

A Tutorial on the Estimation and Detection Theory with Applications to Radar, Array Processing, and Imaging

Jean-Philippe Ovarlez^{1,2}

¹SONDRA, CentraleSupélec, France

²ONERA, The French Aerospace Lab, DEMR/MATS, France

jeanphilippe.ovarlez@centralesupelec.fr
<https://www.jeanphilippeovarlez.com>

T03 Tutorial - Conference Radar 2024 - 21 October 2024



Plan

1 Motivations and General Introduction

- Radar and Imaging Sensors
- Radar and Imaging Sensors - New challenges
- Applicative Context
- Methodological Context

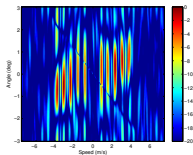
2 Tutorial Description

Radar and Imaging Sensors

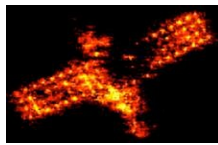
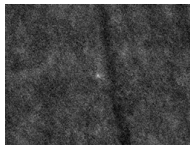
RADAR = **RA**dio **D**etection **A**nd **R**anging



- emits and receives electromagnetic waves,
- detects the presence of targets,



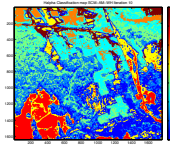
Detection maps



ISAR Image



SAR Image



SAR Classification

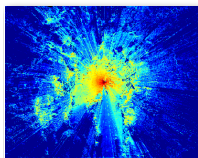
- but also: estimates parameters (range, radial velocity, angles of presentation, acceleration, amplitude (related to Radar Cross Section), etc.),
- images, classifies, recognizes.

Note: Almost all the conventional Statistical Signal Processing methodologies and background modeling tools are based on the Gaussian hypothesis (**standard conditions**).

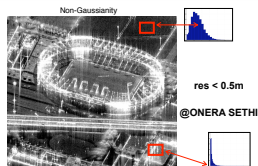
Radar and Imaging Sensors - New challenges

Positioning: facing the new **non-standard** conditions

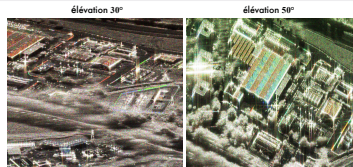
- **Complex Environments:** ground, dynamic environments (sea, ionosphere), heterogeneous, non-Gaussian, reverberating.
- **Complex targets:** small RCS, extended targets, fluctuating, dispersive, anisotropic, off-grid targets.
- **Sensor Diversity:** temporal, spatial, polarimetric, interferometric, spectral.
- **Improvement of sensor resolution:** spatial, spectral, angular.
- **Outliers, jamming**
- **Increase of the dimension and the size of signals to analyze.**



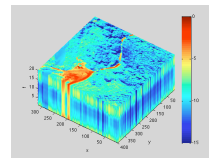
**Heterogeneous
Environments**



**Non-Gaussian
Environments**

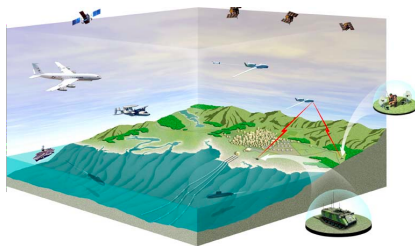


**Non-Stationary Targets
and Environments**



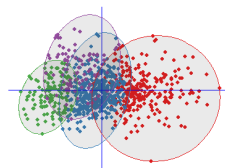
Curse of Dimensionality

Applicative Context



Air, ground, sea Surveillance

- ▣ Radar Detection, Space-Time Adaptive Processing
- ▣ Synthetic Aperture Radar
- ▣ Sources Localization
- ▣ Interferometric, Polarimetric Classification
- ▣ Change Detection, Infrastructure Monitoring
- ▣ Anomaly Detection in Hyperspectral Imaging
- ▣ MIMO Radar
- ▣ Tracking



Big Data

- ▣ Recognition
- ▣ Classification, Clustering
- ▣ Dimension Reduction
- ▣ Machine Learning, Deep Learning
- ▣ Graphs Analysis
- ▣ Learning Techniques

Finance

- ▣ Time Series
- ▣ Portfolio Optimization
- ▣ Risk Management
- ▣ Classification
- ▣ Prediction



Methodological Context

Goals: Improvement of sensors performance and their processing

- To model thanks **statistics** the variability of the unknown environment and data,
- To estimate the **spectral properties** of the environment (ionosphere, sea, wind through forest, etc.),
- To elaborate **estimators** and **detectors** that are *robust* and *adaptive* to these environments,
- To **regulate the False Alarm** on these *heterogeneous, non-stationary, non-Gaussian* environments,
- To **improve** the classification, the clustering techniques.

