

Schémas de Détection Adaptative Robuste en Environnement non Gaussien, hétérogène et en présence d'outliers - Application au Traitement Radar Adaptatif Spatio-Temporel (STAP)

Jean-Philippe Ovarlez^{1,2}

¹SONDRA, CentraleSupélec, France

²French Aerospace Lab, ONERA DEMR/TSI, France

Joint works with F. Pascal, P. Forster, G. Ginolhac, M. Mahot, A. Breloy, and many others

Journée GDR *Traitement d'antenne : Signaux Non-Gaussiens,
Non-Circulaires, Non-Stationnaires*
8 décembre 2016, Telecom ParisTech

Contents

- 1 Preliminaries
 - Motivations
 - Some Background on Detection Theory
 - Case of Adaptive Gaussian Detection in Gaussian Background
- 2 Adaptive Robust Detection Schemes in non-Gaussian Background
 - CES distributions
 - M -estimators and Tyler (FP) Estimator
 - Robustness of M -estimators and ANMF
 - Exploiting Prior Information: Covariance Structure
 - Low Rank Detectors
 - Shrinkage of M -estimator
- 3 Applications
 - Surveillance Radar
 - STAP Applications
- 4 Conclusions and Perspectives

Outline

1 Preliminaries

■ Motivations

- Some Background on Detection Theory
- Case of Adaptive Gaussian Detection in Gaussian Background

2 Adaptive Robust Detection Schemes in non-Gaussian Background

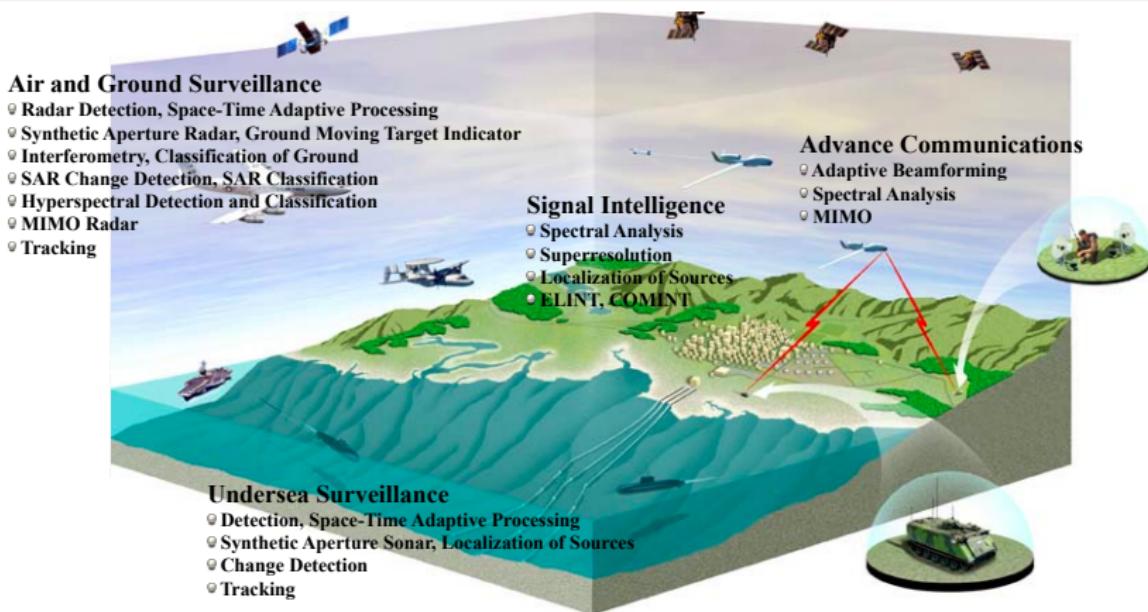
- CES distributions
- M -estimators and Tyler (FP) Estimator
- Robustness of M -estimators and ANMF
- Exploiting Prior Information: Covariance Structure
- Low Rank Detectors
- Shrinkage of M -estimator

3 Applications

- Surveillance Radar
- STAP Applications

4 Conclusions and Perspectives

Motivations: Almost all algorithms and systems analysis for detection, estimation and classification rely on Covariance-Based methods



Under Gaussian assumptions $\mathcal{CN}(\mathbf{0}, \Sigma)$, the Sample Covariance Matrix (SCM) is the most likely covariance matrix estimate (MLE) and is the empirical mean of the cross-correlation of n m -vectors \mathbf{z}_k :

$$\widehat{\mathbf{S}}_n = \frac{1}{n} \sum_{k=1}^n \mathbf{z}_k \mathbf{z}_k^H$$

- This estimate is unbiased, efficient, Wishart distributed,
- n can represent any samples support: in time, spatial, angular domain, \mathbf{z}_k a vector of any information collected in any domain:
 - in Radar Detection, it can represent the time returns collected in a given range bin of interest, n is here the range bin support
 - in Array Processing, it can represent the spatial information collected by the antenna array at a given time, n is here the time support,
 - in STAP, it can represent the joint spatial and time information collected in a given range bin of interest, n is here the time support,
 - in SAR or Hyperspectral imaging, it can represent the polarimetric and/or interferometric, or spectral information collected for a given pixel of the spatial image, n is here the spatial support.

- To have a SCM estimate invertible (whitening process), the number n of samples has to be bigger than the size m of the information collected \mathbf{z}_k ,
- To improve the quality of the estimate, n has to be high but it means also that the space support has also to respect the initial Gaussian hypothesis (has to be statistically homogeneous) that is not always the case in the real world !
- Due to the increase of the radar resolution or due to the illumination angle, the number of the scatterers present in each cell (random walk) can become very small, the Central Limit Theorem being no longer valid. Even if the number of scatterers is large enough to apply the CLT, this number can also randomly fluctuate from one resolution cell to another, leading to a backscattered signal locally Gaussian with random power (heterogeneous support)
- Robustness of the SCM: The n secondary data used to estimate the SCM may also contain another target returns, jammers, strong undesired scatterers which can lead to a poor or a biased estimate.

Preliminaries

Adaptive Robust Detection Schemes in non-Gaussian Background
Applications
Conclusions and Perspectives

Motivations

Some Background on Detection Theory
Case of Adaptive Gaussian Detection in Gaussian Background

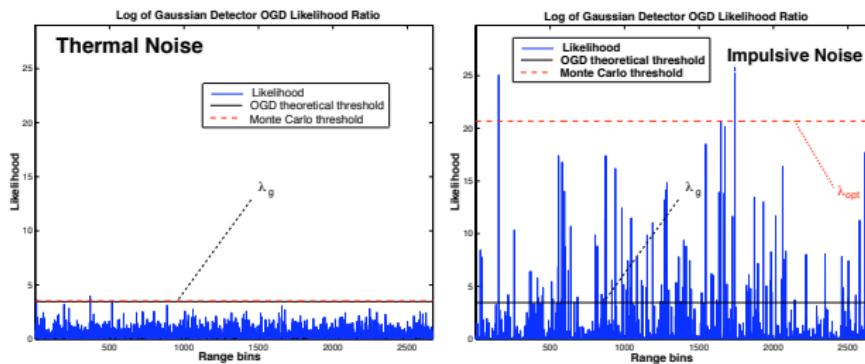


Figure: Failure of the Gaussian detector ($\lambda_g = -\sigma^2 \log P_{fa}$): (left) Adjustment of the detection threshold, (right) K-distributed clutter with same power as the Gaussian noise

- ⇒ Bad performance of the conventional Gaussian detector in case of mis-modeling
- ⇒ Need/Use of non-Gaussian distributions
- ⇒ Need/Use of robust estimates

Outline

1 Preliminaries

- Motivations
- Some Background on Detection Theory
- Case of Adaptive Gaussian Detection in Gaussian Background

2 Adaptive Robust Detection Schemes in non-Gaussian Background

- CES distributions
- M -estimators and Tyler (FP) Estimator
- Robustness of M-estimators and ANMF
- Exploiting Prior Information: Covariance Structure
- Low Rank Detectors
- Shrinkage of M -estimator

3 Applications

- Surveillance Radar
- STAP Applications

4 Conclusions and Perspectives

Problem Statement

- In a m -vector \mathbf{z} , detecting a unknown complex deterministic signal $\mathbf{s} = \mathbf{A}\mathbf{p}$ embedded in an additive noise \mathbf{y} (with covariance matrix Σ) , can be written as the following statistical test:

$$\begin{cases} \text{Hypothesis } H_0: \mathbf{z} = \mathbf{y} & \mathbf{z}_i = \mathbf{y}_i, \quad i = 1, \dots, n \\ \text{Hypothesis } H_1: \mathbf{z} = \mathbf{s} + \mathbf{y} & \mathbf{z}_i = \mathbf{y}_i, \quad i = 1, \dots, n \end{cases}$$

where the \mathbf{z}_i 's are n "signal-free" independent secondary data used to estimate the noise parameters .

⇒ Neyman-Pearson criterion

- Detection test:** comparison between the Likelihood Ratio $\Lambda(\mathbf{z})$ and a detection threshold λ :

$$\Lambda(\mathbf{z}) = \frac{p_{\mathbf{z}}(\mathbf{z}/H_1)}{p_{\mathbf{z}}(\mathbf{z}/H_0)} \stackrel{H_1}{\gtrless} \lambda,$$

- Probability of False Alarm (type-I error): $P_{fa} = \mathbb{P}(\Lambda(\mathbf{z}) > \lambda/H_0)$
- Probability of Detection: $P_d = \mathbb{P}(\Lambda(\mathbf{z}) > \lambda/H_1)$ for different Signal-to-Noise Ratios (SNR).

Well known Gaussian Detectors (Σ known)

- Homogeneous Gaussian case (Matched Filter - Optimum Gaussian Detector): if $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, \Sigma)$ then

$$\Lambda(\mathbf{z}) = \frac{|\mathbf{p}^H \Sigma^{-1} \mathbf{z}|^2}{\mathbf{p}^H \Sigma^{-1} \mathbf{p}} \stackrel{H_1}{\underset{H_0}{\gtrless}} \lambda_g$$

with $\lambda_g = \sqrt{-\ln P_{fa}}$.

