# Robust Covariance Matrix Estimation in Signal Processing

Thesis defense of Mélanie Mahot


Supervisors :

| | | |
|---|---|---|
| PASCAL Frédéric | Assistant Professor | Supervisor |
| FORSTER Philippe | Professor | Thesis Director |
| OVARLEZ Jean-Philippe | Professor | Thesis Co-Director |


The authors would like to thank the Direction Générale de l'Armement (DGA) for funding this project.

# Motivations

Many signal processing applications require the estimation of statistical parameters such as the covariance matrix of received data.

## Classically, data are considered to be Gaussian

- The Maximum Likelihood Estimator (MLE) is the Sample Covariance Matrix (SCM).
- It is easy to manage and has well-known statistical properties.

## However, studies show that :

- high resolution techniques provide data which do not have a Gaussian behavior.
- Outliers and other parasites are not been taken into account with this model.
- The SCM can give poor results as soon as data are not completely Gaussian.

- We need a more flexible and adjustable model to take into account most of the contexts.
- Unless the corresponding MLE can be obtained, we need, a robust family of estimators.

# Presentation outline

# Outline

# Statistical modelling and associated estimators in signal processing

## Classical assumption

Traditionally, data are assumed to be (circular) complex Gaussian.

### Gaussian distribution

Let $\mathbf{z}$ be a complex circular random vector of length $m$. $\mathbf{z}$ has a complex Gaussian distribution if its probability density function (pdf) can be written

$$f_c(\mathbf{z}) = \frac{1}{\pi^m |\mathbf{\Sigma}|} \exp\left(-\left(\mathbf{z}-\boldsymbol{\mu}\right)^H \mathbf{\Sigma}^{-1} (\mathbf{z}-\boldsymbol{\mu})\right) \quad (1)$$

where $\boldsymbol{\mu}$ is the statistical mean and $\mathbf{\Sigma}$ is the covariance matrix.

*This complex Gaussian distribution will be denoted* $\mathcal{CN}(\boldsymbol{\mu}, \mathbf{\Sigma})$.

# Statistical modelling and associated estimators in signal processing

## Classical assumption

The MLE of the Gaussian distribution is the SCM.

### The classical covariance matrix estimator : the SCM

Assuming that $(\mathbf{z}_1, \mathbf{z}_2, ...\mathbf{z}_N)$ is a $N$-sample of independant complex Gaussian vectors of length $m$ and zero mean, the SCM is

$$\widehat{\mathbf{\Sigma}}_{SCM} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{z}_n\mathbf{z}_n^H. \tag{2}$$

- It is easy to manage,
- and has well-known statistical properties.

# Statistical modelling and associated estimators in signal processing

## Compound-Gaussian distributions

### A better modelling of the data : Compound-Gaussian Vector (GCV)

The GCV $\mathbf{c}$ can be written :

$$\mathbf{c} = \sqrt{\tau}\mathbf{z}. \tag{3}$$

where

- $\tau$, the texture, is a scalar positive random variable.
- $\mathbf{z}$, the speckle, is a $m$-dimension random circular complex Gaussian vector with zero mean and covariance matrix $\mathbf{\Sigma}$.

The data are considered locally Gaussian, with a spatially variable power.

*This Complex Compound-Gaussian distribution will be denoted* $\mathcal{CCG}(\mathbf{0}, \mathbf{\Sigma}, \tau)$.

# Statistical modelling and associated estimators in signal processing

Compound-Gaussian distributions

The MLE depends on $\tau$ which is often unknown. Therefore, an approached MLE is mostly used.

## The Fixed Point Estimator (FPE)

Let $(\mathbf{c}_1, ..., \mathbf{c}_N)$ be a $N$-sample of GCV. The FPE is defined as the unique solution, up to a scale factor, of the equation :

$$\widehat{\mathbf{\Sigma}}_{FPE} = \frac{m}{N} \sum_{n=1}^{N} \frac{\mathbf{c}_n . \mathbf{c}_n^H}{\mathbf{c}_n^H \widehat{\mathbf{\Sigma}}_{FPE}^{-1} \mathbf{c}_n} \tag{4}$$

- Also known as Tyler's estimator.
- The existence and uniqueness of the solution has been proven (F. Pascal 2006).
- This estimator is unbiased , consistent and asymptotically complex Gaussian when data are GCV.

# Statistical modelling and associated estimators in signal processing

Complex elliptical circular distributions

## A more general model : complex elliptical distributions

Let $\mathbf{z}$ be a complex circular random vector of length $m$. $\mathbf{z}$ has a complex elliptical distribution if its pdf can be written

$$\mathbf{z} \longrightarrow f_{\mathbf{z}}(\mathbf{z}) = c_{2m,g} 2^m |\mathbf{\Sigma}|^{-1} . g\left( (\mathbf{z} - \boldsymbol{\mu})^H \mathbf{\Sigma}^{-1} (\mathbf{z} - \boldsymbol{\mu}) \right) . \tag{5}$$

where $g : [0, \infty) \to [0, \infty)$ is the density generator such as (5) defines a pdf and $c_{2m,g}$ is a normalization constant. $\boldsymbol{\mu}$ is the statistical mean and $\mathbf{\Sigma}$ the scatter matrix (proportional to the covariance matrix if it exists).