Methods: Statistical Signal Processing

- **Robust Estimation Techniques** of spectral and statistic characteristics of the environment and targets: adaptivity, statistic learning, cognitive, maximal exploitation of the *a priori*,
- **Optimal Detection Schemes** (Likelihood, Bayesian) for stealthy targets embedded in these complex environments,
- **Exploitation of emerging statistical Signal Processing techniques**: Minimal Estimation Bounds, Time-Frequency Analysis, Random Matrix Theory, Clustering, Compressive Sensing, Artificial Intelligence, Riemannian and Differential Geometry, etc.

Methodological Context

Goals: Improvement of sensors performance and their processing

- To model thanks **statistics** the variability of the unknown environment and data,
- To estimate the **spectral properties** of the environment (ionosphere, sea, wind through forest, etc.),
- To elaborate **estimators** and **detectors** that are *robust* and *adaptive* to these environments,
- To **regulate the False Alarm** on these *heterogeneous, non-stationary, non-Gaussian* environments,
- To **improve** the classification, the clustering techniques.

Methods: Statistical Signal Processing

- **Robust Estimation Techniques** of spectral and statistic characteristics of the environment and targets: adaptivity, statistic learning, cognitive, maximal exploitation of the *a priori*,
- **Optimal Detection Schemes** (Likelihood, Bayesian) for stealthy targets embedded in these complex environments,
- **Exploitation of emerging statistical Signal Processing techniques**: Minimal Estimation Bounds, Time-Frequency Analysis, Random Matrix Theory, Clustering, Compressive Sensing, Artificial Intelligence, Riemannian and Differential Geometry, etc.

Plan

1 Motivations and General Introduction

- Radar and Imaging Sensors
- Radar and Imaging Sensors - New challenges
- Applicative Context
- Methodological Context

2 Tutorial Description

General Introduction

Survey on

- Background on conventional Gaussian statistical modeling: Radar modeling, Random Noise Modeling, Maximum Likelihood Estimation, Detection Schemes, etc.
- Recent methodologies on more recent robust estimation and detection schemes: Complex Elliptically Symmetric distributions, M -Estimators,
- If time left: more advanced techniques: Robust COMET (RCOMET), Random Matrix Theory, Robust Low-Rank modeling, Riemannian Geometry, etc.

Two Main Parts

- **Part A:** Background on Statistical Processing for Radar, Array Processing, SAR and Hyperspectral Imaging,
- **Part B:** Recent Methodologies on Robust Estimation and Detection in non-Gaussian Environment - Applications and Results in Radar, STAP and Array Processing, SAR Imaging, Hyperspectral Imaging

A Tutorial on the Estimation and Detection Theory with Applications to Radar, Array Processing, and Imaging

Jean-Philippe Ovarlez^{1,2}

¹SONDRA, CentraleSupélec, France

²ONERA, The French Aerospace Lab, DEMR/MATS, France

jeanphilippe.ovarlez@centralesupelec.fr

<https://www.jeanphilippeovarlez.com>

T03 Tutorial - Conference Radar 2024 - 21 October 2024



Contents

- **Part A:**
Background on Statistical Processing for Radar, Array Processing, SAR and Hyperspectral Imaging,
- **Part B:**
Recent Methodologies on Robust Estimation and Detection in non-Gaussian Environment
- Applications and Results in Radar, STAP and Array Processing, SAR Imaging, Hyperspectral Imaging

Part A

Background on Statistical Processing for Radar, Array Processing, SAR and Hyperspectral Imaging

Part A: Contents

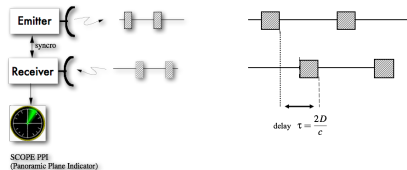
- 1 Radar basis
- 2 Conventional Radar and Imaging Processing
- 3 Some Background on Detection Theory
- 4 Motivations for more robust detection schemes