- Partially Homogeneous Gaussian case (Normalized Matched Filter): if $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, \alpha \Sigma)$ with α unknown:

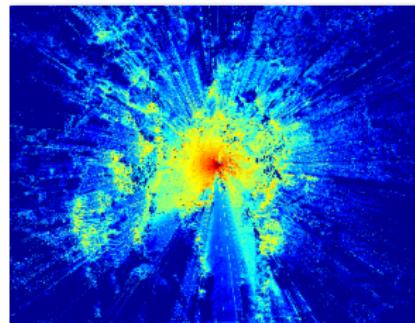
$$\Lambda(\mathbf{z}) = \frac{|\mathbf{p}^H \Sigma^{-1} \mathbf{z}|^2}{(\mathbf{p}^H \Sigma^{-1} \mathbf{p})(\mathbf{z}^H \Sigma^{-1} \mathbf{z})} \stackrel{H_1}{\underset{H_0}{\gtrless}} \lambda_{NMF}$$

The False Alarm regulation can be theoretically done thanks to

$$\lambda_{NMF} = 1 - P_{fa}^{\frac{1}{m-1}}.$$

Going to adaptive detection

Generally, some parameters (**say Σ !**) are unknown.



⇒ Covariance Matrix Estimation

Requirements:

- Background modeling: Gaussian, SIRV (K-distribution, Weibull, etc.), CES (Multidimensional Generalized Gaussian Distributions, etc.),
- Estimation procedure: ML-based approaches, M -estimation, LS-based methods, etc.
- Adaptive detectors derivation and adaptive performance evaluation.

Outline

1 Preliminaries

- Motivations
- Some Background on Detection Theory
- **Case of Adaptive Gaussian Detection in Gaussian Background**

2 Adaptive Robust Detection Schemes in non-Gaussian Background

- CES distributions
- M -estimators and Tyler (FP) Estimator
- Robustness of M -estimators and ANMF
- Exploiting Prior Information: Covariance Structure
- Low Rank Detectors
- Shrinkage of M -estimator

3 Applications

- Surveillance Radar
- STAP Applications

4 Conclusions and Perspectives

Homogeneous Gaussian noise/clutter

The Sample Covariance Matrix (SCM)

$$\widehat{\mathbf{S}}_n = \frac{1}{n} \sum_{i=1}^n \mathbf{z}_i \mathbf{z}_i^H$$

where \mathbf{z}_i are complex independent circular zero-mean Gaussian with covariance matrix Σ , i.e. $p_{\mathbf{z}_i}(\mathbf{z}_i) = \frac{1}{(\pi)^m |\Sigma|} \exp(-\mathbf{z}_i^H \Sigma^{-1} \mathbf{z}_i)$.

The Shrinkage or Diagonal Loading SCM [O. Ledoit and M. Wolf]

$$\widehat{\mathbf{S}}_{Sh.} = (1 - \beta) \frac{1}{n} \sum_{i=1}^n \mathbf{z}_i \mathbf{z}_i^H + \beta \mathbf{I} \quad \text{or} \quad \widehat{\mathbf{S}}_{DL} = \frac{1}{n} \sum_{i=1}^n \mathbf{z}_i \mathbf{z}_i^H + \beta \mathbf{I}$$

Standard approaches: Gaussian noise/clutter

Properties of the SCM

- Simple Covariance Matrix estimator,
- Very tractable,
- Wishart distributed,
- Well-known statistical properties: unbiased and efficient.

Then, $\sqrt{n} \operatorname{vec}(\widehat{\mathbf{S}}_n - \Sigma) \xrightarrow{d} \mathcal{CN}(\mathbf{0}, \mathbf{C}, \mathbf{P})$

where $\mathbf{C} = (\Sigma^* \otimes \Sigma)$
 $\mathbf{P} = (\Sigma^* \otimes \Sigma) \mathbf{K}_{m,m}$

Adaptive Gaussian Detection

$$\text{Gaussian model} \Rightarrow \hat{\mathbf{S}}_n = \frac{1}{n} \sum_{i=1}^n \mathbf{z}_i \mathbf{z}_i^H$$

- two-step GLRT AMF test [F. C. Robey *et al.*, 1992]

$$\Lambda_{AMF}(\mathbf{z}) = \frac{\left| \mathbf{p}^H \hat{\mathbf{S}}_n^{-1} \mathbf{z} \right|^2}{\left(\mathbf{p}^H \hat{\mathbf{S}}_n^{-1} \mathbf{p} \right)} \stackrel[H_1]{>}{\stackrel[H_0]{<}{}} \lambda_{AMF}. \quad (1)$$

- GLRT Kelly test [E. J. Kelly, 1986]

$$\Lambda_{Kelly}(\mathbf{z}) = \frac{\left| \mathbf{p}^H \hat{\mathbf{S}}_n^{-1} \mathbf{y} \right|^2}{\left(\mathbf{p}^H \hat{\mathbf{S}}_n^{-1} \mathbf{p} \right) \left(n + \mathbf{z}^H \hat{\mathbf{S}}_n^{-1} \mathbf{z} \right)} \stackrel[H_1]{>}{\stackrel[H_0]{<}{}} \lambda_{Kelly}. \quad (2)$$

Outline

1 Preliminaries

- Motivations
- Some Background on Detection Theory
- Case of Adaptive Gaussian Detection in Gaussian Background

2 Adaptive Robust Detection Schemes in non-Gaussian Background

■ CES distributions

- *M*-estimators and Tyler (FP) Estimator
- Robustness of *M*-estimators and ANMF
- Exploiting Prior Information: Covariance Structure
- Low Rank Detectors
- Shrinkage of *M*-estimator

3 Applications

- Surveillance Radar
- STAP Applications

4 Conclusions and Perspectives

Modeling the background

Let \mathbf{z} be a complex circular random vector of length m . \mathbf{z} has a complex elliptically symmetric (CES) distribution ($CE(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g.)$) if its PDF is

$$g_{\mathbf{z}}(\mathbf{z}) = \pi^{-m} |\boldsymbol{\Sigma}|^{-1} h_z((\mathbf{z} - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{z} - \boldsymbol{\mu})), \quad (3)$$

where $h_z : [0, \infty) \rightarrow [0, \infty)$ is the density generator, where $\boldsymbol{\mu}$ is the statistical mean (generally known or $= \mathbf{0}$) and $\boldsymbol{\Sigma}$ is the scatter matrix. In general, $E[\mathbf{z} \mathbf{z}^H] = \alpha \boldsymbol{\Sigma}$ where α is known.

- Large class of distributions: Gaussian ($h_z(z) = \exp(-z)$), SIRV, MGGD ($h_z(z) = \exp(-z^\alpha)$), etc.
- Closed under affine transformations,
- Stochastic representation theorem: $\boxed{\mathbf{z} =_d \boldsymbol{\mu} + \mathcal{R} \mathbf{A} \mathbf{u}^{(k)}}$, where $\mathcal{R} \geq 0$, independent of $\mathbf{u}^{(k)}$ and $\boldsymbol{\Sigma} = \mathbf{A} \mathbf{A}^H$ is a factorisation of $\boldsymbol{\Sigma}$, where $\mathbf{A} \in \mathbb{C}^{m \times k}$ with $k = \text{rank}(\boldsymbol{\Sigma})$.

SIRV: a CES subclass

The m -vector \mathbf{z} is a complex Spherically Invariant Random Vector if its PDF can be put in the following form:

$$g_{\mathbf{z}}(\mathbf{z}) = \int_0^{\infty} \frac{1}{\pi^m |\Sigma| \tau^m} \exp\left(-\frac{(\mathbf{z} - \boldsymbol{\mu})^H \Sigma^{-1} (\mathbf{z} - \boldsymbol{\mu})}{\tau}\right) p_{\tau}(\tau) d\tau \quad (4)$$

where $p_{\tau} : [0, \infty) \rightarrow [0, \infty)$ is the texture generator.

- Large class of distributions: Gaussian ($p_{\tau}(\tau) = \delta(\tau - 1)$), K-distribution (p_{τ} gamma), Weibull (no closed form), Student-t (p_{τ} inverse gamma), etc.
 Main Gaussian Kernel: closed under affine transformations,
- The texture random scalar is modeling the variation of the power of the Gaussian vector \mathbf{x} along his support (e.g. heterogeneity of the noise along range bins, time, spatial domain, etc.),
- Stochastic representation theorem: $\boxed{\mathbf{z} =_d \boldsymbol{\mu} + \sqrt{\tau} \mathbf{A} \mathbf{x}}$, where $\tau \geq 0$ is the texture, independent of \mathbf{x} and $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \Sigma)$.

Outline

- 1 Preliminaries
 - Motivations
 - Some Background on Detection Theory
 - Case of Adaptive Gaussian Detection in Gaussian Background
- 2 Adaptive Robust Detection Schemes in non-Gaussian Background
 - CES distributions
 - **M -estimators and Tyler (FP) Estimator**
 - Robustness of M-estimators and ANMF
 - Exploiting Prior Information: Covariance Structure
 - Low Rank Detectors
 - Shrinkage of M -estimator
- 3 Applications
 - Surveillance Radar
 - STAP Applications
- 4 Conclusions and Perspectives

Estimating the covariance matrix: Conventional estimators

Assuming n available SIRV secondary data $\mathbf{z}_k = \sqrt{\tau_k} \mathbf{x}_k$ where $\mathbf{x}_k \sim \mathcal{CN}(\mathbf{0}, \Sigma)$ and where τ_k scalar random variable.

- The Sample Covariance Matrix SCM may be a poor estimate of the Elliptical/SIRV Scatter/Covariance Matrix because of the texture contamination:

$$\hat{\Sigma}_n = \frac{1}{n} \sum_{k=1}^n \mathbf{z}_k \mathbf{z}_k^H = \frac{1}{n} \sum_{k=1}^n \tau_k \mathbf{x}_k \mathbf{x}_k^H \neq \frac{1}{n} \sum_{k=1}^n \mathbf{x}_k \mathbf{x}_k^H$$

- The Normalized Sample Covariance Matrix (NSCM) may be a good candidate of the Elliptical SIRV Scatter/Covariance Matrix:

$$\hat{\Sigma}_{NSCM} = \frac{1}{n} \sum_{k=1}^n \frac{\mathbf{z}_k \mathbf{z}_k^H}{\mathbf{z}_k^H \mathbf{z}_k} = \frac{1}{n} \sum_{k=1}^n \frac{\mathbf{x}_k \mathbf{x}_k^H}{\mathbf{x}_k^H \mathbf{x}_k}$$

This estimate does not depend on the texture τ_k but it is biased and share the same eigenvectors but have different eigenvalues, with the same ordering [Bausson *et al.* 2006].

Estimating the covariance matrix

Let $(\mathbf{z}_1, \dots, \mathbf{z}_n)$ be a n -sample $\sim CE_m(\mathbf{0}, \Sigma, g_{\mathbf{z}(\cdot)})$ (Secondary data).