*This complex elliptical distribution will be denoted* $\mathcal{CE}(\boldsymbol{\mu}, \mathbf{\Sigma}, g)$

# Statistical modelling and associated estimators in signal processing

## Complex elliptical circular distributions

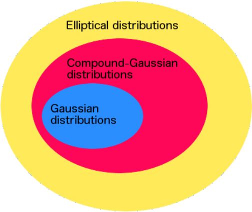
The MLE of the elliptical distributions depends on the density generator $g$.

### MLE of elliptical distributions

Let $(\mathbf{z}_1, ..., \mathbf{z}_N)$ be a $N$-sample of independant complex elliptical vectors such as $\mathbf{z}_i \sim \mathcal{CE}(\mathbf{0}, \boldsymbol{\Sigma}, g)$ for $i = 1...N$. The MLE of $\boldsymbol{\Sigma}$ can be written,

$$\widehat{\boldsymbol{\Sigma}}_{ML} = \frac{1}{N}\sum_{i=1}^{N} \frac{-\,g'(\mathbf{z}_i^H \widehat{\boldsymbol{\Sigma}}_{ML}^{-1}\mathbf{z}_i)}{g(\mathbf{z}_i^H \widehat{\boldsymbol{\Sigma}}_{ML}^{-1}\mathbf{z}_i)}\mathbf{z}_i\mathbf{z}_i^H \tag{6}$$

# Statistical modelling and associated estimators in signal processing




$$
\text{MLE} \begin{cases}
= & \textit{SCM} \text{ for Gaussian distributions} \\
& \widehat{\boldsymbol{\Sigma}}_{SCM} = \dfrac{1}{N}\displaystyle\sum_{n=1}^{N} \mathbf{z}_n \mathbf{z}_n^H \\
\approx & \textit{FPE} \text{ for Compound-Gaussian distributions} \\
& \widehat{\boldsymbol{\Sigma}}_{FPE} = \dfrac{m}{N}\displaystyle\sum_{n=1}^{N} \dfrac{\mathbf{c}_n.\mathbf{c}_n^H}{\mathbf{c}_n^H \widehat{\boldsymbol{\Sigma}}_{FPE}^{-1} \mathbf{c}_n} \\
= & \text{depend on g for elliptical distributions} \\
& \widehat{\boldsymbol{\Sigma}}_{ML} = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} \dfrac{-\,g'(\mathbf{z}_i^H \widehat{\boldsymbol{\Sigma}}_{ML}^{-1} \mathbf{z}_i)}{g(\mathbf{z}_i^H \widehat{\boldsymbol{\Sigma}}_{ML}^{-1} \mathbf{z}_i)} \mathbf{z}_i \mathbf{z}_i^H
\end{cases}
$$

Elliptical distributions provide a more general model, but the MLE can not be used if $g$ is unknown.

*Estimators which are robust all over elliptical distributions, must be used.*

# Outline

# Robustness Theory

## Gross error model (Huber, 1964)

Given a distribution model *G*, the real distribution *F*, is in the neighborhood of the model :

$$F = (1 - \varepsilon)G + \varepsilon H$$

- $0 \leq \varepsilon < 1$
- $H$ is an unknown distribution

### Robust estimator

- gives good results* when the distribution is $G$
- and all over $F$.

*bias controlled or negligible, low asymptotic variance.

# Robustness Theory

*M*-estimators

*M*-estimators are a robust generalization of MLE.

## *M*-estimators

Let $(\mathbf{z}_1, ..., \mathbf{z}_N)$ be an $N$-sampling of complex independant circular random vectors of length $m$. We assume that $\mathbf{z}_i \sim CE(\mathbf{0}, \mathbf{\Sigma}, g))$, $i = 1, ..., N$. The complex $M$-estimator of $\mathbf{\Sigma}$ is defined as the solution of

$$\mathbf{V}_N = \frac{1}{N} \sum_{n=1}^{N} u\left(\mathbf{z}_n^H \mathbf{V}_N^{-1} \mathbf{z}_n^H\right) \mathbf{z}_n \mathbf{z}_n^H, \tag{7}$$

Existence and uniqueness of the solution has been shown in the real case, provided weight function $u$ satisfies a set of general assumptions (Maronna, 1976). Ollila (2003) has shown that these conditions hold also in the complex case.

## Link with the scatter matrix

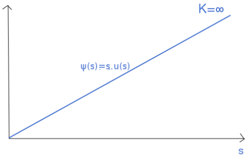
Let $\mathbf{V} = E\left[u(\mathbf{z}'\mathbf{V}^{-1}\mathbf{z})\,\mathbf{z}\mathbf{z}'\right]$ where $\mathbf{z} \sim CE(\mathbf{0}, \mathbf{\Sigma}, g)$.
The previous assumptions been respected,

- This equation admits a unique solution $\mathbf{V}$ and $\mathbf{V} = \sigma\mathbf{\Sigma}$, where $\sigma$ is given by Tyler(1982),
- A simple iterative procedure provides $\mathbf{V}_N$,
- $\mathbf{V}_N$ is a consistent estimate of $\mathbf{V}$.

