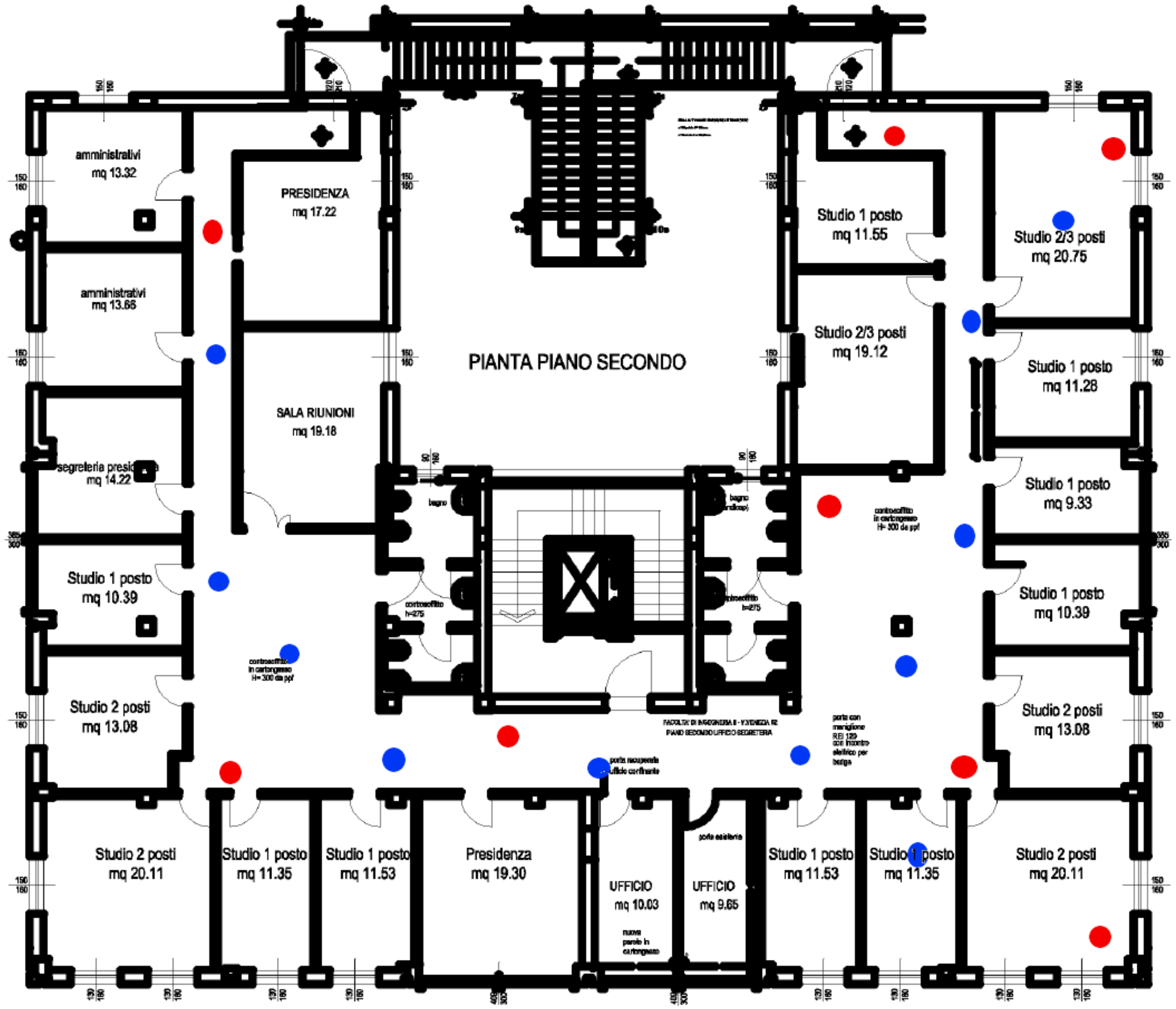
$\mathbb{S}_q^{\mathsf{r}}$ guaranteed to contain $\mathbf{p}^*$.