PDF $g_{\mathbf{z}}(\cdot)$ specified: ML-estimator of Σ

$$\widehat{\Sigma} = \frac{1}{n} \sum_{i=1}^n \frac{-g'_{\mathbf{z}} \left(\mathbf{z}_i^H \widehat{\Sigma}^{-1} \mathbf{z}_i \right)}{g_{\mathbf{z}} \left(\mathbf{z}_i^H \widehat{\Sigma}^{-1} \mathbf{z}_i \right)} \mathbf{z}_i \mathbf{z}_i^H,$$

PDF $g_{\mathbf{z}}(\cdot)$ not specified: M-estimator of Σ

$$\widehat{\Sigma} = \frac{1}{n} \sum_{i=1}^n u \left(\mathbf{z}_i^H \widehat{\Sigma}^{-1} \mathbf{z}_i \right) \mathbf{z}_i \mathbf{z}_i^H,$$

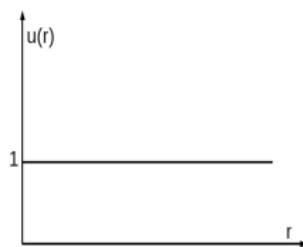
Maronna (1976), Kent and Tyler (1991)

- Existence,
- Uniqueness,
- Convergence of the recursive algorithm, etc.

Examples of M -estimators

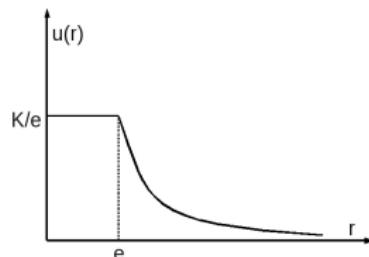
SCM:

$$u(r) = 1$$



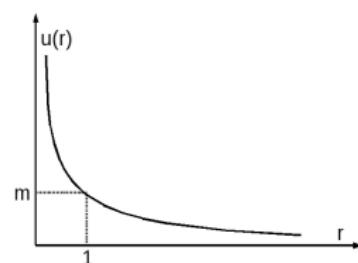
Huber's M -estimator:

$$u(r) = \begin{cases} K/e & \text{if } r \leq e \\ K/r & \text{if } r > e \end{cases}$$



FPE (Tyler):

$$u(r) = \frac{m}{r}$$



Remarks:

- Huber = mix between SCM and FPE,
- FPE and SCM are “not” (theoretically) M -estimators,
- FPE is the most robust while SCM is the most efficient.

Estimating the covariance matrix: Tyler's M -estimators

Let $(\mathbf{z}_1, \dots, \mathbf{z}_n)$ be a n -sample $\sim CE_m(\mathbf{0}, \Sigma, g_{\mathbf{z}})$ (**Secondary data**).

FP Estimate (Tyler, 1987; Pascal, 2008)

$$\widehat{\Sigma}_{FPE} = \frac{m}{n} \sum_{k=1}^n \frac{\mathbf{z}_k \mathbf{z}_k^H}{\mathbf{z}_k^H \widehat{\Sigma}_{FPE}^{-1} \mathbf{z}_k}$$

- The FPE does not depend on the texture (SIRV or CES distributions),
- Existence,
- Uniqueness,
- Convergence of the recursive algorithm,
- True MLE under SIRV distributed noise with unknown deterministic texture $\{\tau_k\}_{k \in [1, n]}$.

Asymptotic distribution of complex M -estimators

Using the results of Tyler, we derived the following results (Mahot, 2013):

Theorem 1 (Asymptotic distribution of $\hat{\Sigma}$)

$$\sqrt{n} \operatorname{vec}(\hat{\Sigma} - \Sigma) \xrightarrow{d} \mathcal{CN}_{m^2}(\mathbf{0}, \mathbf{C}, \mathbf{P}), \quad (5)$$

where \mathcal{CN} is the complex Gaussian distribution, \mathbf{C} the CM and \mathbf{P} the pseudo CM:

$$\begin{aligned}\mathbf{C} &= \sigma_1 (\Sigma^* \otimes \Sigma) + \sigma_2 \operatorname{vec}(\Sigma) \operatorname{vec}(\Sigma)^H, \\ \mathbf{P} &= \sigma_1 (\Sigma^* \otimes \Sigma) \mathbf{K} + \sigma_2 \operatorname{vec}(\Sigma) \operatorname{vec}(\Sigma)^T,\end{aligned}$$

where \mathbf{K} is the commutation matrix and where the constant σ_1 and σ_2 are completely defined.

An important property of complex M -estimators

- Let $\widehat{\Sigma}$ an estimate of Hermitian positive-definite matrix Σ that satisfies

$$\sqrt{n} \left(\text{vec}(\widehat{\Sigma} - \Sigma) \right) \xrightarrow{d} \mathcal{CN}(\mathbf{0}, \mathbf{C}, \mathbf{P}), \quad (6)$$

with

$$\begin{cases} \mathbf{C} = \nu_1 \Sigma^* \otimes \Sigma + \nu_2 \text{vec}(\Sigma) \text{vec}(\Sigma)^H, \\ \mathbf{P} = \nu_1 (\Sigma^* \otimes \Sigma) \mathbf{K}_{m,m} + \nu_2 \text{vec}(\Sigma) \text{vec}(\Sigma)^T, \end{cases}$$

where ν_1 and ν_2 are any real numbers.

e.g.

	SCM	M -estimators	FP
ν_1	1	σ_1	$(m+1)/m$
ν_2	0	σ_2	$-(m+1)/m^2$
...	More accurate		More robust

- Let $H(\cdot)$ be a r -multivariate function on the set of Hermitian positive-definite matrices, with continuous first partial derivatives and such as $H(\mathbf{V}) = H(\alpha \mathbf{V})$ for all $\alpha > 0$, e.g. **the ANMF statistic, the MUSIC statistic**, etc.

Theorem 2 (Asymptotic distribution of $H(\widehat{\Sigma})$)

$$\sqrt{n} \left(H(\widehat{\Sigma}) - H(\Sigma) \right) \xrightarrow{d} \mathcal{CN} (\mathbf{0}_{r,1}, \mathbf{C}_H, \mathbf{P}_H), \quad (7)$$

where \mathbf{C}_H and \mathbf{P}_H are defined as

$$\begin{aligned} \mathbf{C}_H &= \textcolor{red}{\nu_1} H'(\Sigma) (\Sigma^T \otimes \Sigma) H'(\Sigma)^H, \\ \mathbf{P}_H &= \textcolor{red}{\nu_1} H'(\Sigma) (\Sigma^T \otimes \Sigma) \mathbf{K}_{m,m} H'(\Sigma)^T, \end{aligned}$$

where $H'(\Sigma) = \left(\frac{\partial H(\Sigma)}{\partial \text{vec}(\Sigma)} \right)$.

CES distribution \Rightarrow two-step GLRT ANMF

ANMF test (ACE, GLRT-LQ) [Conte, 95, Kraut/Scharf 99]

$$H(\widehat{\Sigma}) = \Lambda_{ANMF}(\mathbf{z}, \widehat{\Sigma}) = \frac{|\mathbf{p}^H \widehat{\Sigma}^{-1} \mathbf{z}|^2}{(\mathbf{p}^H \widehat{\Sigma}^{-1} \mathbf{p})(\mathbf{z}^H \widehat{\Sigma}^{-1} \mathbf{z})} \stackrel{H_1}{\gtrless} \lambda_{ANMF}, \quad (8)$$

where $\widehat{\Sigma}$ stands for any M -estimators.

- The ANMF is **scale-invariant (homogeneous of degree 0)**, i.e.
 $\forall \alpha, \beta \in \mathbb{R}, \Lambda_{ANMF}(\alpha \mathbf{z}, \beta \widehat{\Sigma}) = \Lambda_{ANMF}(\mathbf{z}, \widehat{\Sigma}).$
- Its **asymptotic distribution** (conditionally to \mathbf{z} !) is known (F. Pascal and J.P. Ovarlez, IEEE-ICASSP 2015)

$$H(\widehat{\Sigma}) \xrightarrow{d} \mathcal{CN}\left(H(\Sigma), \frac{\sigma_1}{n} H(\Sigma) (H(\Sigma) - 1)^2\right).$$

- It is CFAR w.r.t the covariance/scatter matrix,
- It is CFAR w.r.t the texture.

Some comments:

Perfect (but asymptotic) characterization of several objects properties, such as detectors, classifiers, estimators, etc.

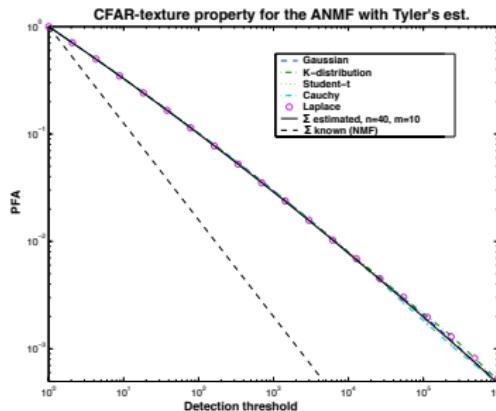
$H(SCM)$ and $H(M\text{-estimators})$ share the same asymptotic distribution (differs from σ_1).



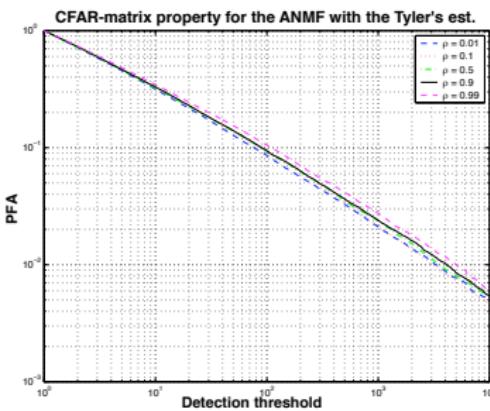
- Link to the classical Gaussian case,
- Quantification of the loss involved by robust estimator.

Illustration of the ANMF CFAR properties for CES process

False Alarm regulation for ANMF built with Tyler's estimate



(a) CFAR-texture



(b) CFAR-matrix

Figure: Illustration of the CFAR properties of the ANMF built with the Tyler's estimator, for a Toeplitz CM whose (i,j) -entries are $\rho^{|i-j|}$.