# Robustness Theory

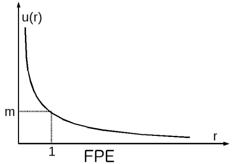
## Examples of *M*-estimators

The SCM is not a *M*-estimator since *K* is infinite.




SCM

The FPE is not a *M*-estimator since *u* is not defined in zero.




FPE

# Robustness Theory

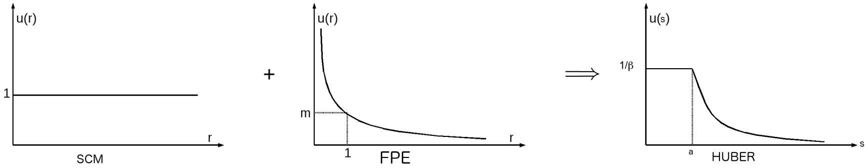
Examples of *M*-estimators

## Huber's *M*-estimator

Is such as its weight function is :

$$u(s) = \frac{1}{\beta} \min_{s \in \mathbb{R}^+} (1, a/s), = \frac{1}{\beta} \left( \mathbb{1}_{s \leq a} + \frac{a}{s} \mathbb{1}_{s > a} \right), \quad (8)$$

It is a mix between the SCM and the FPE.




Parameters $a$ and $\beta$ enable to adjust the proportion of data processed with the SCM or the FPE.

# Robustness Theory

## Huber's *M*-estimator

### Gaussian context settings :

- $q$ : proportion of data, treated as Gaussian and processed with the SCM.
- $1-q$ : proportion of data processed with FPE.
- $s = \mathbf{z}_n^H \mathbf{\Lambda}^{-1} \mathbf{z}_n$ has a $\chi^2$ distribution with $2m$ degrees of freedom.

We can write

$$q = F_{2m}(2a)$$

where $F_{2m}(.)$ is the cumulative distribution function of a $\chi^2$ with $2m$ degrees of freedom.
$\beta$ is set such as $\sigma = 1$ and verifies

$$E[\sigma\psi(|\mathbf{t}|^2)] = m$$

where $\mathbf{t} \sim \mathcal{N}(\mathbf{0}_{m,1}, \mathbf{I}_m)$ and $\psi(s) = s.u(s)$. This leads to

$$\beta = F_{2m+2}(2a) + a\frac{1-q}{m}$$

In the elliptical case, similar computations can be done.

# Robustness Theory

Robustness Criteria

## The Influence Function (IF)

Measures the effect of an infinitesimal disturbance at a given point :

$$\mathrm{IF}_{\mathbf{V},G}(\mathbf{z}_0) = \lim_{\varepsilon \to 0} \frac{\mathbf{V}\left((1-\varepsilon)G + \varepsilon\delta_{\mathbf{z}_0}\right) - \mathbf{V}(G)}{\varepsilon} \tag{9}$$

with $\delta_{\mathbf{z}_0}$ the Dirac in $\mathbf{z}_0$ and $G$ the nominal distribution.
**An estimator is robust according to the IF if its IF is continuous and bounded.**

## The Breakdown Point (BP)

Global criterion which describes how far from a model, the actual distribution can be, the estimator still giving some information on the estimated parameter.
Given the gross error model $F = (1-\varepsilon)G + \varepsilon H$,
**the BP is such as for $\varepsilon < BP$ the eigenvalues of $\mathbf{V}(F)$ remain strictly positive and bounded.**
$BP \leq \frac{1}{2}$

## The asymptotic bias $b_{\mathbf{V},G,H}$

Bias between the eigenvalues of $\mathbf{V}(F)$ and $\mathbf{V}(G)$, for each $\varepsilon < BP$. The maximum asymptotic bias is the maximum over all distributions $H$.
**An estimator is robust according to the asymptotic bias if $b_{\mathbf{V},G,H}$ is close to 1.**

# Robustness Theory

Application of the robustness criteria to the studied estimators

## The SCM

- The IF is continuous but unbounded*.
- $BP = 0$
- The asymptotic bias depends on the power of the outlier*.

**This estimator is not robust.**

## The *M*-estimators

- The IF is continuous and bounded (Ollila).
- $BP \leq \dfrac{1}{m}$
- The maximum asymptotic bias depends on the weight function. Tables can be found.

**This estimator is more robust.**

* An intermediate result is presented in a paper (Mahot 2010, Conference EUSIPCO).

# Robustness Theory

## Application of the robustness criteria to the studied estimators

### The FPE

- The IF is continuous and bounded*.
- $BP = \dfrac{1}{m}^{**}$.
- The maximum asymptotic bias depends only on $m$ and $\varepsilon^{**}$.

**This estimator is the most robust.**

* An intermediate result is presented in a paper (Mahot 2010, Conference EUSIPCO).
** The expression and proof is in the final manuscript.