Example : Localization using UWB signals

Using the Static WP RB database

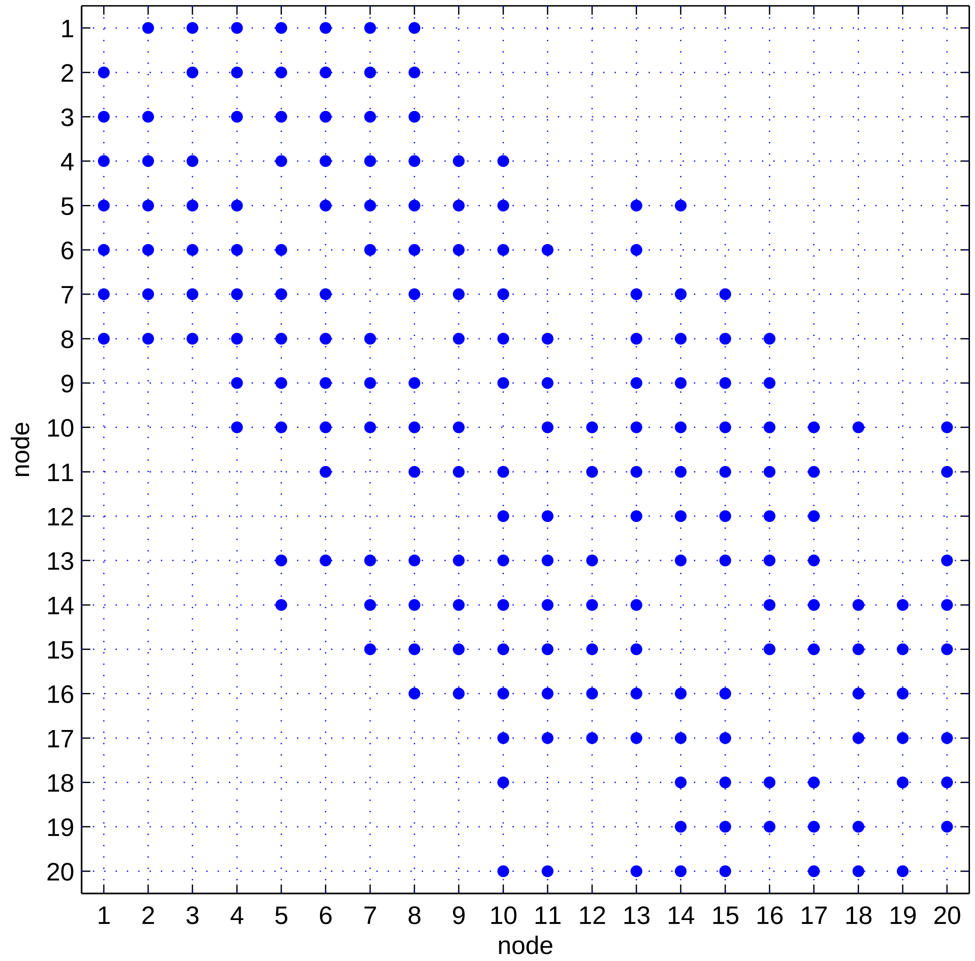
UWB device : PulsOn220

Environment



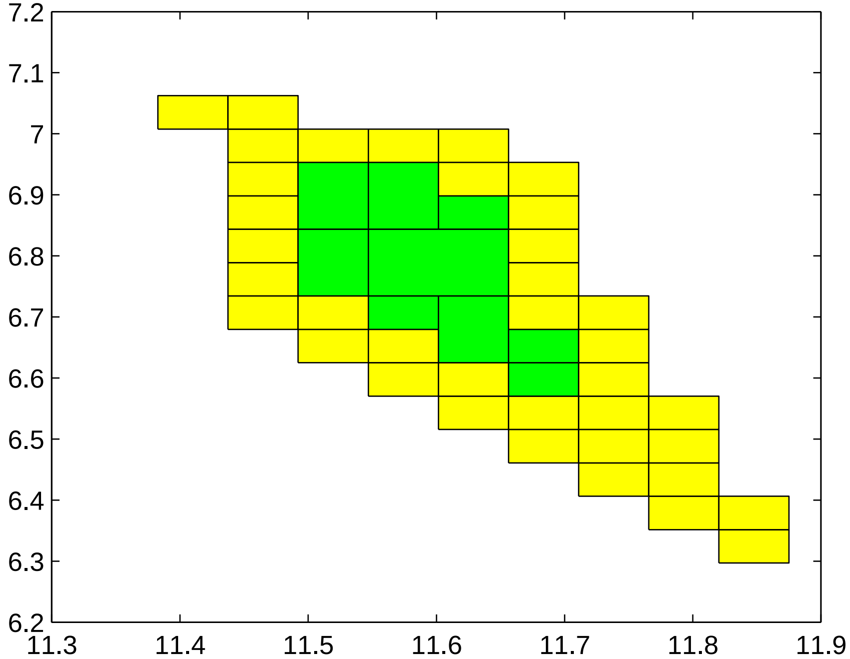
UWB Measurement campaign performed at ISMB, Italy

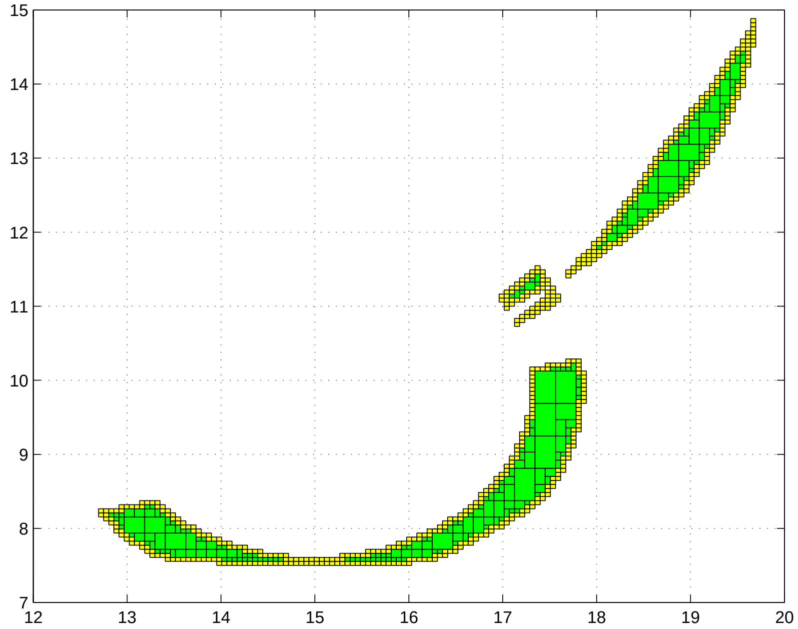
Connectivity



Static scenario, 1000 measurements.