Outline

- 1 Radar basis
 - Parameter Estimation
 - Noise and Clutter in Radar
- 2 Conventional Radar and Imaging Processing
 - Range-Doppler Radar Processing
 - Array and Space-Time Adaptive Processing
 - SAR Image Processing
 - Hyperspectral Image Processing
- 3 Some Background on Detection Theory
 - Problem Statement
 - Modeling Homogeneous Gaussian Noise/Clutter
 - Examples of Detector Derivations
 - Synthesis of CFAR Detection Schemes Under Gaussian Noise
- 4 Motivations for more robust detection schemes
 - Examples of Gaussian Hypothesis Failure
 - Need of Better Approaches

Parameter Estimation - Range Measurement

Electromagnetic wave propagates with speed light c . The two-way propagation delay up to the distance D is $\tau = \frac{2D}{c}$



- Radar emitted signal: $s_e(t) = u(t) \exp(2i\pi f_0 t)$ where f_0 is the carrier frequency, and $u(\cdot)$ the baseband signal,
- Radar received signal: $s_r(t) = \alpha s_e(t - \tau) + b(t)$ where α is the backscattering amplitude of the target and $b(\cdot)$ is an additive noise.

$$s_r(t) = \alpha s_e\left(t - \frac{2D}{c}\right) + b(t).$$

Parameter Estimation - Velocity Measurement

Let us consider an illuminated moving target located for time t at range $D(t) = D_0 + v t$ where v is the radial target velocity.

If $\tau(t)$ is the two-way delay of the received signal at time t , the signal has been reflected at time $t - \tau(t)/2$ and the range $D(t)$ has to verify the following equation:

$$c \tau(t) = 2 D \left(t - \frac{\tau(t)}{2} \right) .$$

We obtain $\tau(t) = 2 \frac{D_0 + v t}{c + v}$ and the model relative to signal return is:

$$s_r(t) = \alpha s_e \left(\frac{c - v}{c + v} t - \frac{2 D_0}{c + v} \right) + b(t) .$$

The moving target is characterized in the signal return by a time-shift-compression/dilation of the emitted signal: action of Affine Group.

Parameter Estimation - Velocity Measurement

Under the so-called *narrow-band* assumptions:

- $f_0 \gg B$, where B is the bandwidth of baseband signal $u(\cdot)$,
- $v \ll c$,
- $2 B T \ll c/v$,

$$\begin{aligned} \text{We have: } s_r(t) &= \alpha s_e \left(\frac{c-v}{c+v} t - \frac{2 D_0}{c+v} \right) + b(t), \\ &= \alpha \exp(i \phi) u \left(t - \frac{2 D_0}{c} \right) \exp(2i \pi f_0 t) \exp \left(-2i \pi \frac{2 v}{c} f_0 t \right) + b(t). \end{aligned}$$

$$s_r(t) = \alpha' s_e \left(t - \frac{2 D_0}{c} \right) \exp(-2i \pi f_d t) + b(t).$$

where $|\alpha'| = |\alpha|$ and where $f_d = \frac{2 v}{c} f_0$ is called the **Doppler frequency** corresponding to moving target. The moving target is so characterized in the signal return by a time-shift/frequency shift of the emitted signal: action of Heisenberg Group .

Distance criterion - Ambiguity function and Matched Filter

One of the most important problem arising in radar theory is to separate targets in range and Doppler spaces. A $\mathcal{L}^2(\mathbb{R})$ distance R between two signals X and Y can be defined:

$$R^2 = \int_{-\infty}^{+\infty} |X(t) - Y(t)|^2 dt.$$

Minimizing this distance leads to maximize the inner product between X and Y (also known as Matched Filter):

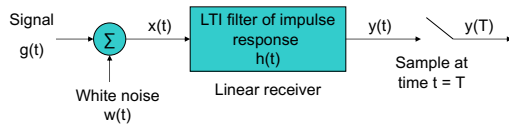
$$\int_{-\infty}^{+\infty} X(t) Y^*(t) dt.$$

According to the physical transformation of X , we obtain the so-called Ambiguity functions [Woodward 53, Kelly 65]:

- Example: $Y(t) = X(t - \tau) e^{2i\pi \nu t}$: $A(\tau, \nu) = \int_{-\infty}^{+\infty} X(t) X^*(t - \tau) e^{-2i\pi \nu t} dt,$
- Example: $Y(t) = \frac{1}{\sqrt{a}} X(a^{-1}t - b)$: $A(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} X(t) X^*(a^{-1}t - b) dt.$