Probability of false alarm

PFA-threshold relation of $\Lambda_{ANMF}(\hat{S}_n)$ (Gaussian case, finite n)

$$P_{fa} = (1 - \lambda)^{a-1} {}_2F_1(a, a-1; b-1; \lambda), \quad (9)$$

where $a = n - m + 2$, $b = n + 2$ and ${}_2F_1$ is the Hypergeometric function defined as

$${}_2F_1(a, b; c; x) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{k=0}^{\infty} \frac{\Gamma(a+k)\Gamma(b+k)x^k}{\Gamma(c+k)} \frac{1}{k!}.$$

Probability of false alarm

For n large enough and for any elliptically distributed noise, the PFA is still given by (??) if we replace n by n/σ_1 .

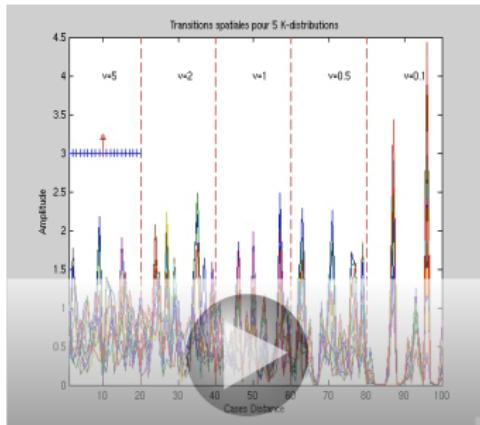
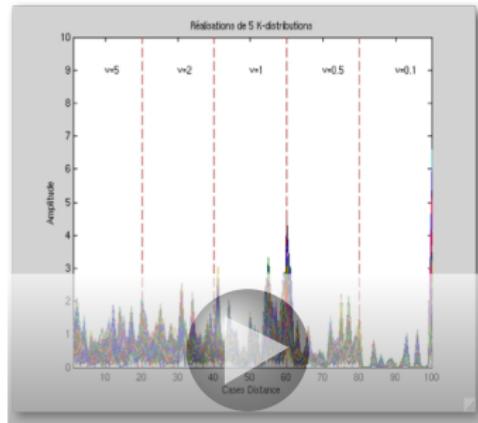
PFA-threshold relation of $\Lambda_{ANMF}(M\text{-est.})$ for CES distributions

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

where $a = \frac{n}{\sigma_1} - m + 2$, $b = \frac{n}{\sigma_1} + 2$ and ${}_2F_1$ is the Hypergeometric function.

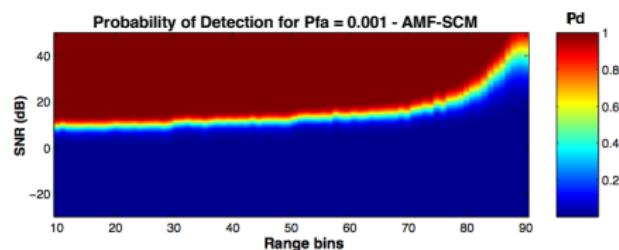
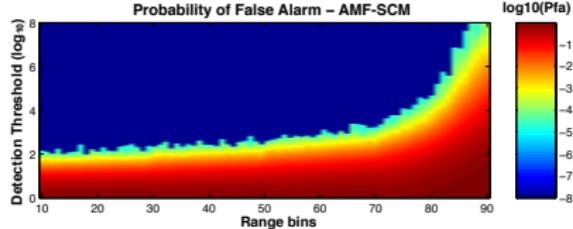
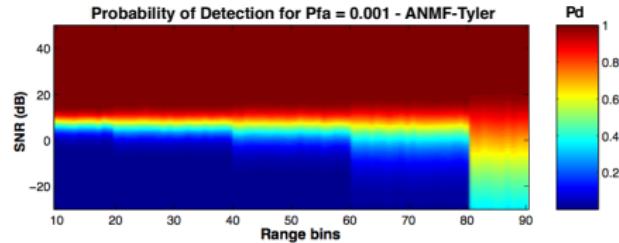
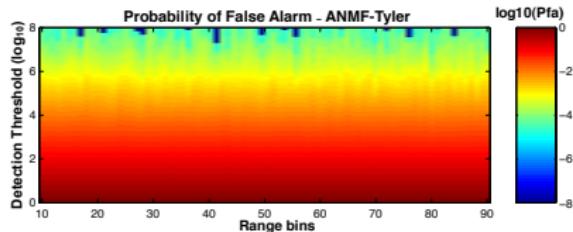
[5] F. Pascal, J.-P. Ovarlez, P. Forster, and P. Larzabal, "Constant false alarm rate detection in spherically invariant random processes," in *Proc. of the European Signal Processing Conf., EUSIPCO-04*, (Vienna), pp. 2143-2146, Sept. 2004.

Properties of ANMF-Tyler Detector on Clutter Transitions



- K-distributed clutter transitions: from Gaussian to impulsive noise,
- Estimation of the covariance matrix onto a range bins sliding window.

Properties of ANMF-Tyler Detector on Clutter Transitions



- ANMF-Tyler: The same detection threshold is guaranteed for a chosen P_{fa} whatever the clutter area,
- ANMF-Tyler: Performance in term of detection is kept for moderate non-Gaussian clutter and improved for spiky clutter.

Outline

- 1 Preliminaries**
 - Motivations
 - Some Background on Detection Theory
 - Case of Adaptive Gaussian Detection in Gaussian Background
- 2 Adaptive Robust Detection Schemes in non-Gaussian Background**
 - CES distributions
 - *M*-estimators and Tyler (FP) Estimator
 - Robustness of M-estimators and ANMF**
 - Exploiting Prior Information: Covariance Structure
 - Low Rank Detectors
 - Shrinkage of *M*-estimator
- 3 Applications**
 - Surveillance Radar
 - STAP Applications
- 4 Conclusions and Perspectives**

Robustness of the M-estimators

Let us suppose that $\{\mathbf{y}_i\}_{i=1,n-1} \sim \mathcal{CN}(\mathbf{0}, \Sigma)$ and that the last secondary data \mathbf{y}_n contains outlier \mathbf{p}_0 :

- Sample Covariance Matrix case:

$$\hat{\mathbf{S}}_n^{pol} = \frac{1}{n} \sum_{k=1}^{n-1} \mathbf{y}_k \mathbf{y}_k^H + \frac{1}{n} \mathbf{p}_0 \mathbf{p}_0^H \quad E \left[\hat{\mathbf{S}}_n^{pol} \right] = \frac{n-1}{n} \Sigma + \frac{1}{n} E \left[\mathbf{p}_0 \mathbf{p}_0^H \right]$$

The power of the outlier \mathbf{p}_0 has a big impact on the quality of the SCM estimation.

- Tyler (or FP) Covariance Matrix case:

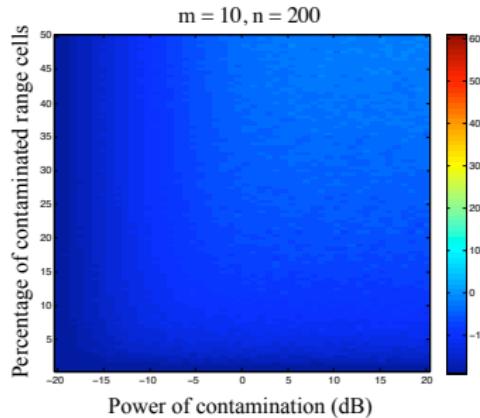
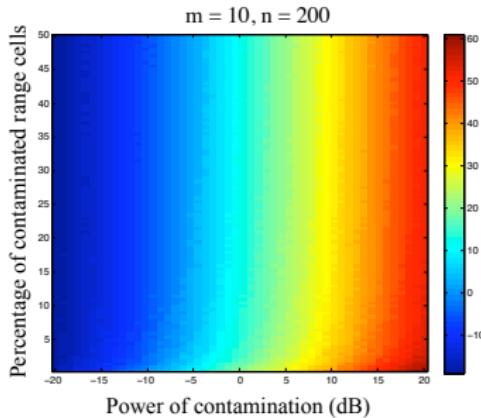
$$\hat{\Sigma}_{FPpol} = \frac{m}{n} \sum_{k=1}^n \frac{\mathbf{y}_k \mathbf{y}_k^H}{\mathbf{y}_k^H \hat{\Sigma}_{FPpol}^{-1} \mathbf{y}_k} \quad E \left[\hat{\Sigma}_{FPpol} \right] = \Sigma + \frac{m+1}{n} \left[E \left[\frac{\mathbf{p}_0 \mathbf{p}_0^H}{\mathbf{p}_0^H \Sigma^{-1} \mathbf{p}_0} \right] - \frac{1}{m} \Sigma \right]$$

The power of the outlier \mathbf{p}_0 has no big impact on the quality of the Tyler estimate.

Robustness of M-estimators

Gaussian vectors \mathbf{y}_k polluted by outliers

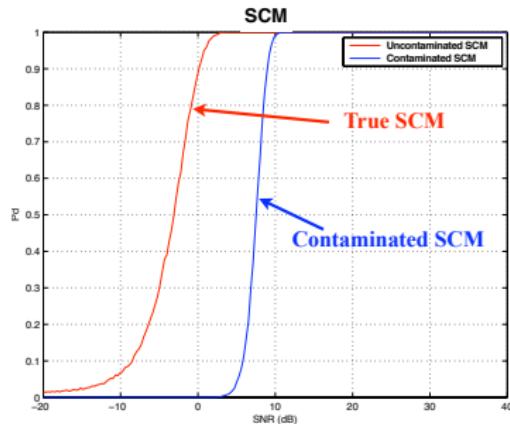
$$\hat{\mathbf{S}}_n = \frac{1}{n} \sum_{k=1}^n \mathbf{y}_k \mathbf{y}_k^H \quad \hat{\Sigma}_{FP} = \frac{m}{n} \sum_{k=1}^n \frac{\mathbf{y}_k \mathbf{y}_k^H}{\mathbf{y}_k^H \hat{\Sigma}_{FP}^{-1} \mathbf{y}_k}$$



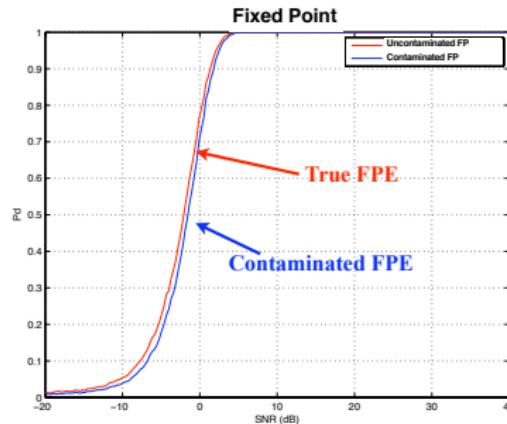
Plot of the error between the covariance matrix estimated with and without outliers.