# Illustration of the BP on a MUSIC application

## The MUSIC Method

Let us consider a Uniform Linear Array (ULA) with $m$ sensors and $p < m$ sources. This ULA receives a narrow band signal :

$$\mathbf{y}(t) = \mathbf{A}(\theta_0)\mathbf{s}(t) + \mathbf{b}(t)$$

with

- $\boldsymbol{\theta}_0 = \begin{pmatrix} \theta_1 & \theta_2 & \dots & \theta_p \end{pmatrix}^T$ with $\theta_k$ the angle of the $k$th signal,
- $\mathbf{A}(\boldsymbol{\theta}_0) = \begin{pmatrix} \mathbf{a}(\theta_1) & \mathbf{a}(\theta_2) & \dots & \mathbf{a}(\theta_p) \end{pmatrix}$, containing the directional vectors,
- $\mathbf{s}(t) = \begin{pmatrix} s_1(t) & s_2(t) & \dots & s_p(t) \end{pmatrix}^T$ vector of the source signals,
- $\mathbf{b}(t)$, additive stationnary noise with zero mean. The vectors$\mathbf{b}(t)$ are independant.

The covariance matrix is

$$\mathbf{R} = \mathbf{A}(\theta)\mathbf{s}\mathbf{s}^H\mathbf{A}^H(\theta) + \sigma^2\mathbf{I} = \sum \lambda_k \mathbf{e}_k\mathbf{e}_k^H = \mathbf{E}_s\Lambda_s\mathbf{E}_s^H + \sigma^2\mathbf{E}_n\mathbf{E}_n^H,$$

with

- $\mathbf{e}_k$ the eigenvector associated to $\lambda_k$.
- $\mathbf{E}_s = [\mathbf{e}_1 \dots \mathbf{e}_p]$ generating the signal subspace.
- $\mathbf{E}_n = [\mathbf{e}_{p+1} \dots . \mathbf{e}_m]$ generating the noise subspace($p < m$).

The angles of the arriving signals are given by the maxima of

$$V_{MUSIC}(\theta) = \frac{1}{\mathbf{a}^H(\theta)\mathbf{E}_n\mathbf{E}_n^H\mathbf{a}(\theta)}$$

# Illustration of the BP on a MUSIC application

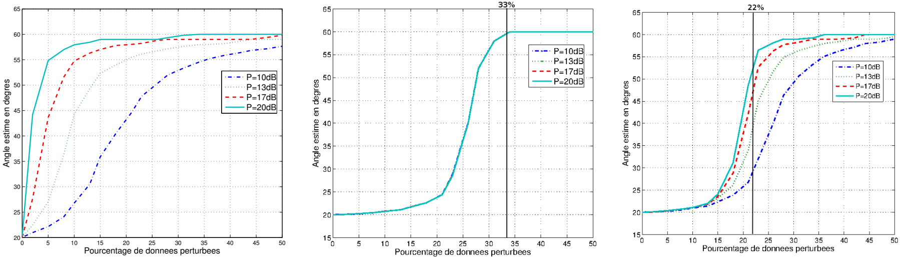
## Application

We consider a ULA with

- $m = 3$ sensors,
- $N = 1000$ independant samples,
- a Gaussian stationary signal with DOA $20^\circ$,
- a Gaussian white additive noise (SNR=3dB).
- A proportion $\varepsilon$ of vectors is replaced by outliers corresponding to a Gaussian stationnary signal of DOA $60^\circ$.
- Outliers with different powers are represented

# Illustration of the BP on a MUSIC application

## Application




$$BP_{SCM} = 0 \qquad BP_{FPE}\,\frac{1}{m} = \frac{1}{3} \qquad BP_{Huber} \simeq 22\% \leq \frac{1}{m} = \frac{1}{3}$$

# Outline

# Asymptotic Performance

When data are complex elliptical, the asymptotic distributions of the studied estimators can be written

$$\sqrt{N}\text{vec}(\widehat{\mathbf{\Sigma}} - \mathbf{\Sigma}) \xrightarrow{d} \mathcal{GCN}\left(\mathbf{0}_{m^2,1}, \mathbf{\Lambda}, \mathbf{\Omega}\right)$$

where $\mathcal{GCN}$ is the general complex Gaussian distribution, $\mathbf{\Lambda}$ the covariance matrix and $\mathbf{\Omega}$ the pseudo-covariance matrix.

## SCM

For Complex Gaussian data, $\mathbf{\Lambda} = (\mathbf{\Sigma}^T \otimes \mathbf{\Sigma})$ $\mathbf{\Omega} = (\mathbf{\Sigma}^T \otimes \mathbf{\Sigma})\mathbf{K}$, where $\mathbf{K}$ is the commutation matrix.