Link with the so-called Matched Filter and Pulse Compression

Let us consider a linear time-invariant filter of impulse response $h(t)$. The filter input $x(t)$ consists of a pulse signal $g(t)$ corrupted by additive zero mean white noise $w(t)$ (with Power Spectral Density $\Phi_w(f) = N_0/2$). The output is $y(t) = g_0(t) + n(t)$, the signal and noise components of the input $x(t)$ for $0 \leq t \leq T$.



$$SNR = \frac{|g_0(T)|^2}{\sigma_n^2} = \frac{|g_0(T)|^2}{E[n^2(t)]},$$

where $|g_0(T)|^2$ is the power of the filtered signal $g(t)$ at $t = T$, and $\sigma_n^2 = E[n^2(t)]$ is the power of the filtered noise.

Since $|g_0(t)|^2 = \left| \int G(f) H(f) e^{j2\pi ft} df \right|^2$ and $\sigma_n^2 = R_n(0)$ where $R_n(\tau) = \int \frac{N_0}{2} |H(f)|^2 e^{2i\pi f\tau} df$, the final expression for the output SNR is:

$$SNR = \frac{\left| \int H(f) G(f) e^{j2\pi fT} df \right|^2}{\frac{N_0}{2} \int |H(f)|^2 df} \leq \frac{2}{N_0} \int |G(f)|^2 df.$$

The SNR output is maximized only for the particular impulse response $h(t)$ that verifies:

$$H(f) = k G^*(f) e^{-2i\pi fT}, \forall k \in \mathbb{C}, \text{ or } h(t) = k g^*(T - t).$$

Range resolution

Let us suppose N targets with amplitude $\{\alpha_i\}_{i \in [1, N]}$ located in range space at distance $\left\{d_i = \frac{c \tau_i}{2}\right\}_{i \in [1, N]}$. The received signal $s_r(t)$ is:

$$s_r(t) = \sum_{i=1}^N \alpha_i s_e(t - \tau_i) \xrightarrow{t \rightarrow f} S_r(f) = \sum_{i=1}^N \alpha_i S_e(f) e^{-2i\pi f \tau_i}.$$

The radar processing leads to evaluate for all τ , the following expression:

$$R(\tau) = \int_{-\infty}^{+\infty} s_r(t) s_e^*(t - \tau) dt \xrightarrow{t \rightarrow f} R(\tau) = \sum_{i=1}^N \alpha_i \int_{-\infty}^{+\infty} |S_e(f)|^2 e^{2i\pi f (\tau - \tau_i)} df.$$

- When $S_e(f) = 1$ for $f \in]-\infty, +\infty[$, $R(\tau) = \sum_{i=1}^N \alpha_i \delta(\tau - \tau_i)$,
- When $S_e(f) = 1$ for $f \in [-B/2, +B/2]$, $R(\tau) = \sum_{i=1}^N \alpha_i \frac{\sin(\pi B (\tau - \tau_i))}{\pi B (\tau - \tau_i)}.$

Range resolution

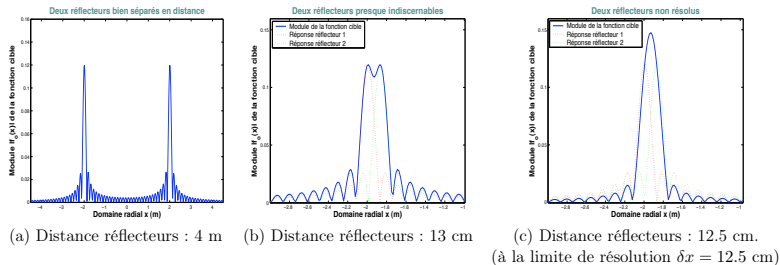


Figure: Here: $\delta D = 0.125$ m and $B = 1.2 \cdot 10^9$ Hz

The range resolution δD (so-called *Range Bin*) is proportional to the inverse of the emitted signal bandwidth B :

$$\delta D = \frac{c}{2} \frac{1}{B}.$$

Velocity resolution

Let us suppose N targets with amplitude $\{\alpha_i\}_{i \in [1, N]}$ with Doppler $\left\{ \nu_i = \frac{2 \nu_i}{c} f_0 \right\}_{i \in [1, N]}$. The received signal $S_r(f)$ is:

$$S_r(f) = \sum_{i=1}^N \alpha_i S_e(f - \nu_i) \xrightarrow{f \rightarrow t} s_r(t) = \sum_{i=1}^N \alpha_i s_e(t) e^{2i\pi \nu_i t}.$$