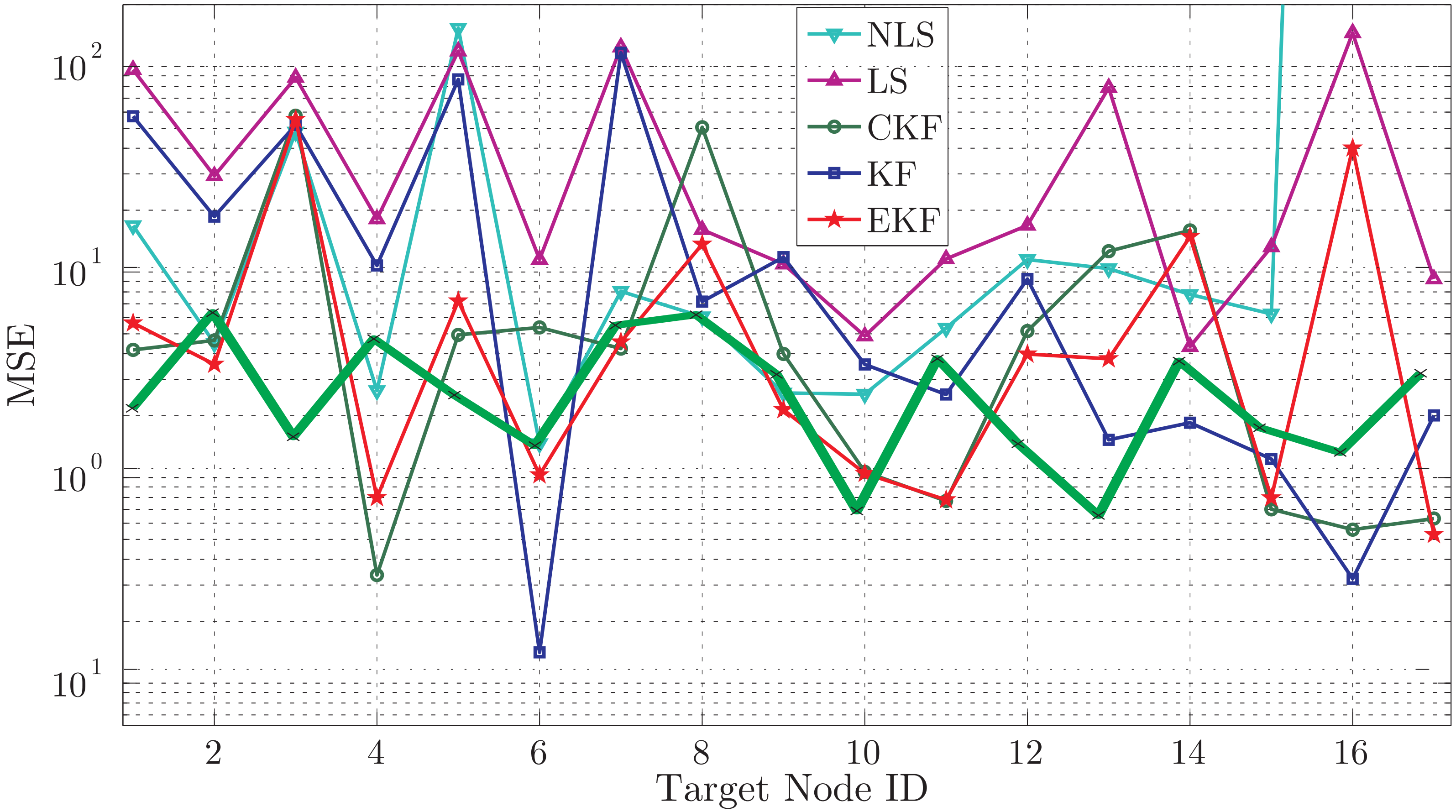
Results





Two typical solution sets

Results



MSE evaluated between middle of bounding box and true location

Robust bounded-error technique not always the best, never the worst...

Difficulties in Robust Distributed parameter estimation

Each sensor should evaluate $\mathbb{S}_n^{\mathrm{r}}$

Main difficulties :

– Direct extension of centralized algorithm not possible

$\hookrightarrow$ involves all measurements.

– Local estimates are exchanged, information progressively available

Assumptions :

– Network is entirely connected

– Sensors exchange estimates

– Number of tolerated outliers fixed *a priori*.

Idealized robust distributed approach

Consider $J \subset [\![1, N]\!]$, and define

$$\mathbb{S}_q^J = \bigcup_{I \subset J, \mathrm{card}(I) = \mathrm{card}(J) - q} \left(\bigcap_{i \in I} \mathbb{S}_i \right), \tag{1}$$

with

$$\mathbb{S}_i = \left\{ \mathbf{p} \in \mathbb{P} | y_i^{\mathsf{m}}(\mathbf{p}) \in [y_i] \right\}.$$

Properties

1. $\mathbb{S}_q^{\mathsf{r}} = \mathbb{S}_q^{[\![1,N]\!]}$
2. $\forall J_1 \subset J_2 \subset [\![1, N]\!]$, one has $\mathbb{S}_q^{J_1} \supset \mathbb{S}_q^{J_2} \supset \mathbb{S}_q^{\mathsf{r}}$.