## FPE

For Compound-Gaussian data,

$$\mathbf{\Lambda} = \frac{m+1}{m}\left[(\mathbf{\Sigma}^T \otimes \mathbf{\Sigma}) - \frac{1}{m}\text{vec}(\mathbf{\Sigma})\text{vec}(\mathbf{\Sigma})^H\right] \quad \mathbf{\Omega} = \frac{m+1}{m}\left[(\mathbf{\Sigma}^T \otimes \mathbf{\Sigma})\mathbf{K} - \frac{1}{m}\text{vec}(\mathbf{\Sigma})\text{vec}(\mathbf{\Sigma})^T\right]$$

## Complex *M*-estimators*

For complex elliptical distributions,

$$\mathbf{\Lambda} = \sigma_1 \mathbf{\Sigma}^T \otimes \mathbf{\Sigma} + \sigma_2 \text{vec}(\mathbf{\Sigma})\text{vec}(\mathbf{\Sigma})^H, \quad \mathbf{\Omega} = \sigma_1 (\mathbf{\Sigma}^T \otimes \mathbf{\Sigma})\mathbf{K} + \sigma_2 \text{vec}(\mathbf{\Sigma})\text{vec}(\mathbf{\Sigma})^T,$$

* Result accepted with major comments in IEEE SP-T.

# Asymptotic Performance

Particular Property

- Let $\mathbf{\Sigma}$ be a fixed hermitian positive-definite matrix and $\mathbf{\Sigma}_N$ a sequence of symmetric positive definite random matrices of order $m$ which satisfies

$$\sqrt{N}\,(\text{vec}(\mathbf{\Sigma}_N - \mathbf{\Sigma})) \xrightarrow{d} \mathcal{GCN}\left(\mathbf{0}_{m^2,1}, \mathbf{\Lambda}, \mathbf{\Omega}\right),$$

  with

$$\mathbf{\Lambda} = \nu_1 \mathbf{\Sigma}^T \otimes \mathbf{\Sigma} + \nu_2 \text{vec}(\mathbf{\Sigma})\text{vec}(\mathbf{\Sigma})^H \quad \mathbf{\Omega} = \nu_1 (\mathbf{\Sigma}^T \otimes \mathbf{\Sigma})\mathbf{K} + \nu_2 \text{vec}(\mathbf{\Sigma})\text{vec}(\mathbf{\Sigma})^T,$$

  where $\nu_1$ and $\nu_2$ are any real numbers.
- Let $H(\mathbf{\Sigma})$ be a $r$-dimensional multivariate function on the set of $m \times m$ complex hermitian positive-definite matrices, possessing continuous first partial derivatives and such as $H(\mathbf{\Sigma}) = H(\alpha\mathbf{\Sigma})$ for all $\alpha > 0$.

# Particular property

Link with the studied estimators

The asymptotic distribution of $H(\mathbf{\Sigma}_N)$ is given by*

$$\sqrt{N}\left(H(\mathbf{\Sigma}_N) - H(\mathbf{\Sigma})\right) \xrightarrow{d} \mathcal{GCN}\left(\mathbf{0}_{r,1}, \mathbf{\Lambda}_H, \mathbf{\Omega}_H\right).$$

where

$$\mathbf{\Lambda}_H = \nu_1 H'(\mathbf{\Sigma})(\mathbf{\Sigma}^T \otimes \mathbf{\Sigma})H'(\mathbf{\Sigma})^H \quad \mathbf{\Omega}_H = \nu_1 H'(\mathbf{\Sigma})(\mathbf{\Sigma}^T \otimes \mathbf{\Sigma})\mathbf{K}H'(\mathbf{\Sigma})^T,$$

and $H'(\mathbf{V}) = \left(\dfrac{dH(\mathbf{V})}{d\text{vec}(\mathbf{V})}\right)$.

This result has been obtained by Tyler in the real case and we have extended it to the complex elliptical case.

## Link with the studied estimators

- The SCM verifies the conditions and $\nu_1 = 1$ in the Gaussian context.
- The FPE verifies the conditions and $\nu_1 = \dfrac{m+1}{m}$ in the Compound-Gaussian context.
- The *M*-estimators verifie the conditions and $\nu_1 = \sigma_1$ in the complex elliptical context.

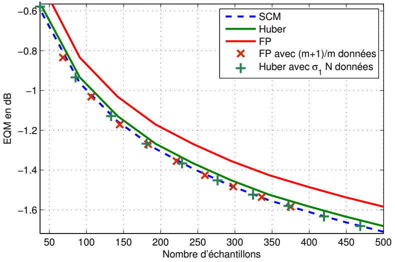
* Result accepted with major comments in IEEE SP-T..

# Asymptotic Performance

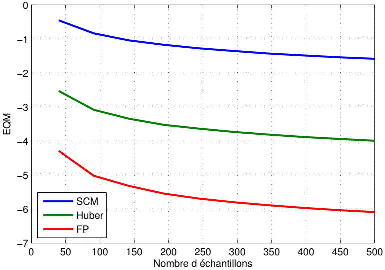
## Illustration with the MUSIC method

We consider a ULA with

- $m = 3$ sensors,
- a Gaussian stationary signal with DOA $20^{\circ}$,
- a Gaussian white additive noise in the first figure, and a K-distributed additive noise of parameter $\nu = 0.1$ in the second figure (SNR=5dB/sensor).
- The x-axis gives the number $N$ of independant data used to estimate the covariance matrix.