Robustness of ANMF: Impact on detection performance

Same target $\mathbf{y}_k = \mathbf{p}_0$ (SNR 20dB) than those in the cell under test in the reference cells (case of convoy for example)



AMF + SCM



ANMF + FPE

The SCM can whiten the target to detect,
 The ANMF built with FPE is more robust.

Outline

- 1 Preliminaries**
 - Motivations
 - Some Background on Detection Theory
 - Case of Adaptive Gaussian Detection in Gaussian Background
- 2 Adaptive Robust Detection Schemes in non-Gaussian Background**
 - CES distributions
 - M -estimators and Tyler (FP) Estimator
 - Robustness of M -estimators and ANMF
 - Exploiting Prior Information: Covariance Structure**
 - Low Rank Detectors
 - Shrinkage of M -estimator
- 3 Applications**
 - Surveillance Radar
 - STAP Applications
- 4 Conclusions and Perspectives**

Motivations

The estimation of Σ does not take into account any prior knowledge on the covariance matrix:

How to improve detection performance by exploiting prior information on Σ ?

⇒ Use of some prior knowledge on the structure of the covariance matrix:

- Toeplitz: Burg [1982] for estimation, Furhmann [1991] for detection in Gaussian case,
- known rank $r < m$ (ex: subspace detector),
- **Persymmetry**: Nitzberg [1980] for estimation, Kai-Wang [1992] for detection in Gaussian case, Conte and De Maio [2003, 2004], Pailloux *et al.* [2010] in non-Gaussian noise.

Using Persymmetry Property

Under persymmetric considerations (ex: symmetrically spaced linear array, symmetrically spaced pulse train, ...), the Hermitian covariance matrix Σ verifies: $\Sigma = \mathbf{J}_m \Sigma^* \mathbf{J}_m$, where \mathbf{J}_m is the m -dimensional antidiagonal matrix having 1 as non-zero elements. If the unitary matrix \mathbf{T} is defined by:

$$\mathbf{T} = \begin{cases} \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{I}_{m/2} & \mathbf{J}_{m/2} \\ i\mathbf{I}_{m/2} & -i\mathbf{J}_{m/2} \end{pmatrix} & \text{for } m \text{ even} \\ \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{I}_{(m-1)/2} & 0 & \mathbf{J}_{(m-1)/2} \\ 0 & \sqrt{2} & 0 \\ i\mathbf{I}_{(m-1)/2} & 0 & -i\mathbf{J}_{(m-1)/2} \end{pmatrix} & \text{for } m \text{ odd}, \end{cases} \quad (11)$$

then:

- $\mathbf{s} = \mathbf{T}\mathbf{p}$ is a real vector (if \mathbf{p} is centrosymmetric, i.e. $\mathbf{p} = \mathbf{J}_m \mathbf{p}^*$),
- $\mathbf{R} = \mathbf{T}\Sigma\mathbf{T}^H$ is a real symmetric matrix.

Equivalent Detection Problem

Using previous transformation \mathbf{T} , the original problem can be reformulated as:

Original Problem	\mathbf{T}	Equivalent Problem
$\begin{cases} H_0 : \mathbf{y} = \mathbf{c}, & \mathbf{c}_1, \dots, \mathbf{c}_n \\ H_1 : \mathbf{y} = \mathbf{A}\mathbf{p} + \mathbf{c}, & \mathbf{c}_1, \dots, \mathbf{c}_n \end{cases}$	\rightarrow	$\begin{cases} H_0 : \mathbf{z} = \mathbf{n}, & \mathbf{n}_1, \dots, \mathbf{n}_n \\ H_1 : \mathbf{z} = \mathbf{A}\mathbf{s} + \mathbf{n}, & \mathbf{n}_1, \dots, \mathbf{n}_n \end{cases}$

where

- $\mathbf{z} = \mathbf{T}\mathbf{y} \in \mathbb{C}^m$,
- $\mathbf{n} = \sqrt{\tau}\mathbf{x}$ and $\mathbf{n}_k = \sqrt{\tau_k}\mathbf{x}_k$ with $\mathbf{x}, \mathbf{x}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{R})$ where \mathbf{R} is an unknown real symmetric matrix,
- $\mathbf{s} = \mathbf{T}\mathbf{p}$ is a real vector.

The main motivation for introducing the transformed data is that the original persymmetric complex covariance matrix of the Gaussian speckle Σ is transformed through \mathbf{T} onto a real covariance matrix \mathbf{R} .

The Persymmetric FP Covariance Matrix Estimate

From the estimate $\widehat{\mathbf{R}}_{FP}$ of the real covariance matrix \mathbf{R} , solution of the following equation:

$$\widehat{\mathbf{R}} = \frac{m}{n} \sum_{k=1}^n \frac{\mathbf{n}_k \mathbf{n}_k^H}{\mathbf{n}_k^H \widehat{\mathbf{R}}^{-1} \mathbf{n}_k},$$

the Persymmetric Fixed-Point Covariance Matrix Estimate can be defined as:

$$\widehat{\mathbf{R}}_{PFP} = \mathcal{R}e(\widehat{\mathbf{R}}_{FP}).$$

Statistical performance of $\widehat{\mathbf{R}}_{PFP}$ [Pascal et al. 2008]:

- $\widehat{\mathbf{R}}_{PFP}$ is a consistent estimate of \mathbf{R} when n tends to infinity,
- $\widehat{\mathbf{R}}_{PFP}$ is an unbiased estimate of \mathbf{R} ,
- Its asymptotic distribution is the same as the asymptotic distribution of a real Wishart matrix with $\frac{m}{m+1} 2n$ degrees of freedom.

The Persymmetric Adaptive Normalized Matched Filter

The resulting P-ANMF for the transformed problem is based on the PFP estimate and can be defined as:

$$\Lambda(\widehat{\mathbf{R}}_{PFP}) = \frac{|\mathbf{s}^\top \widehat{\mathbf{R}}_{PFP}^{-1} \mathbf{z}|^2}{(\mathbf{s}^\top \widehat{\mathbf{R}}_{PFP}^{-1} \mathbf{s})(\mathbf{z}^\top \widehat{\mathbf{R}}_{PFP}^{-1} \mathbf{z})} \stackrel{H_1}{\gtrless} \stackrel{H_0}{\lessdot} \lambda. \quad (12)$$

Properties:

- $\Lambda(\widehat{\mathbf{R}}_{PFP})$ is texture-CFAR,
- $\Lambda(\widehat{\mathbf{R}}_{PFP})$ is matrix-CFAR,
- The use of PFP estimate in the ANMF allows to **virtually double the number n of secondary data** and improve the performance of the ANMF detector built with the FP matrix estimate.

$\Lambda(\widehat{\mathbf{R}}_{PFP})$ is SIRV-CFAR and is called the P-ANMF.

Statistical study of the P-ANMF

The analytical expression for the Probability Density Function of the test statistic $\Lambda(\widehat{\mathbf{R}}_{PFP})$ is really not easy to derive in a closed form but the following results gives some insight about its distribution.

$\Lambda(\widehat{\mathbf{R}}_{PFP})$ has the same distribution as $\frac{F}{F+1}$ where

$$F = \frac{(\alpha_1 u_{22} - \alpha_2 u_{21})^2 + \left(1 + \left(\frac{\beta_3}{u_{33}}\right)^2\right) (a u_{22} - b u_{21})^2}{(\alpha_2 u_{11})^2 + \left(t_{11} u_{22} \frac{\beta_3}{u_{33}}\right)^2 + u_{11}^2 \left(1 + \left(\frac{\beta_3}{u_{33}}\right)^2\right) b^2} \quad (13)$$

and where: $a, b, \alpha_1, u_{21} \sim \mathcal{N}(0, 1)$, $\alpha_2^2 \sim \chi^2_{m-1}$, $\beta_3^2 \sim \chi^2_{m-2}$, $u_{11}^2 \sim \chi^2_{n'-m+1}$, $u_{22}^2 \sim \chi^2_{n'-m+2}$, $u_{33}^2 \sim \chi^2_{n'-m+3}$ with $n' = \frac{m}{m+1} 2n$.

Outline

- 1 Preliminaries**
 - Motivations
 - Some Background on Detection Theory
 - Case of Adaptive Gaussian Detection in Gaussian Background
- 2 Adaptive Robust Detection Schemes in non-Gaussian Background**
 - CES distributions
 - *M*-estimators and Tyler (FP) Estimator
 - Robustness of *M*-estimators and ANMF
 - Exploiting Prior Information: Covariance Structure
 - Low Rank Detectors**
 - Shrinkage of *M*-estimator
- 3 Applications**
 - Surveillance Radar
 - STAP Applications
- 4 Conclusions and Perspectives**

Conventional Low Rank Detectors

Principle of Low Rank Matched Filter approaches found for example in [Kirstein *et al.*, 94] (Principal Component Inverse) and [Haimovich, 96] (Eigencanceler) and [Rangaswami *et al.*, 04].

Let suppose the rank r of clutter covariance matrix Σ is known:

- Example of sidelooking STAP with M pulses measurements and N sensors,
 $r = N + (M - 1)\beta$ (Brennan's rule) where $\beta = 2\nu T_r/d$.