The radar processing leads to evaluate for all ν , the following expression:

$$R(\nu) = \int_{-\infty}^{+\infty} S_r(f) S_e^*(f - \nu) df \xrightarrow{t \rightarrow f} R(\nu) = \sum_{i=1}^N \alpha_i \int_{-\infty}^{+\infty} |s_e(t)|^2 e^{-2i\pi t(\nu - \nu_i)} dt.$$

The velocity resolution δV (so-called *Doppler Bin*) is proportional to the inverse of the emitted signal duration (or integration time) T :

$$\delta V = \frac{c}{2 f_0} \frac{1}{T}.$$

Joint range and Velocity resolution

Let us suppose N targets with amplitude $\{\alpha_i\}_{i \in [1, N]}$ moving at velocity $\{v_i\}_{i \in [1, N]}$ and located in range space at distance $\left\{d_i = \frac{c \tau_i}{2}\right\}_{i \in [1, N]}$. The received signal $S_r(f)$ is:

$$s_r(t) = \sum_{i=1}^N \alpha_i s_e(t - \tau_i) e^{2i \pi v_i t}.$$

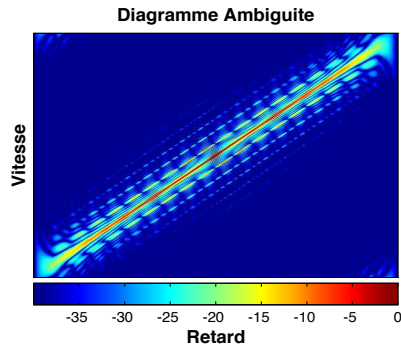
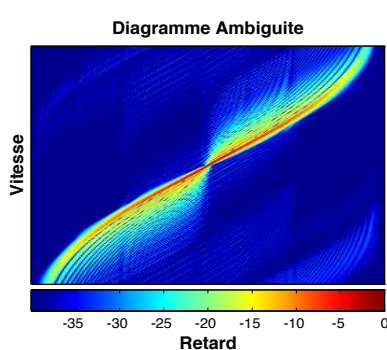
The radar processing (Matched Filter) leads to evaluate for all (τ, v) , the following expression:

$$R(\tau, v) = \int_{-\infty}^{+\infty} s_r(t) s_e^*(t - \tau) e^{-2i \pi v t} dt.$$

This last equation is the superposition of the ambiguity functions [Rihaczek 1969] centered at $\{(\tau_i, v_i)\}_{i \in [1, N]}$

$$R(\tau, v) = \sum_{i=1}^N \alpha_i A(\tau - \tau_i, v - v_i).$$

Some examples of Ambiguity Functions



- Best radar waveforms are those which look like a *thumbtack* form ($A(\tau, \nu) = \delta(\tau) \delta(\nu)$) but they definitely don't exist :-)
- Range and Doppler sidelobes can be troublesome for high density targets detection because of their superposition at different ranges and Doppler [Rihaczek 1969].

Link with Minimal Bounds (Cramer Rao bounds)

- Let us define the second order moments (centered) of the signal

$$\sigma_t^2 = \int_{-\infty}^{+\infty} t^2 |s_e(t)|^2 dt \approx T^2, \quad \sigma_f^2 = \int_{-\infty}^{+\infty} f^2 |S_e(f)|^2 df \approx B^2 \text{ and the modulation index}$$

$m = \frac{-1}{2\pi} \text{Im} \int_{-\infty}^{+\infty} t s_e(t) \frac{ds_e^*(t)}{dt} dt$. Under white Gaussian noise with variance σ^2 , range and doppler accuracies are given by the following **Cramer-Rao Bounds** [Kay 93, Kay 98]:

$$E[(\nu - \hat{\nu})^2] = \frac{\sigma^2}{4\pi^2 \alpha^2} \frac{\sigma_f^2}{\sigma_f^2 \sigma_t^2 - (m - t_0 f_0)^2} \geq \frac{\sigma^2}{4\pi^2 \alpha^2} \frac{1}{\sigma_t^2}, \quad (1)$$