MSE of the DOA estimated with the SCM, Huber and the FPE, with Gaussian white additive noise. $\sigma_1 \simeq 1.066$, $\dfrac{m+1}{m} \simeq 1,33$.




MSE of the DOA estimated with the SCM, Huber and the FPE, K-distributed additive noise of parameter 0.1.

- Gaussian case : same results up to a scale factor.
- Non-Gaussian case : FPE and *M*-estimators give better results.

# Outline

# Context

## Binary hypothesis test

$$\begin{cases} H_0 : \mathbf{z} = \mathbf{c} & \mathbf{z}_i = \mathbf{c}_i \quad i = 1, ..., N \\ H_1 : \mathbf{z} = \mathbf{s} + \mathbf{c} & \mathbf{z}_i = \mathbf{c}_i \quad i = 1, ..., N \end{cases}$$

with

- $\mathbf{s} = A\mathbf{p}$ the complex known signal characterizing the target. $A$ is the amplitude of the signal an $\mathbf{p}$ the steering vector containing all the other information
- $\mathbf{c}$ the additive noise (*clutter*),
- $\mathbf{z}_i$ the observation vectors. They are assumed i.i.d.

The purpose of the detection is to decide which hypothesis is the more likely to be true.

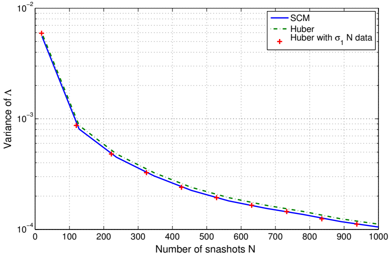
## The Adaptive Normalized Match Filter (ANMF)

$$\Lambda(\widehat{\mathbf{M}}|\mathbf{z}) = \frac{|\mathbf{p}^H \widehat{\mathbf{M}}^{-1} \mathbf{z}|^2}{(\mathbf{p}^H \widehat{\mathbf{M}}^{-1} \mathbf{p})(\mathbf{z}^H \widehat{\mathbf{M}}^{-1} \mathbf{z})}.$$

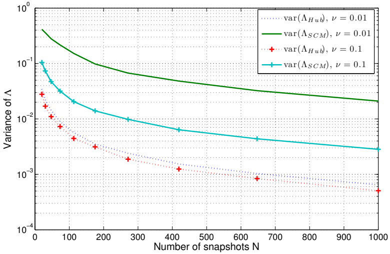
# Application : detection problem in adaptive processing

## Performances of the ANMF with different estimators

Here the function $H(.)$ of the particular property, is such as $H(\widehat{\mathbf{M}}) = \Lambda(\widehat{\mathbf{M}}|\mathbf{z}) = \Lambda(\alpha\widehat{\mathbf{M}}|\mathbf{z})$




Variance of the ANMF detector for Huber estimate and the SCM. Additive white Gaussian noise. Here $\sigma_1 = 1.067$.




Variance of the ANMF detector for Huber estimate and the SCM. Additive K-distributed noise

# Second example : $P_{fa}$-threshold relationship

Background

## Probability of false alarme ($P_{fa}$)

Probability of detecting a signal in the noise only case.

## $P_{fa}$-threshold relationship

**In the Gaussian case and with the SCM, it is known to be :**

$$P_{fa} = (1-\lambda)^{a-1} \, {}_2F_1(a, a-1; b-1; \lambda), \tag{10}$$

where

$$a = N - m + 2, \, b = N + 2$$

and ${}_2F_1$ is the Hypergeometric function defined as

$${}_2F_1(a, b; c; x) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{k=0}^{\infty} \frac{\Gamma(a+k)\Gamma(b+k)}{\Gamma(c+k)} \frac{x^k}{k!}$$

# Second example : $P_{fa}$-threshold relationship

Generalization

## Generalization*

- For any estimator verifying the conditions of the particular property,
- any complex elliptically distributed signal
- and for $N$ large enough

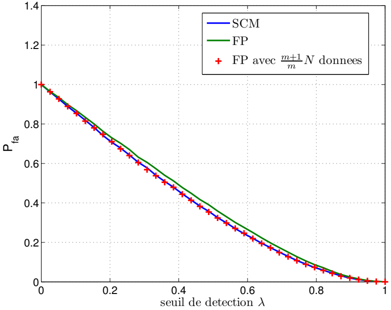
the $P_{fa}$-threshold relationship is still given by (10) if we replace $N$ by $N/\nu_1$.

* Result published in a conference (Mahot 2012, conference SAM).

# Application : detection problem in adaptive processing

Illustration of the approached $P_{fa}$-threshold relationship

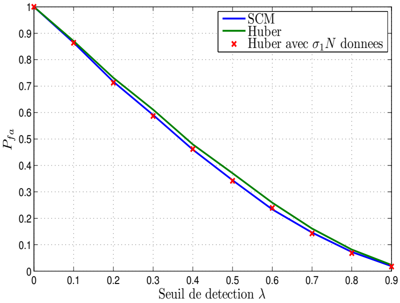
In the case of the FPE, for $N$ large enough equation (10) stands by replacing $N$ by $\dfrac{m}{m+1}N$. In the case of $M$-estimators, $N$ must be replaced by $N/\sigma_1$.