The idea is to **project the data onto the orthogonal subspace of the clutter**.

$$\hat{\Sigma}_n = \frac{1}{n} \sum_{k=1}^n \mathbf{y}_k \mathbf{y}_k^H = (\mathbf{U}_r \mathbf{U}_0) \begin{pmatrix} \Sigma_r & \mathbf{0} \\ \mathbf{0} & \Sigma_0 \end{pmatrix} (\mathbf{U}_r \mathbf{U}_0)^H,$$

If we denote by $\Pi_{SCM} = \mathbf{U}_r \mathbf{U}_r^H$ the projector onto the clutter subspace, the Low-Rank ANMF detector is given by:

$$\Lambda_{LR-ANMF-SCM}(\mathbf{z}) = \frac{|\mathbf{p}^H (\mathbf{I} - \Pi_{SCM}) \mathbf{z}|^2}{(\mathbf{p}^H (\mathbf{I} - \Pi_{SCM}) \mathbf{p})(\mathbf{z}^H (\mathbf{I} - \Pi_{SCM}) \mathbf{z})} \stackrel{H_1}{\underset{H_0}{\gtrless}} \lambda.$$

Extended Low Rank Detectors

In a case of heterogeneous and non-Gaussian clutter, we know that $\hat{\Sigma}_{SCM}$ or Π_{SCM} are not good estimates. If we denote the Normalized Sample Covariance Matrix by:

$$\Sigma_{NSCM} = \frac{NM}{n} \sum_{k=1}^n \frac{\mathbf{y}_k \mathbf{y}_k^H}{\mathbf{y}_k^H \mathbf{y}_k} = (\mathbf{U}_r \mathbf{U}_0) \begin{pmatrix} \Sigma_r & \mathbf{0} \\ \mathbf{0} & \Sigma_0 \end{pmatrix} (\mathbf{U}_r \mathbf{U}_0)^H$$

[Ginolhac et al., 12] proved that $\Pi_{NSCM} = \mathbf{U}_r \mathbf{U}_r^H$ is a consistent estimate projector onto the clutter subspace. We can define the extended Low-Rank ANMF-NSCM:

$$\Lambda_{LR-ANMF-NSCM}(\mathbf{y}) = \frac{|\mathbf{p}^H (\mathbf{I} - \Pi_{NSCM}) \mathbf{z}|^2}{(\mathbf{p}^H (\mathbf{I} - \Pi_{NSCM}) \mathbf{p})(\mathbf{z}^H (\mathbf{I} - \Pi_{NSCM}) \mathbf{z})} \stackrel{H_1}{\gtrless} \lambda.$$

This detector is found to be **texture-CFAR** and is **asymptotically Σ -CFAR**. Moreover, he has another nice **robustness property** when outliers and targets are present in the secondary data. The Normalized Sample Covariance Matrix is a good candidate for adaptive version of Rangaswami's Low Rank Matched Filter and Low Rank Normalized Matched Filter.

Extended ML Low Rank Detectors

When the texture is assumed to be deterministic and unknown, this problem can be addressed by deriving the exact clutter subspace projector estimation [A. Breloy, 2016]. The problem to solve can thus be stated as an optimization problem on the exact likelihood:

$$\arg \min_{\Sigma, \{\tau_k\}} -\log(f(\{\mathbf{z}_k\} | \Sigma, \{\tau_k\})) = \arg \min_{\Sigma, \{\tau_k\}} -\log \left(\prod_{k=1}^n \frac{\exp(-\mathbf{z}_k^H (\tau_k \Sigma + \sigma^2 \mathbf{I})^{-1} \mathbf{z}_k)}{\pi^m |\tau_k \Sigma + \sigma^2 \mathbf{I}|} \right)$$

under constraints $\begin{cases} \text{rank}(\Sigma) = r \\ \Sigma \succeq \mathbf{0} \\ \tau_k > 0 \end{cases}$.

The problem is not convex: a solution can be proposed and consists in analysing

$$\Sigma = \sum_{i=1}^r c_i \mathbf{v}_i \mathbf{v}_i^H \text{ and estimating the MLE of } \{c_i\}, \{\mathbf{v}_i\} \text{ and } \{\tau_i\}.$$

- When no enough secondary data are available (undersampled case, $n < m$), this procedure can be applied on a regularized M -estimators.

Outline

- 1 Preliminaries**
 - Motivations
 - Some Background on Detection Theory
 - Case of Adaptive Gaussian Detection in Gaussian Background
- 2 Adaptive Robust Detection Schemes in non-Gaussian Background**
 - CES distributions
 - *M*-estimators and Tyler (FP) Estimator
 - Robustness of *M*-estimators and ANMF
 - Exploiting Prior Information: Covariance Structure
 - Low Rank Detectors
 - Shrinkage of *M*-estimator**
- 3 Applications**
 - Surveillance Radar
 - STAP Applications
- 4 Conclusions and Perspectives**

Shrinkage of Tyler's estimators

Case of small number of observations or under-sampling $n < m$: matrix is not invertible \Rightarrow Problem when using M -estimators or Tyler's estimator!

Chen estimator

$$\widehat{\Sigma}_C = (1 - \beta) \frac{m}{n} \sum_{i=1}^n \frac{\mathbf{z}_i \mathbf{z}_i^H}{\mathbf{z}_i^H \widehat{\Sigma}_C^{-1} \mathbf{z}_i} + \beta \mathbf{I}$$

subject to the constraint $\text{Tr}(\widehat{\Sigma}) = m$ and for $\beta \in (0, 1]$.

- Originally introduced in [Y. Abramovich *et al.*, IEEE-ICASSP-07],
- Existence, uniqueness and algorithm convergence proved in [Y. Chen, A. Wiesel, and A. O. Hero, IEEE-TSP 2011],
- Active research [Y. Abramovich, O. Besson, R. Couillet, M. McKay, A. Wiesel, F. Pascal].

Shrinkage Tyler's estimators

Pascal estimator [F. Pascal et al, IEEE-TSP 2014]

$$\widehat{\Sigma}_P = (1 - \beta) \frac{m}{n} \sum_{i=1}^n \frac{\mathbf{z}_i \mathbf{z}_i^H}{\mathbf{z}_i^H \widehat{\Sigma}_P^{-1} \mathbf{z}_i} + \beta \mathbf{I}$$

subject to the **no** trace constraint but for $\beta \in (\bar{\beta}, 1]$, where
 $\bar{\beta} := \max(0, 1 - n/m)$.

- $\widehat{\Sigma}_P$ (naturally) verifies $\text{Tr}(\widehat{\Sigma}_P^{-1}) = m$ for all $\beta \in (0, 1]$,
- Existence, uniqueness and algorithm convergence proved,
- The main challenge is to find the optimal β ! [R. Couillet and M. R. McKay, 2015].

Outline

1 Preliminaries

- Motivations
- Some Background on Detection Theory
- Case of Adaptive Gaussian Detection in Gaussian Background

2 Adaptive Robust Detection Schemes in non-Gaussian Background

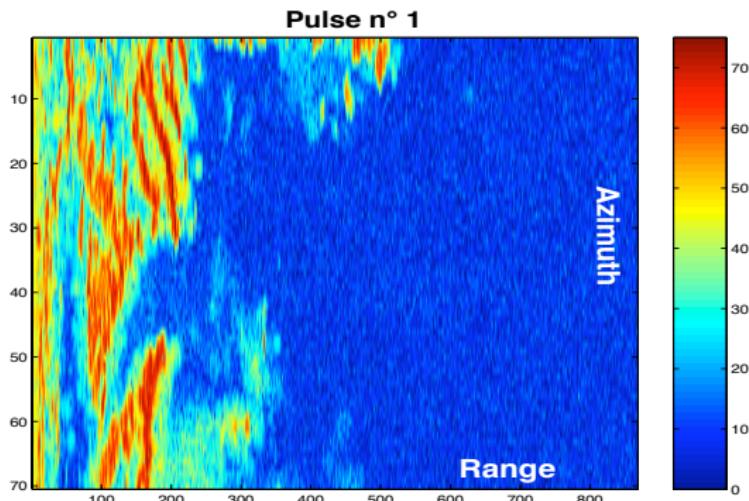
- CES distributions
- M -estimators and Tyler (FP) Estimator
- Robustness of M-estimators and ANMF
- Exploiting Prior Information: Covariance Structure
- Low Rank Detectors
- Shrinkage of M -estimator

3 Applications

- Surveillance Radar
- STAP Applications

4 Conclusions and Perspectives

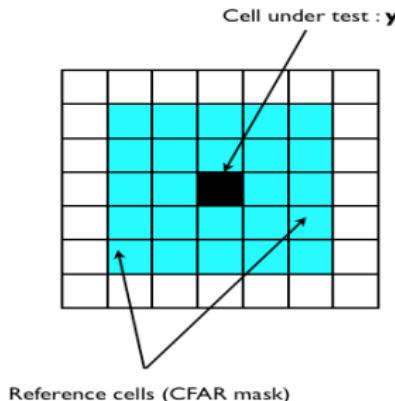
Data Description



- "Range-azimuth" map from ground clutter data collected by a radar from THALES Air Defence, placed 13 meters above ground and illuminating area at low grazing angle.
- Ground clutter complex echoes collected in 868 range bins for 70 different azimuth angles and for $m = 8$ pulses.

Data processing

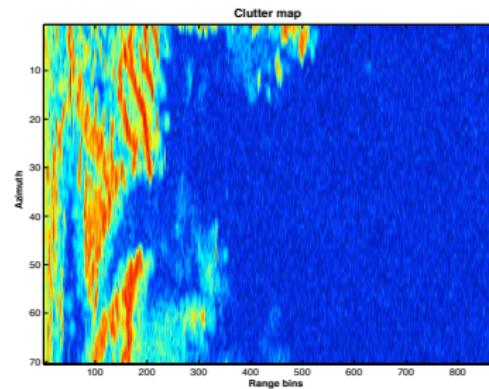
- Rectangular CFAR mask 5×5 for $0 \leq k \leq m$ different steering vectors \mathbf{p}_k .



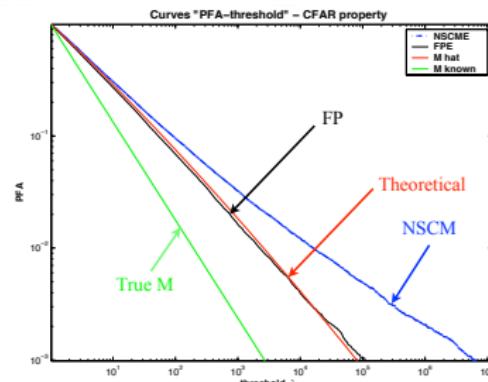
$$\mathbf{p}_k = \begin{pmatrix} 1 \\ \exp\left(\frac{2i\pi(k-1)}{m}\right) \\ \exp\left(\frac{2i\pi(k-1)2}{m}\right) \\ \vdots \\ \exp\left(\frac{2i\pi(k-1)(m-1)}{m}\right) \end{pmatrix}$$

- For each y , computation of associated detectors $\Lambda_{ANMF}(\widehat{\Sigma}_{Tyler})$ and $\Lambda_{ANMF}(\widehat{\Sigma}_{NSCM})$
- Mask moving all over the map.

False Alarm Regulation Results on Experimental Data (Surveillance Radar)



Azimut/range bins map



Relationship " P_{fa} -threshold"

Figure: False alarm regulation for $p_0 = (1 \dots 1)^T$.

Black curve fits red curve until $PFA = 10^{-3}$.