$$E[(\tau - \hat{\tau})^2] = \frac{\sigma^2}{4\pi^2 \alpha^2} \frac{\sigma_t^2}{\sigma_f^2 \sigma_t^2 - (m - t_0 f_0)^2} \geq \frac{\sigma^2}{4\pi^2 \alpha^2} \frac{1}{\sigma_f^2}, \quad (2)$$

$$E[(\nu - \hat{\nu})(\tau - \hat{\tau})] = \frac{\sigma^2}{4\pi^2 \alpha^2} \cdot \frac{m - t_0 f_0}{\sigma_f^2 \sigma_t^2 - (m - t_0 f_0)^2} \quad (3)$$

- Radar uses to emit signal characterized with high time-bandwidth product $B T$.

Outline

- 1 Radar basis
 - Parameter Estimation
 - **Noise and Clutter in Radar**
- 2 Conventional Radar and Imaging Processing
 - Range-Doppler Radar Processing
 - Array and Space-Time Adaptive Processing
 - SAR Image Processing
 - Hyperspectral Image Processing
- 3 Some Background on Detection Theory
 - Problem Statement
 - Modeling Homogeneous Gaussian Noise/Clutter
 - Examples of Detector Derivations
 - Synthesis of CFAR Detection Schemes Under Gaussian Noise
- 4 Motivations for more robust detection schemes
 - Examples of Gaussian Hypothesis Failure
 - Need of Better Approaches

Noise and Clutter in Radar

Thermal noise

Thermal noise for most radars corresponds to additive complex white Gaussian noise $\mathcal{CN}(\mathbf{0}_m, \mathbf{I}_m)$. This noise is generated by electronic devices in radar receivers.

What is the clutter?

Clutter refers to radio frequency (RF) echoes returned from targets which are uninteresting to the radar operators and interfere with the observation of useful signals.

Such targets include natural objects such as ground, sea, precipitations (rain, snow or hail), sand storms, animals (especially birds), atmospheric turbulence, and other atmospheric effects, such as ionosphere reflections and meteor trails.

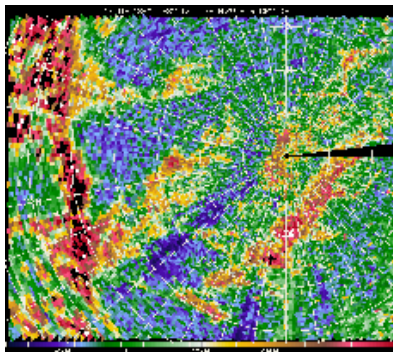
Clutter may also be returned from man-made objects such as buildings and, intentionally, by radar countermeasures such as chaff.

A statistical model for the clutter is necessary: can we consider the clutter as Gaussian process, non-Gaussian process, iid, correlated, stationary ????

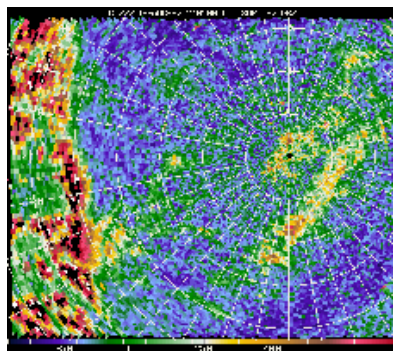
Noise and Clutter in Radar

Example of clutter map for different azimuth resolutions

resolution 3°



resolution 1°



Outline

- 1 Radar basis
 - Parameter Estimation
 - Noise and Clutter in Radar
- 2 Conventional Radar and Imaging Processing
 - **Range-Doppler Radar Processing**
 - Array and Space-Time Adaptive Processing
 - SAR Image Processing
 - Hyperspectral Image Processing
- 3 Some Background on Detection Theory
 - Problem Statement
 - Modeling Homogeneous Gaussian Noise/Clutter
 - Examples of Detector Derivations
 - Synthesis of CFAR Detection Schemes Under Gaussian Noise
- 4 Motivations for more robust detection schemes
 - Examples of Gaussian Hypothesis Failure
 - Need of Better Approaches