$P_{fa}$-threshold curve obtained with

- the SCM and the FPE with $N = 10$ data,
- the FPE with $\frac{m+1}{m}N \simeq 13$ data.

White Gaussian noise. $m = 3$.




$P_{fa}$-threshold curve obtained with

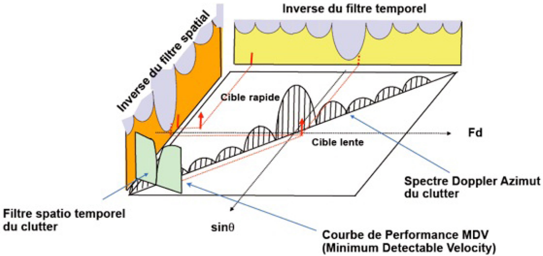
- the SCM and Huber's $M$-estimator with $N = 10$ data,
- Huber's $M$-estimator with $\sigma_1 N \simeq 12$ data.

White Gaussian noise, $\sigma_1 = 1.23$, $q = 0.25$ and $m = 3$.

# Application of different estimators to Space Time Adaptive Processing (STAP)

## Presentation

- STAP are used, to detect moving targets.
- Instead of using space processing and time processing separatly, it uses the spatio-temporal link between arriving signals.
- Results in a better filtering of the noise and ground echo.

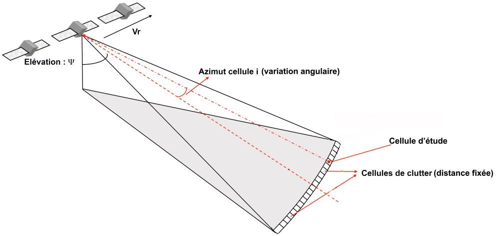



Space processing, time processing and space-time processing

# Application of different estimators to Space Time Adaptive Process (STAP)

Context

- For one range bin (distance), 4 sensors receive 48 pulses.
- The resulting steering vectors allow to detect the DOA and Doppler speed of potential targets.
- the covariance matrix is estimated using the data of the over range bins.
- The results of the detector is observed on one range bin.

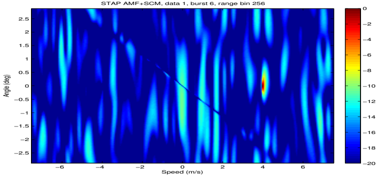



Angle scanning for a side looking radar

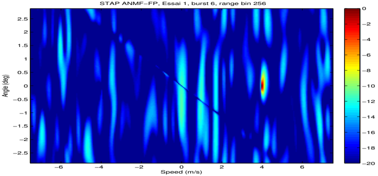
# Application of different estimators to Space Time Adaptive Process (STAP)

## Gaussian Background

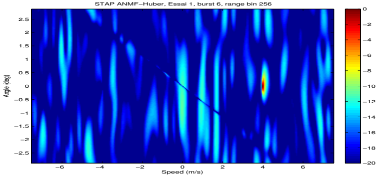
- Gaussian background (with no outlier).
- Unique target of DOA $0^\circ$, Doppler speed $V_c = 4$ m/s and range bin 256.



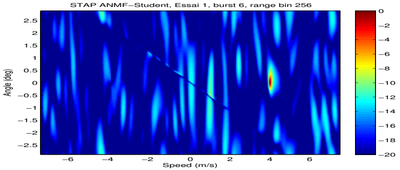

Detection of the target with the SCM




Detection of the target with the FPE




Detection of the target with Huber's $M$-estimator, $q = 0.6$

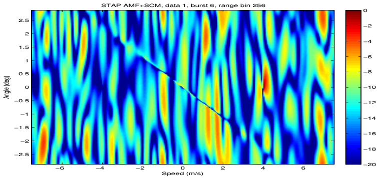



Detection of the target with Student's $M$-estimator, $\nu = 2$

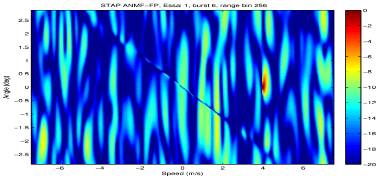
# Application of different estimators to Space Time Adaptive Process (STAP)

## Gaussian Background with outlier

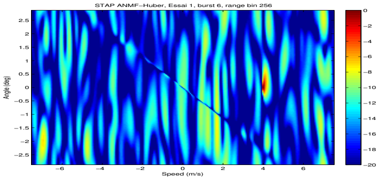
- Same background but with an outlier in the learning data : A second target with same DOA and Doppler in a different range bin.