Combining robust estimates

Assume that Sensor 1

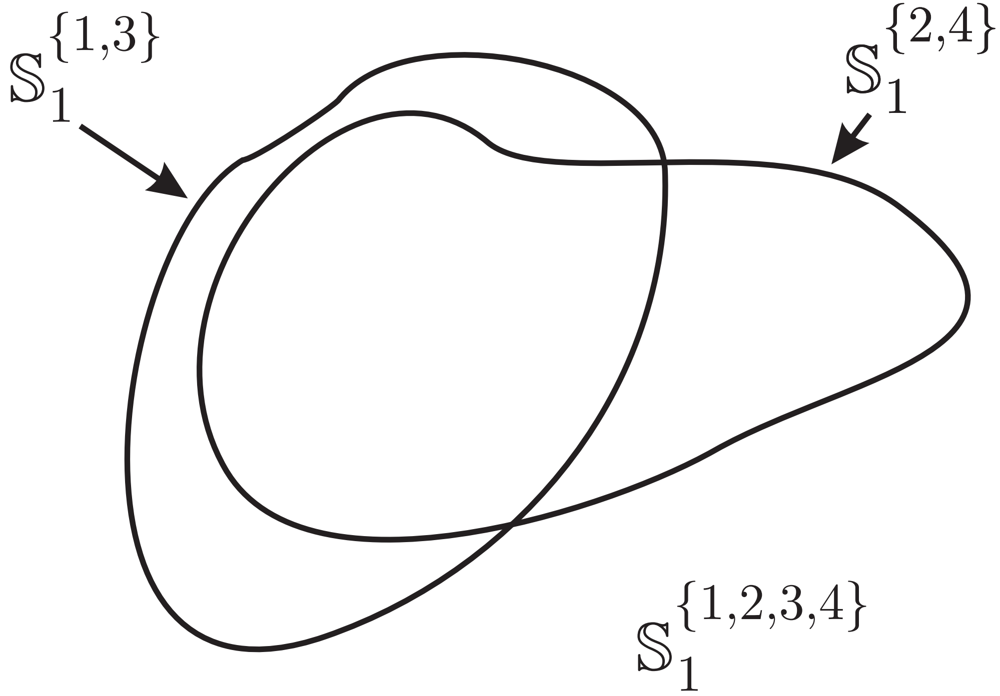
– has evaluated $\mathbb{S}_q^{J_1}$

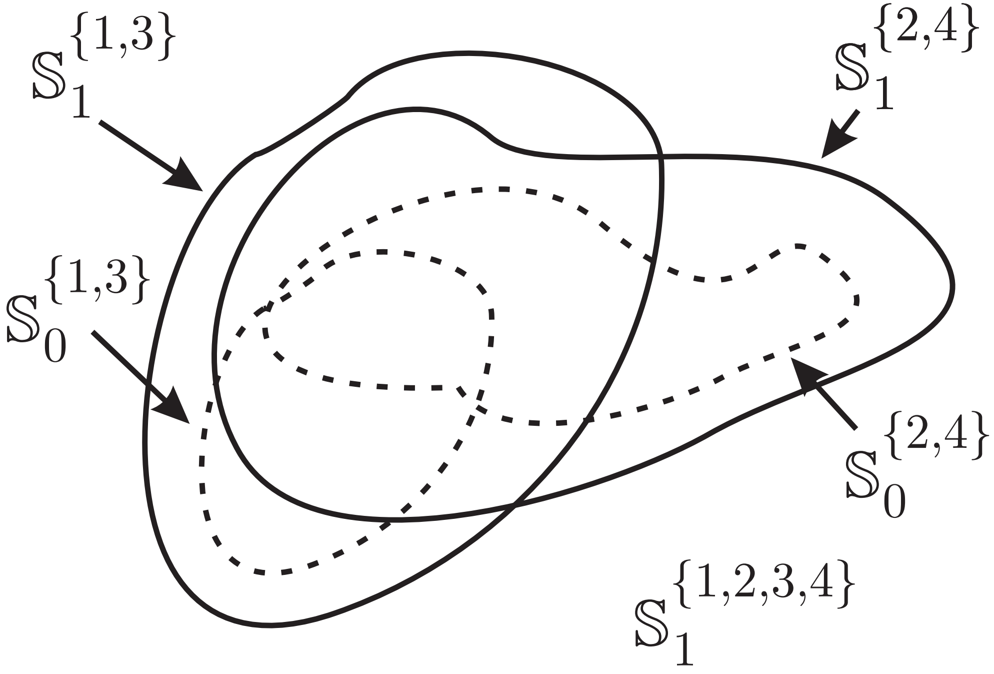
– has received $\mathbb{S}_q^{J_2}$ from Sensor 2.

To get a better approximation of $\mathbb{S}_q^{\mathrm{r}}$, evaluate $\mathbb{S}_q^{J_1 \cup J_2}$.

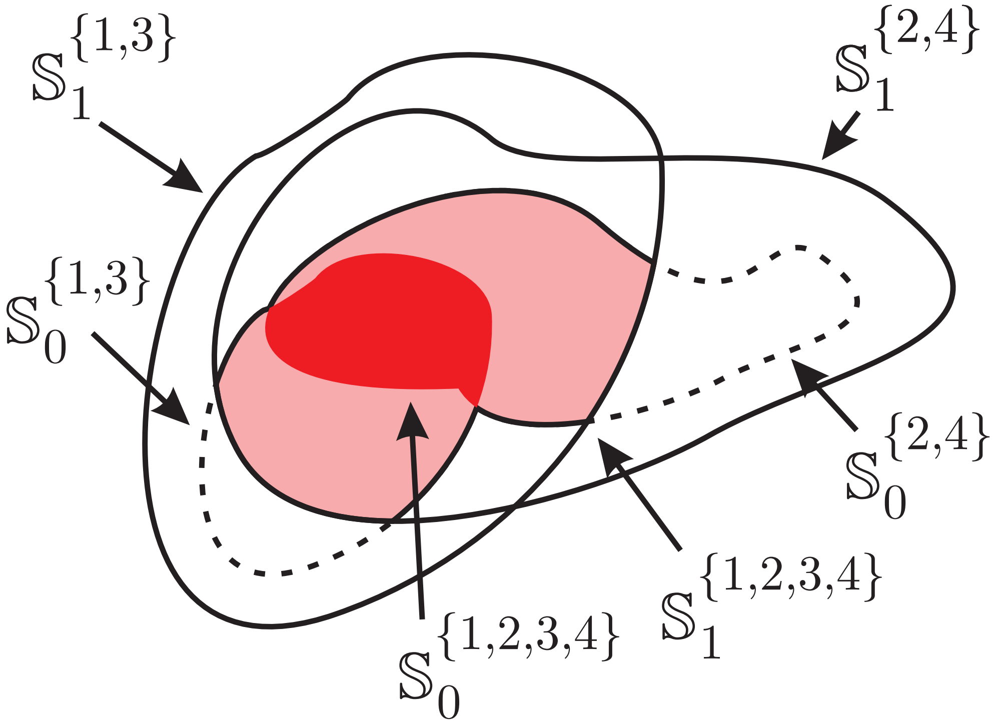
– If $J_1 \cap J_2 \neq \emptyset$, no simple relation between $\mathbb{S}_q^{J_1} \cap \mathbb{S}_q^{J_2}$, $\mathbb{S}_q^{J_1} \cup \mathbb{S}_q^{J_2}$ and $\mathbb{S}_q^{J_1 \cup J_2}$.

– If $J_1 \cap J_2 = \emptyset$, one has $\mathbb{S}_q^{J_1 \cup J_2} \subset \mathbb{S}_q^{J_1} \cap \mathbb{S}_q^{J_2}$, but in general $\mathbb{S}_q^{J_1 \cup J_2} \neq \mathbb{S}_q^{J_1} \cap \mathbb{S}_q^{J_2}$.