Outline

1 Preliminaries

- Motivations
- Some Background on Detection Theory
- Case of Adaptive Gaussian Detection in Gaussian Background

2 Adaptive Robust Detection Schemes in non-Gaussian Background

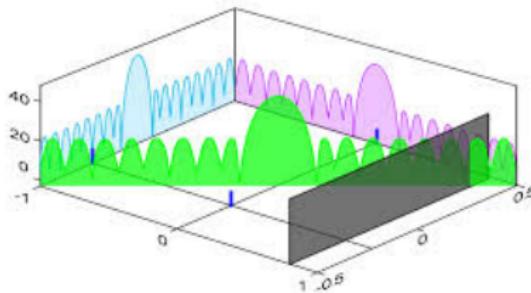
- CES distributions
- M -estimators and Tyler (FP) Estimator
- Robustness of M-estimators and ANMF
- Exploiting Prior Information: Covariance Structure
- Low Rank Detectors
- Shrinkage of M -estimator

3 Applications

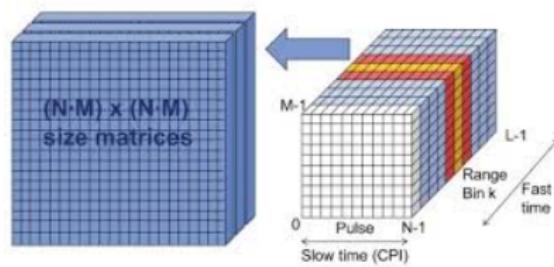
- Surveillance Radar
- **STAP Applications**

4 Conclusions and Perspectives

Space Time Adaptive Processing: Principles



(a) STAP principles



(b) STAP datacube

$$\mathbf{p}(\theta, f_d) = \begin{pmatrix} 1 \\ \exp(-2i\pi d \sin(\theta)/\lambda) \\ \vdots \\ \exp(-2i\pi(N-1)d \sin(\theta)/\lambda) \end{pmatrix} \otimes \begin{pmatrix} 1 \\ \exp(-2i\pi f_d T_r) \\ \vdots \\ \exp(-2i\pi f_d (M-1) T_r) \end{pmatrix}$$

STAP Principles

Problem: Using joint spatial and time measurements, estimate the position (angle) and the Doppler frequency (speed) of the target
⇒ use of the ANMF with a particular steering vector

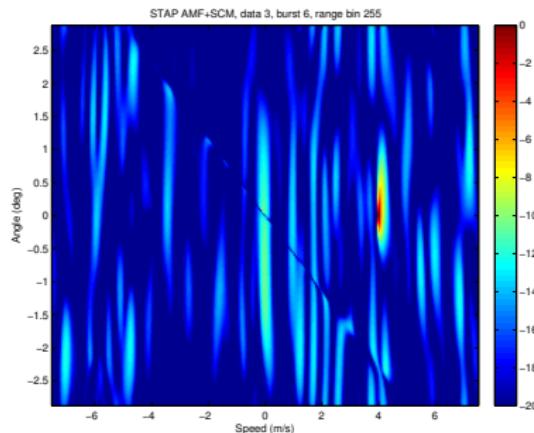
Data parameters: real clutter with synthetic target

X-Band $\simeq 10^9$ Hz, wavelength $\lambda = 0.03\text{m}$, flight speed $v = 100\text{m/s}$, distance to the scene 30km, 5 deg of incidence, PRF (Pulse Repetition Frequency) of 1 kHz, inter-sensor distance $d = 0.3\text{m}$, 12 trials with $n = 410$ range bins, $M = 64$ pulses and $N = 4$ sensors.

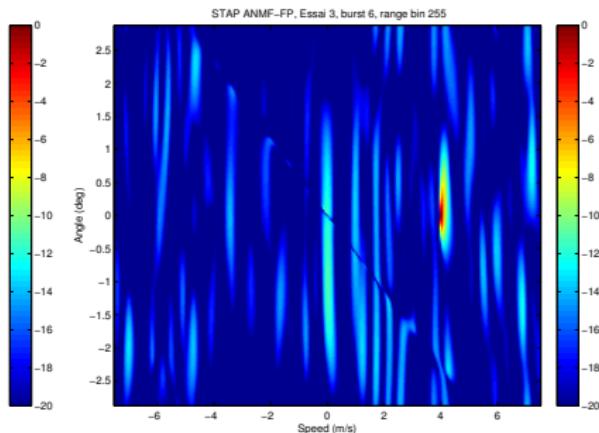
This means observations of size $m = 256$ while $n \leq 410$!

Clutter more or less homogeneous **BUT** some targets (outliers) could be present in the secondary data

No target is present in the secondary data - homogeneous noise



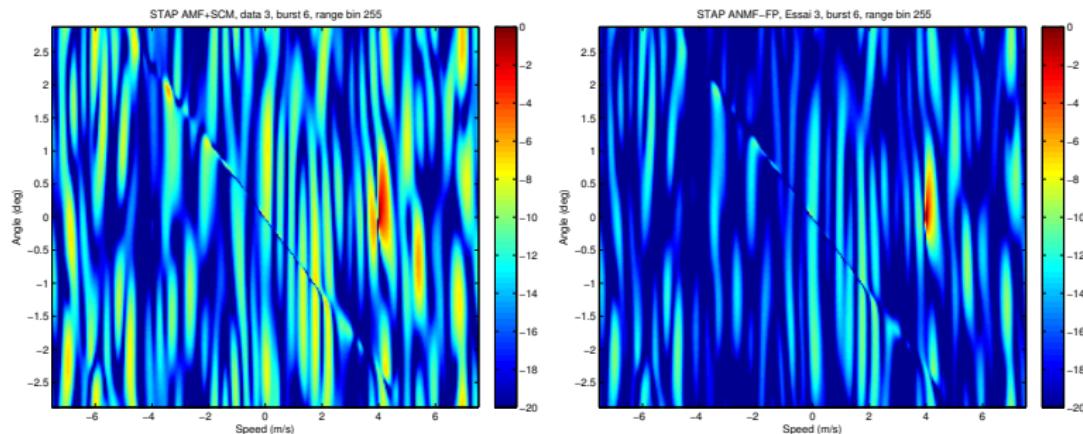
(c) AMF detector with the SCM



(d) ANMF detector with Tyler's est.

Figure: Doppler-angle map for the range bin 255 with $n = 404$ secondary data (targets and guard cells are removed) and $m = 256$

Two targets (4m/s and -4m/s) are present in the secondary data -
homogeneous noise



(a) AMF detector with the SCM (b) ANMF detector with Tyler's est.

Figure: Doppler-angle map for the range bin 255 with $n = 404$ secondary data
(guard cells are removed) and $m = 256$

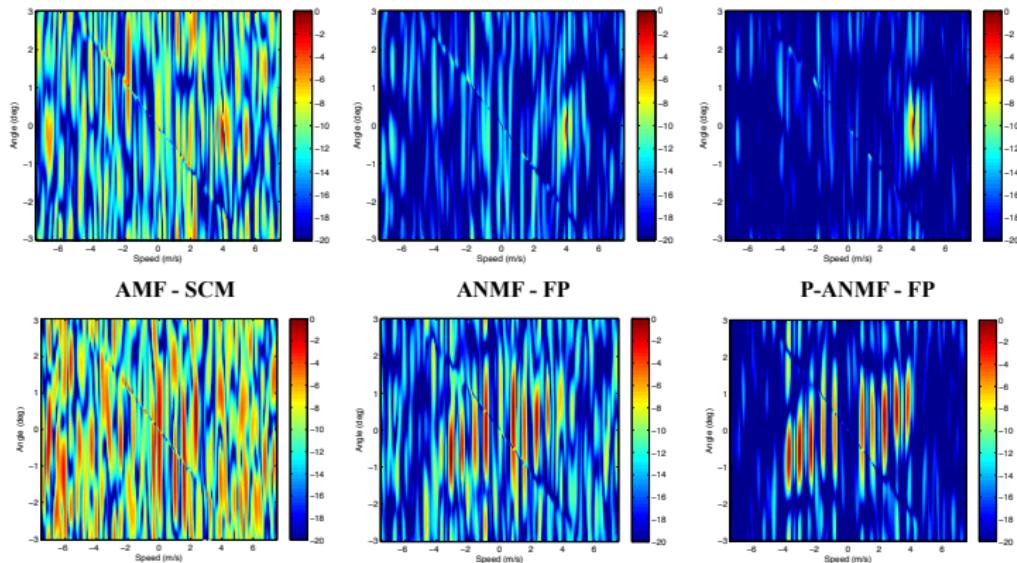


Figure: Doppler-angle map for the range bin 255 with $n = 404$ secondary data (guard cells are removed) and $m = 256$

Extended Low Rank Detectors

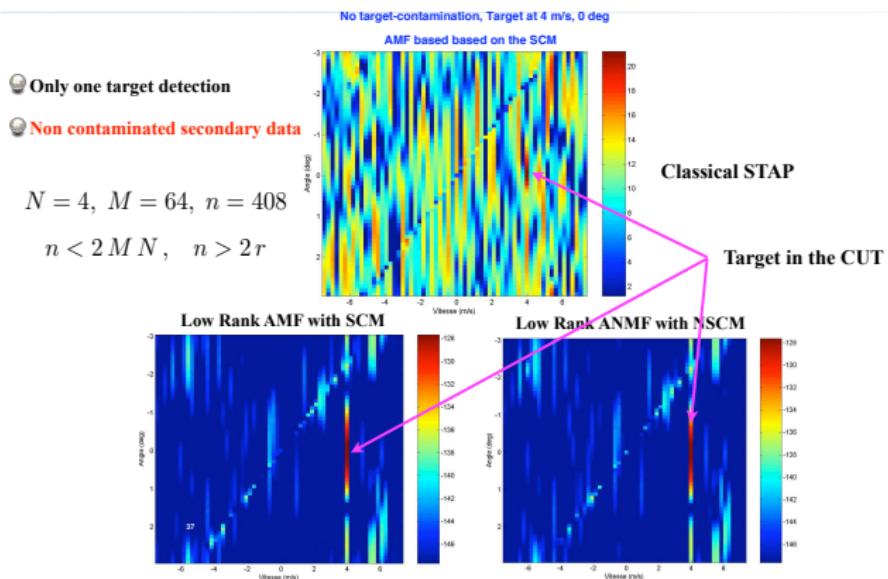


Figure: Doppler-angle map for the range bin 255 with $n = 408$ secondary data (guard cells are removed) and $m = 256$