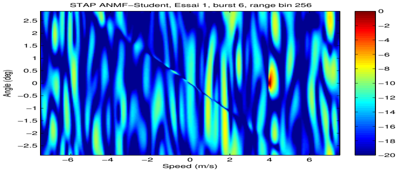

Detection of the target with the SCM




Detection of the target with the FPE




Detection of the target with Huber's $M$-estimator, $q = 0.6$

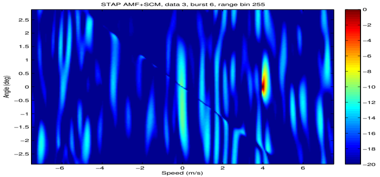



Detection of the target with Student's $M$-estimator, $\nu = 2$

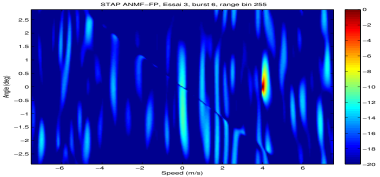
# Application of different estimators to Space Time Adaptive Process (STAP)

non-Gaussian Background

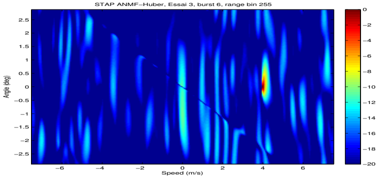
- (slightly) non-Gaussian data obtained by the SAR (Synthetic Aperture Radar)THR RAMSES.
- Unique target with range bin 255, DOA 0° and Doppler speed $V_c = 4$ m/s.



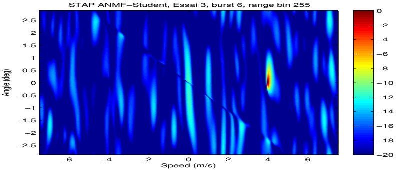

Detection of the target with the SCM




Detection of the target with the FPE




Detection of the target with Huber's $M$-estimator, $q = 0.6$




Detection of the target with Student's $M$-estimator, $\nu = 2$

# Outline

# Conclusion

- The classical Gaussian modelling and its MLE the SCM, giving often poor results, more general models have been studied (GCV, elliptical distributions).
- Since the exact distribution of the data is unknown, the MLE can not be used. Estimors which are robust all over a class of distribution have been studied
- The asymptotic performance of these estimators has been studied. A particular property with interesting consequences has been highlighted.
- These results have been illustrated with the detection problem and with STAP data.

# Prospects

The FPE and $M$-estimators could be used in many signal processing applications, for more robust results. However, some problems remain :

- Non-stationnary data case.
- The FPE or the $M$-estimators give bad results if there is not enough data. (structure of the matrix, low rank techniques)
- The robustness of the FPE and the $M$-estimators decreases with the dimension of the vectors. (other estimators, random matrix)

This study could be generalized to applications where the joint estimation of the mean and covariance matrix is needed.

Thank you for your attention.

# My publications

## Conferences


1. M. Mahot, P. Forster, J.-P. Ovarlez and F. Pascal, "*Robustness Analysis of Covariance Matrix Estimates, European Signal Processing Conference (EUSIPCO)*", Aalborg, Denmark, August 2010.
2. M. Mahot, P. Forster, J.-P. Ovarlez and F. Pascal, "*Robust Covariance Matrix Estimation in Presence of Outliers*", Workshop SONDRA, Cargese, France, May 2010.
3. M. Mahot, P. Forster, F. Pascal and J.-P. Ovarlez, "*Asymptotic Properties of Robust Complex Covariance Matrix Estimate*" (in french), Proc. GRETSI, Bordeaux, France, September 2011.
4. M. Mahot, P. Forster, F. Pascal and J.-P. Ovarlez, "*Robust Covariance Matrix Estimates with Attractive Asymptotic Properties*", Worskop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), San Juan, Porto Rico, December 2011.
5. M. Mahot, F. Pascal, P. Forster and J.-P. Ovarlez, "*Robust ANMF test using Huber's M-estimator*", Sensor Array and Multichannel Signal Processing Workshop (SAM), Hboken, NJ, June 2012.
6. J. Frontera-Pons, M. Mahot, J.P. Ovarlez, F. Pascal, J. Chanussot, "*Robust detection using M-estimators for Hyperspectral Imaging*",Workshop on Hyperspectral Imaging and Signal Processing : Evolution in Remote Sensing (WHISPER), Shanghai, China, June 2012.
7. J. Frontera-Pons, M. Mahot, J.P. Ovarlez, F. Pascal, S.K. Pang, J. Chanussot, *"A class of robust estimates for detection in hyperspectral images using elliptical distributions background"*, IEEE Geoscience and Remote Sensing Society (IGARSS), Munich, Germany, July 2012
8. J.-P. Ovarlez, F. Pascal, P. Forster, M. Mahot and G. Pailloux, "*A Review of Non-Gaussian Detectors on Different Set of Experimental Data*", Workshop SONDRA, Cargese, France, May 2010.


## Journals


1. M. Mahot, P. Forster, F. Pascal and J.-P. Ovarlez, "*Asymptotic properties of robust complex covariance matrix estimates*", IEEE Trans. On Signal Processing, 2012, accepted with major revision.
2. J.P. Ovarlez, F. Pascal, P. Forster, G. Ginolhac, M. Mahot, "*Traitements STAP et Modélisation SIRV : Robustesse et Persymétrie*", Special Issue "Traitements Spatio-Temporels Adaptatifs (STAP)", Traitement du Signal, Vol.28, No.1-2, pp.113-142, 2011.