Range-Doppler Radar Processing

- The cross-correlation operation is closely related to the so-called *Matched Filter* (filter which maximizes the SNR at its output). This is also known as the *pulse compression* processing. This matched filter offers the gain $B T$ on the noise power σ^2 ,
- The Doppler resolution is inversely proportional to the integration time. For monostatic radar (both emission and reception on the same antenna), radar prefers to cut off this long integration time into m pulses of duration T with Pulse Repetition Frequency (PRF) $F_r = 1/T_r$ (total integration time $m T_r$):

$$s(t) = \sum_{k=0}^{m-1} s_e(t - k T_r).$$

Considering the signal return $s_r(t)$, the radar processing consists in evaluating:

$$R(\tau, \nu) = \int_{-\infty}^{+\infty} s_r(t) s_r^*(t - \tau) e^{-2i\pi\nu t} dt = \sum_{n=0}^{m-1} e^{-2i\pi\nu n T_r} \int_0^{T_r} s_r(u + n T_r) s_r^*(u - \tau) \cancel{e^{-2i\pi\nu u}} du.$$

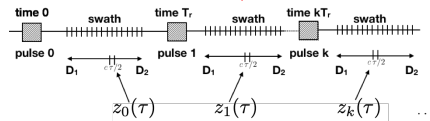
Neglecting the Doppler into the pulse duration leads to adapting the processing to the 0-Doppler: missing high-speed targets, bias in range estimation due to the ambiguity.

Range-Doppler Radar Processing

When supposing non migrating target and neglecting the Doppler variation in the pulse, we can rewrite the processing as:

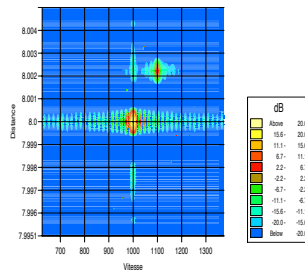
$$R(\tau, \nu) = \sum_{n=0}^{m-1} e^{-2i\pi\nu n T_r} \underbrace{\int_0^{T_r} s_r(u + n T_r + \tau) s_e^*(u) du}_{z_n(\tau)} = \mathbf{p}^H \mathbf{z},$$

where $\mathbf{z} = (z_0(\tau), z_1(\tau), \dots, z_{m-1}(\tau))^T$ and $\mathbf{p} = (1, e^{2i\pi\nu T_r}, \dots, e^{2i\pi\nu (m-1) T_r})^T$.



- For each range bin $c\tau/2$ (time T_r can be sampled at resolution $\delta\tau = 1/B$) on the range support $[D_1, D_2]$ of the analyzed swath, compute $z_n(\tau)$ corresponding to the time correlation between received signal and emitted pulse $s_e(t)$ at time $n T_r$,
- For each range bin $c\tau/2$, compute the Discrete Fourier Transform ($\mathbf{p}^H \mathbf{z}$) over the m coefficients $\{z_n(\tau)\}_{n \in [0, m-1]}$ to characterize Doppler spectrum in the spectral support $\nu \in [0, 1/T_r]$.

Range-Doppler Radar Processing



Example of the so-called Range-Doppler map of the processing data.

- Coherent Doppler processing brings an improvement of m on the Doppler resolution with regards to the one pulse processing ($\delta v = 1/(m T_r)$) as well as a gain m in SNR.
- Range resolution does not change. Always related by the pulse bandwidth,
- Appearance of the range ambiguities at ranges $c k T_r/2$,
- Appearance of the Doppler ambiguities at Doppler frequency k/T_r .

Outline

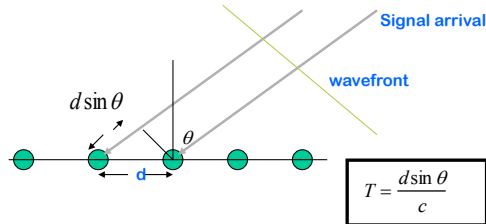
- 1 Radar basis
 - Parameter Estimation
 - Noise and Clutter in Radar
- 2 Conventional Radar and Imaging Processing
 - Range-Doppler Radar Processing
 - **Array and Space-Time Adaptive Processing**
 - SAR Image Processing
 - Hyperspectral Image Processing
- 3 Some Background on Detection Theory
 - Problem Statement
 - Modeling Homogeneous Gaussian Noise/Clutter
 - Examples of Detector Derivations
 - Synthesis of CFAR Detection Schemes Under Gaussian Noise
- 4 Motivations for more robust detection schemes
 - Examples of Gaussian Hypothesis Failure
 - Need of Better Approaches