$\mathbb{S}_1^{\{1,3\}}$
$\mathbb{S}_1^{\{2,4\}}$
$\mathbb{S}_1^{\{1,2,3,4\}}$



$\mathbb{S}_1^{\{1,3\}}$
$\mathbb{S}_1^{\{2,4\}}$
$\mathbb{S}_0^{\{1,3\}}$
$\mathbb{S}_0^{\{2,4\}}$
$\mathbb{S}_1^{\{1,2,3,4\}}$



To compute $\mathbb{S}_q^{J_1 \cup J_2}$, all $\mathbb{S}_0^{J_1}, \ldots, \mathbb{S}_q^{J_1}$ and $\mathbb{S}_0^{J_2}, \ldots, \mathbb{S}_q^{J_2}$ are needed

$$\mathbb{S}_{q'}^{J_1 \cup J_2} = \bigcup_{q_1+q_2=q'} \mathbb{S}_{q_1}^{J_1} \cap \mathbb{S}_{q_2}^{J_2}, \; q' = 0, \ldots, q.$$

Each sensor has to transmit $\Gamma_q^J = \left(\mathbb{S}_0^J, \mathbb{S}_1^J, \cdots, \mathbb{S}_q^J\right)$

– Initially sensor i processes its own measurement,

$\hookrightarrow \mathbb{S}_i$,

$\hookrightarrow \Gamma_q^{\{i\}} = (\mathbb{S}_i, \mathbb{P}, \cdots, \mathbb{P})$.

– Then, it broadcasts $\Gamma_q^{\{i\}}$ and receives similar structures.

– The sensor i is able to improve its estimates if $J_1 \cap J_2 = \emptyset$,

$$\forall q' \in [\![0, q]\!] \qquad \mathbb{S}_{q'}^{J_1 \cup J_2} = \bigcup_{q_1 + q_2 = q'} \mathbb{S}_{q_1}^{J_1} \cap \mathbb{S}_{q_2}^{J_2}.$$

– Then sensor broadcasts the best Γ_q^J (with the largest $\mathrm{card}(J)$).

– and so on until convergence.

Estimation and communication performed until all sensors have obtained $\mathbb{S}_q^{[\![1,N]\!]} = \mathbb{S}_q^{\mathsf{r}}$

$\hookrightarrow$ occurs in finite time when the network is connected.

If computations stopped before convergence, each sensor of the network has an outer-approximation of $\mathbb{S}_q^{\mathsf{r}}$, which improves when more data are exchanged.

Implementation issues

Representation of sets

Sets such as $\mathbb{S}_q^J$ have to be transmitted : not possible in general.

$$\Downarrow$$

Wrappers $\overline{\mathbb{S}}_q^J$ for $\mathbb{S}_q^J$ are considered

$$\Downarrow$$

Allows to get $\overline{\mathbb{S}}_q^{\mathrm{r}}$.

Wrappers : ellipsoids, polytopes, or subpavings [JKDW01].

$\overline{\Gamma}_q^J$ represented by a regular paving of $\mathbb{P}$

– Paving described by a binary tree.
– Each leaf labeled with ℓ (corresponding box is a subset of $\overline{\mathbb{S}}_\ell^J$).

Computation for each sensor by only unions and intersections of subpavings (of trees).

Transmission protocol

Each node :

– stores intermediate results with $\overline{\Gamma}_q^J = \left(\overline{\mathbb{S}}_0^J, \overline{\mathbb{S}}_1^J, \cdots, \overline{\mathbb{S}}_q^J\right)$ in place of Γ_q^J

– transmits $\overline{\Gamma}_q^J$ to neighboring nodes.

Optimized Link State Routing Protocol [CJL$^+$01], used to satisfy the combination constraint more easily.

$\hookrightarrow$ Nodes of the network use *multipoint distribution relays* (MPR) for transmission.

$\hookrightarrow$ For a given node, only a subset of its neighbors relay its message.

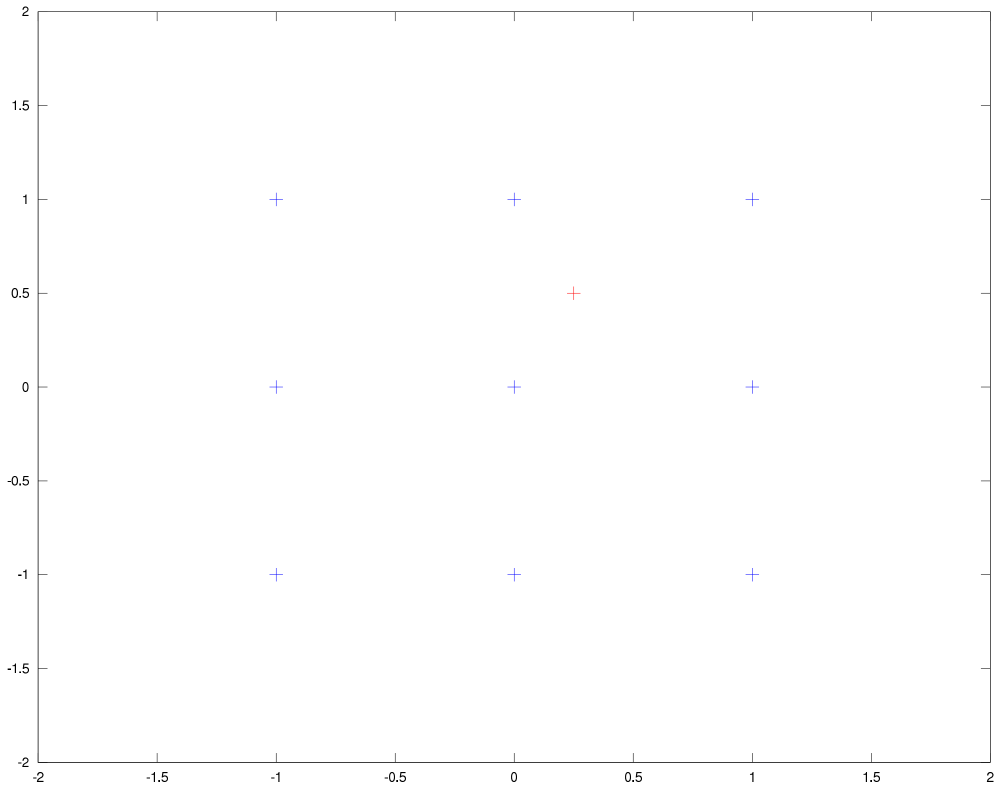
$\hookrightarrow$ Selection of MPRs adaptive and done in real time.