Extended Low Rank Detectors

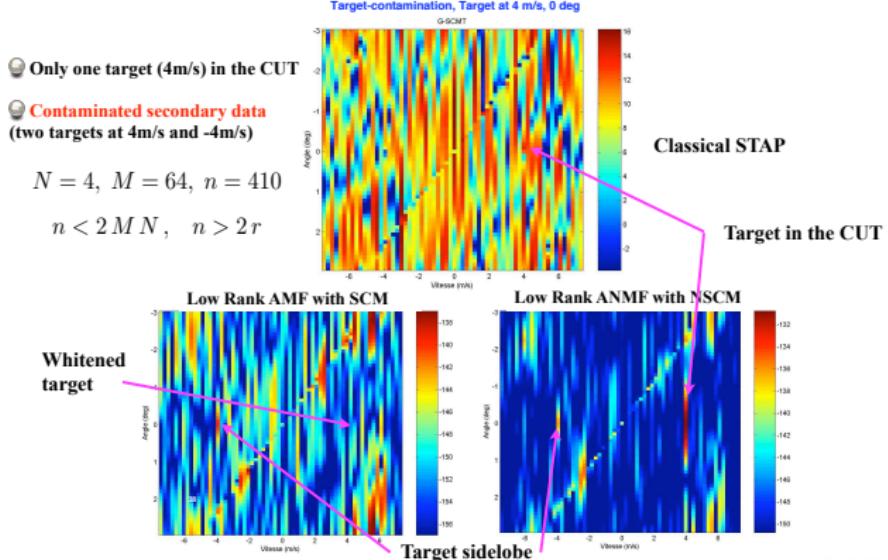


Figure: Doppler-angle map for the range bin 255 with $n = 410$ secondary data (guard cells are removed) and $m = 256$

Extended ML Low Rank Detectors

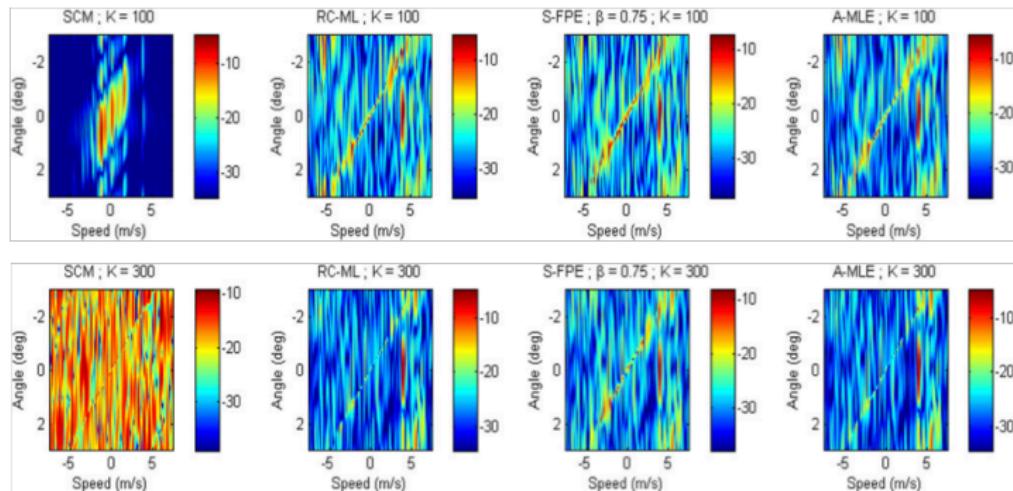
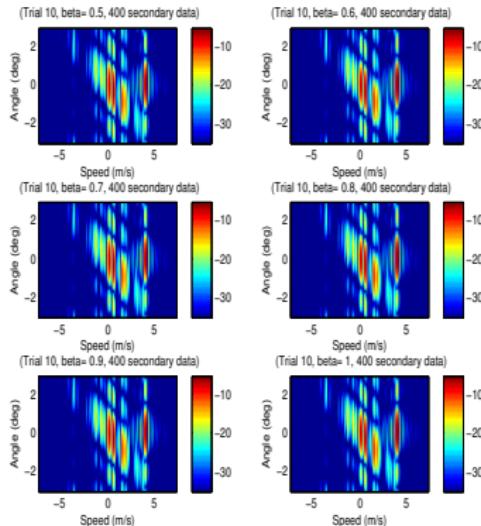


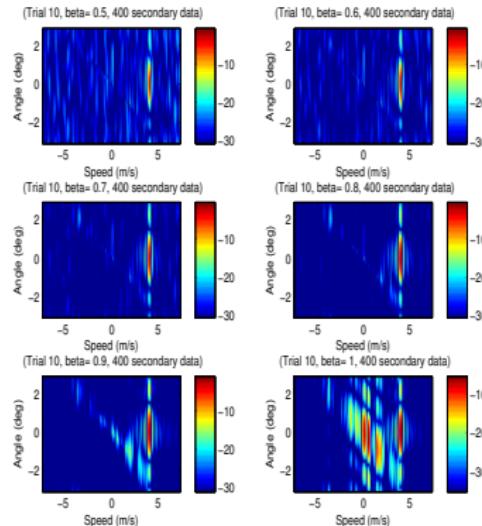
Figure: Comparison of various STAP detectors (Clubstap dataset) for two sets of secondary data ($n = K = 100$ and $n = K = 300$): SCM, RC-ML from [B. Kang, V. Monga, and M. Rangaswamy, TAES 2014] and S-FPE/A-MLE from [A. Breloy, IEEE-TSP 2016]

Application of Shrinkage to STAP

Applications to STAP data for \neq values of β , $m = 256$ and $n = 400$



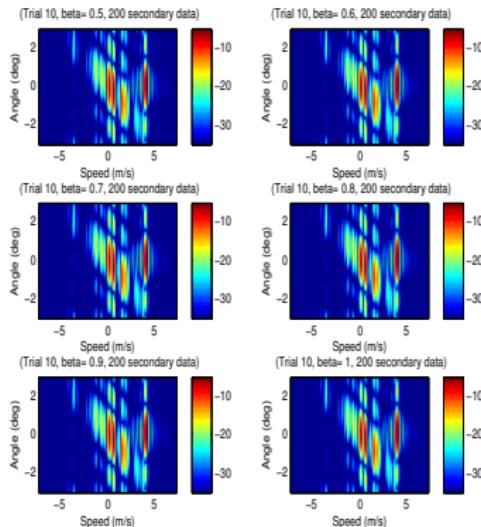
(a) SCM



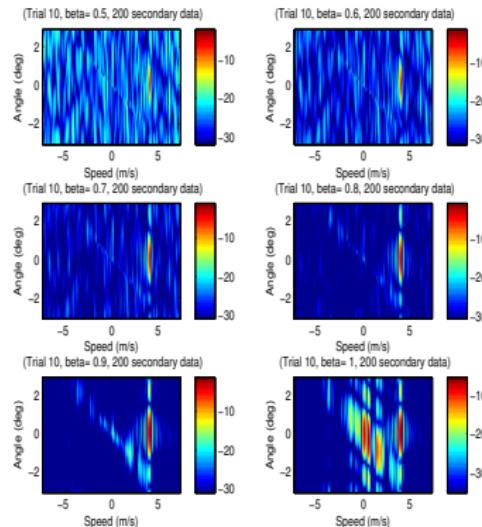
(b) Shrinkage FPE

Application of Shrinkage to STAP

Applications to STAP data for \neq values of β , $m = 256$ and $n = 200 \leq m$



(c) SCM



(d) Shrinkage FPE

Outline

- 1 Preliminaries
 - Motivations
 - Some Background on Detection Theory
 - Case of Adaptive Gaussian Detection in Gaussian Background
- 2 Adaptive Robust Detection Schemes in non-Gaussian Background
 - CES distributions
 - M -estimators and Tyler (FP) Estimator
 - Robustness of M-estimators and ANMF
 - Exploiting Prior Information: Covariance Structure
 - Low Rank Detectors
 - Shrinkage of M -estimator
- 3 Applications
 - Surveillance Radar
 - STAP Applications
- 4 Conclusions and Perspectives

Conclusions

When the background is non-Gaussian and/or heterogeneous, the conventional detectors (AMF or sub-optimal CFAR tests) are not at all optimal and lead to poor false alarm regulation and poor detection performance,

The SIRV and CES background modeling allows to take into account the background complexity: the non-Gaussianity, the temporal background fluctuations and the spatial background power fluctuations,

Using this model, the ANMF detector built with the Fixed Point (or other M-estimators) background covariance matrix estimator is shown to be CFAR-texture, CFAR-matrix and exhibits nice properties (robustness) and very good detection performance,

Conclusions

Taking into account additional *a priori* properties on the covariance matrix structure (low rank, persymmetry, Toeplitz, ...) can lead to a appreciable gain for small numbers of secondary data,

These methods have been applied for many problems involving covariance matrix estimation: STAP detection, SAR detection (FOPEN), Polarimetric/Interferometric SAR detection and classification, SAR and Hyperspectral Change Detection, SAR and Hyperspectral time-series analysis, Hyperspectral Anomaly detection, Hyperspectral detection.

On-going works and Perspectives

Link with **Random Matrix Theory**: for high dimensionality data (ex: hyperspectral, STAP), strong statistical connexion with Robust Estimation theory: see current works of R. Couillet, and F. Pascal,

Robust estimation of structured covariances matrices [Y. Sun, D. P. Palomar, A. Breloy, G. Ginolhac, F. Pascal, P. Forster],

Joint location and scale with **M-Estimators** (non-centered multivariate data, e.g. hyperspectral data) [J. Frontera, F. Pascal, J.P. Ovarlez],

How to deal with non i.i.d secondary data? RMT approach: [R. Couillet, F. Pascal, J.P. Ovarlez], VARMA approach: [W. Ben-Abdallah, P. Bondon, J.P. Ovarlez],

On-going works and Perspectives

No secondary data: [C. Ren, N. El-Korso, P. Forster, A. Breloy, J.P. Ovarlez],

M-Estimators and Riemannian Geometry: [F. Barbaresco], [F. Pascal, G. Ginolhac, A. Renaux],

Outliers: [C. Culan, C. Adnet],

Shrinkage of M-Estimators: [A. Wiesel, Y. Abramovitch, O. Besson, F. Pascal, E. Ollila, ...], [Q. Hoarau, G. Ginolhac],

Sparsity and high dimension: [A. Bitar, J.P. Ovarlez].

Acknowledgements

- Frédéric Pascal, LSS CentraleSupelec, Gif sur Yvette, France,
- Philippe Forster, ENS Cachan, France,
- Guillaume Ginolhac, Annecy University, France,
- and former PhD Students: M. Mahot, P. Formont, J. Frontera-Pons, A. Breloy, ...

References

Many references relative to this seminar can be found on my homepage:

<http://www.jeanphilippeovarlez.com>