Dynamic hierarchical structure :

a sensor selects his MPRs and sends its sets $\overline{\Gamma}$ only to its MPRs.

Simulations

Single source localization in 2D-environment with NWS.



Considered network of 9 sensors (blue) and one source (red)

Each sensor measures the power it receives from the source.

All measurement errors are bounded :

– received power y_i by the i-th sensor,
– noise-free measurement belongs to $[y_i] = [y_i/w, y_i w]$ with $w = 1.7$.

Two outliers are introduced by hand (Sensors 4 and 6).

Location of the source $\mathbf{p}^* = (\theta_1, \theta_2)$ has to be estimated.

Measurement model for the i-th sensor

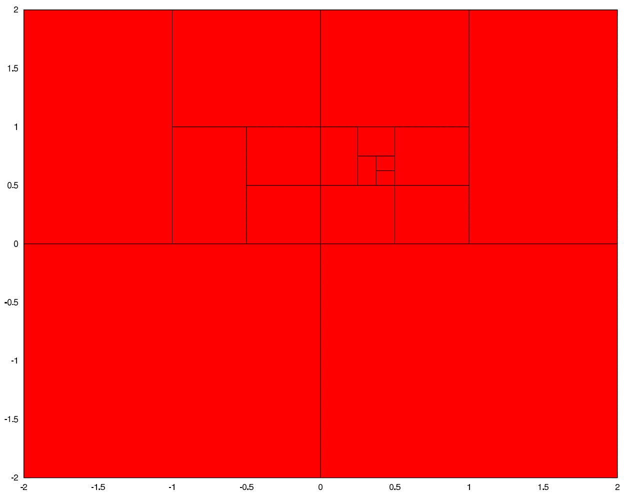
$$y_i^{\mathsf{m}}(\theta_1, \theta_2) = \frac{P_0}{\left(\sqrt{(\theta_1 - \theta_{1i})^2 + (\theta_2 - \theta_{2i})^2}\right)^{\eta}} \tag{2}$$

where $(\theta_{1i}, \theta_{2i})$ is location of Sensor i.

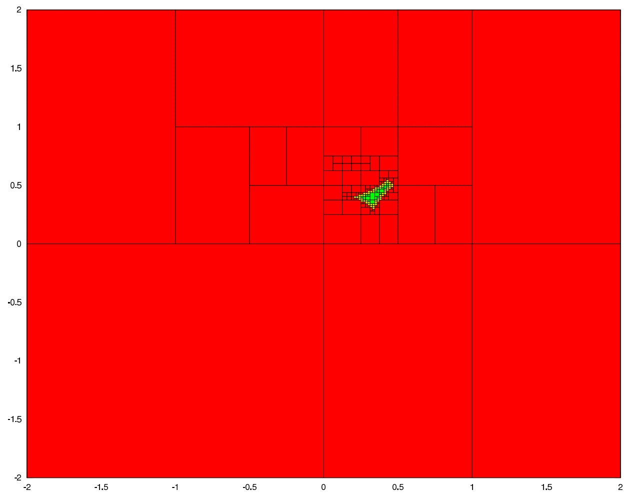
$P_0 = 1$ and $\eta = 2$ are assumed to be known.

Centralized estimation

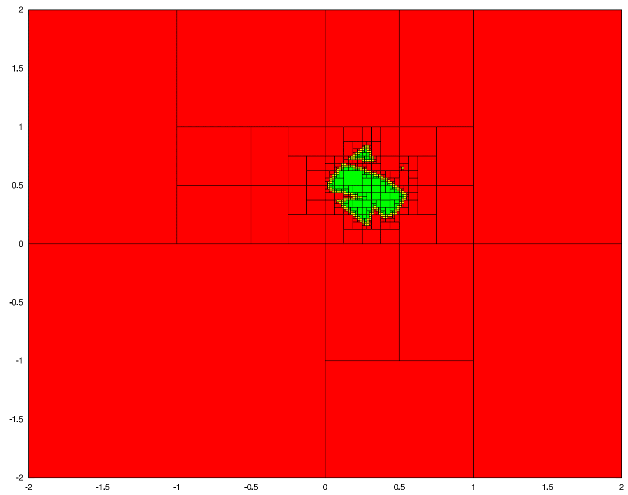
Number of tolerated outliers $q \in \{0, 1, 2\}$



$q = 0$



$q = 1$

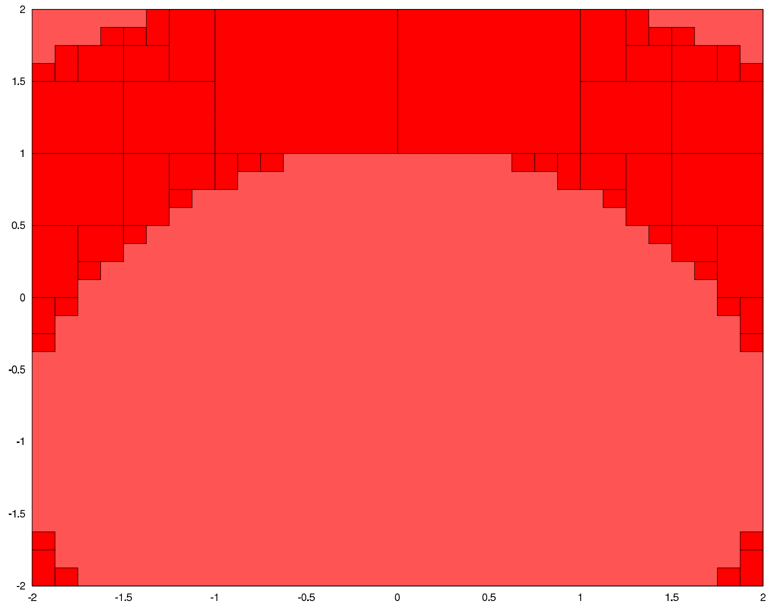


$q = 2$

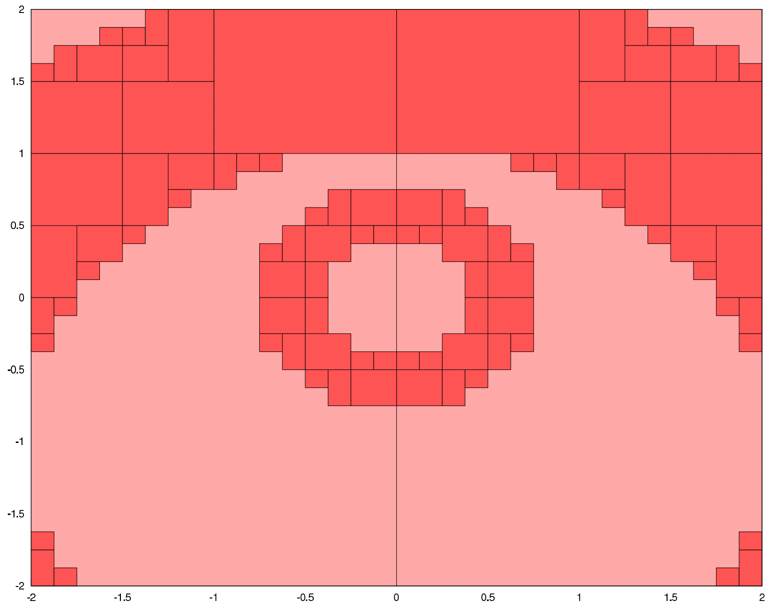
all plots are in $[-2, 2]^2$

Distributed estimation

Estimates $\overline{\Gamma}_2^J$ available at the 4-th sensor using the proposed distributed estimator. Number of tolerated outliers $q = 2$

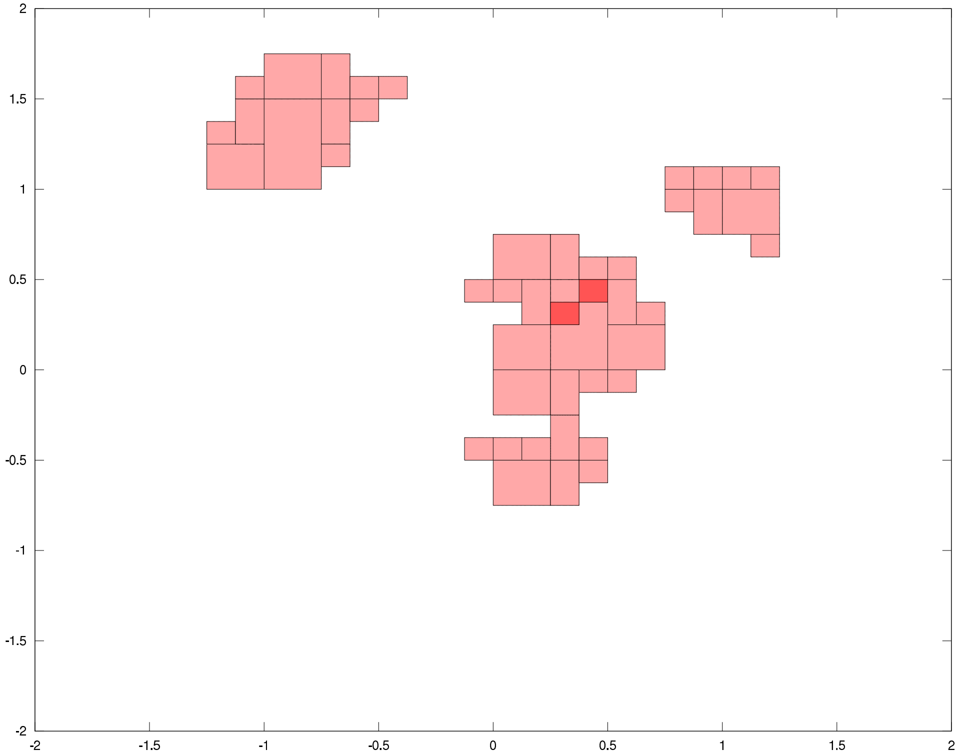


$\overline{\Gamma}_2^{\{4\}}$

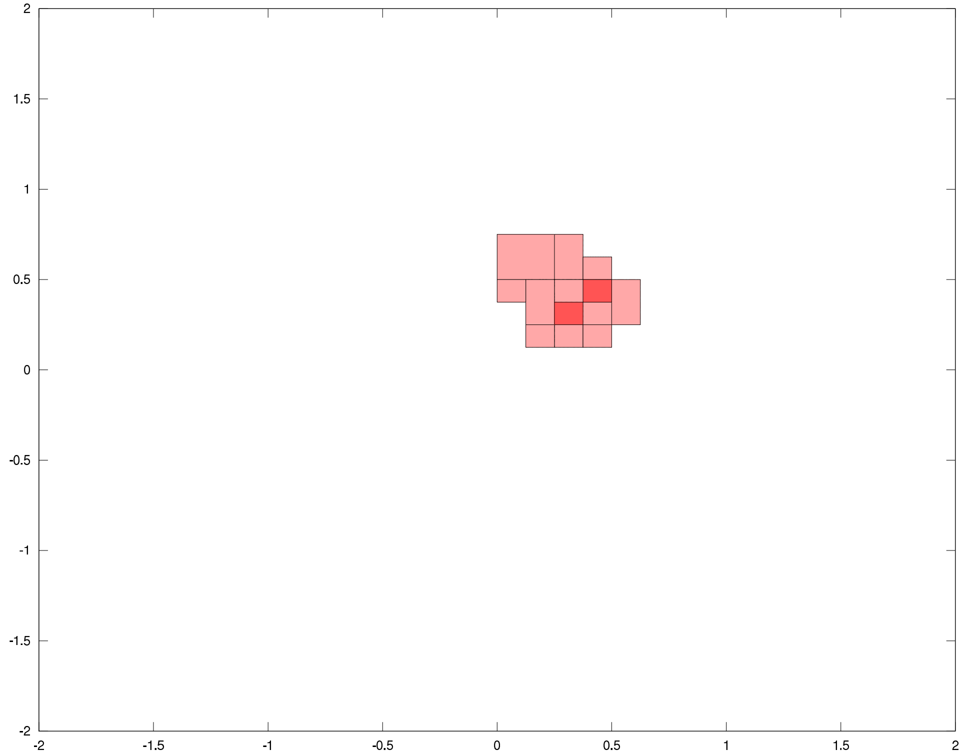


$\overline{\Gamma}_2^{\{4,5\}}$

With $\overline{\mathbb{S}}_0^J$ in *dark-red*, $\overline{\mathbb{S}}_1^J$ in *red*, and $\overline{\mathbb{S}}_2^J$ in *light-red*.

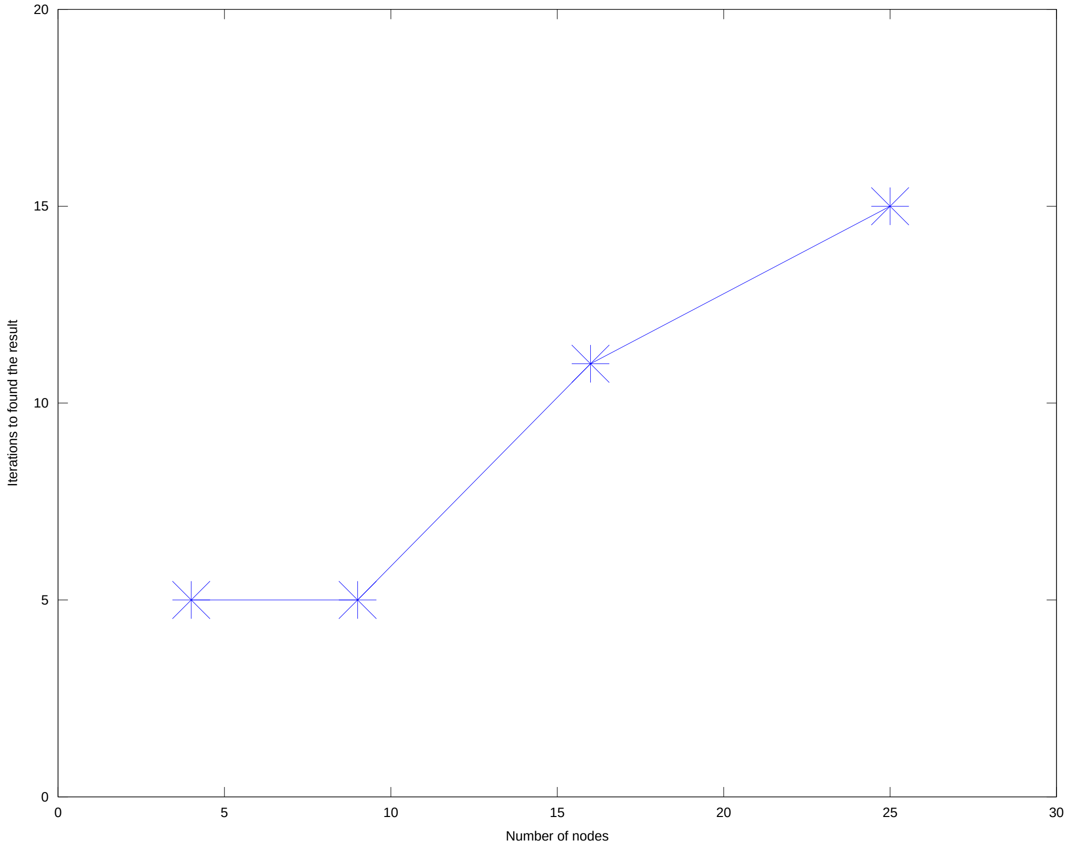


$\overline{\Gamma}_2^{\{2,4,5,6,8\}}$



$\overline{\Gamma}_2^{[\![1,N]\!]} = \Gamma_2$

$\overline{\mathbb{S}}_0^J$ in *dark-red*, $\overline{\mathbb{S}}_1^J$ in *red*, and $\overline{\mathbb{S}}_2^J$ in *light-red*



Evolution of convergence time with number of sensor

Conclusion

– Guaranteed robust bounded-error distributed estimation algorithm

$\hookrightarrow$ Robust to any number q of outliers

$\hookrightarrow$ Outer-approximation of $\mathbb{S}_q^{\mathrm{r}}$ at each iteration

– Not ,necessary to specify *a priori* the measurements which are deemed as outliers.

– Number of outliers has to be specified *a priori*.

– Complexity of the algorithm has to be evaluated more carefully.

$\hookrightarrow$ Sets $\overline{\Gamma}$ are quite complex, transmission requires some resources.

– Robust estimator by exchanging measurements within clusters, and exchange estimates between cluster heads.

Références

[CJL+01] T. Clausen, P. Jacquet, A. Laouiti, P. Muhlethaler, A. Qayyum, and L. Viennot. Optimized link state routing protocol for ad hoc networks. *Multi Topic Conference, 2001. IEEE INMIC 2001. Technology for the 21st Century. Proceedings. IEEE International*, pages 62–68, 2001.

[JKDW01] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter. *Applied Interval Analysis*. Springer-Verlag, London, 2001.