Array and Space-Time Adaptive Processing

Source locating in azimuth θ , at Doppler ν and in range bin $c\tau/2$

If the radar receives a signal on an antenna array, each antenna is collecting $s_r(t)$ delayed by the time shift $T = nd \sin \theta / c$ depending on its spatial position nd ($n \in [0, N_s]$) on the array. Supposing that the array is non-dispersive ($N_s d \sin \theta \ll c/B$), the concatenated $N_s \times m$ -observation vector \mathbf{y} collected by the radar on the antenna array for a given range bin $c\tau/2$ and Doppler ν is then:

$$\mathbf{y} = A \mathbf{p} \otimes \left(1, e^{2i\pi f_0 d \sin \theta / c}, \dots, e^{2i\pi f_0 (N_s-1) d \sin \theta / c} \right)^T + \mathbf{b}.$$



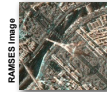
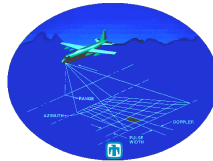
Outline

- 1 Radar basis
 - Parameter Estimation
 - Noise and Clutter in Radar
- 2 Conventional Radar and Imaging Processing
 - Range-Doppler Radar Processing
 - Array and Space-Time Adaptive Processing
 - **SAR Image Processing**
 - Hyperspectral Image Processing
- 3 Some Background on Detection Theory
 - Problem Statement
 - Modeling Homogeneous Gaussian Noise/Clutter
 - Examples of Detector Derivations
 - Synthesis of CFAR Detection Schemes Under Gaussian Noise
- 4 Motivations for more robust detection schemes
 - Examples of Gaussian Hypothesis Failure
 - Need of Better Approaches

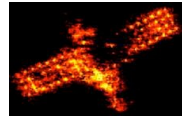
Background on SAR and Radar Imaging



ONERA RAMSES Image



RAMSES Image



ONERA ISAR Image



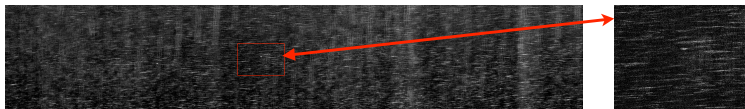
ONERA RAMSES Image

Radar Imaging [*Mensa 81, Soumekh 94, 99*] allows to build more and more complex images:

- Current use of **very high spectral bandwidth** and **very high angular bandwidth** leading to very high spatial resolution,
- Application to monitoring (detection, change detection), classification, 3D reconstruction, EM analysis, etc.

These applications require some physical diversity to reach good performances.

Detection in monochannel SAR Images



Conventional SAR detection framework on a **mono-channel SAR image** mainly consists in locally comparing the complex amplitude of pixel x_i . In Gaussian homogeneous environment:

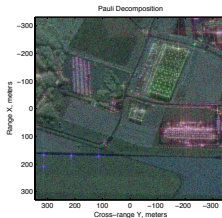
- Global thresholding $\Lambda(x_i) = |x_i|^2 \underset{H_0}{\overset{H_1}{\geq}} \lambda$, leads to $\lambda = -\sigma^2 \log P_{fa}$,
- Local thresholding $\Lambda(x_i) = \frac{|x_i|^2}{\frac{1}{N} \sum_{k \neq i}^N |x_k|^2} \underset{H_0}{\overset{H_1}{\geq}} \lambda$ leads to $\lambda = N \left(P_{fa}^{-1/N} - 1 \right)$.
- Statistic-based thresholding (other distributions): $\Lambda(x_i) = g(x_i) \underset{H_0}{\overset{H_1}{\geq}} \lambda$ leads to $\lambda = f(P_{fa})$.

Detection scheme on mono-channel SAR image only consists in thresholding the intensity pixel of the image.

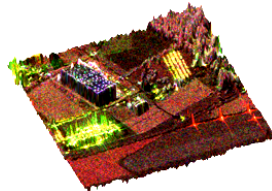
Multi-Channel SAR Images

Multi-channel SAR images automatically propose this diversity through:

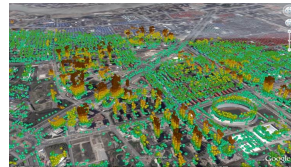
- polarimetric channels (POLSAR), interferometric channels (INSAR), polarimetric and interferometric channels (POLINSAR),
- multi-temporal, multi-bands, multi-passes SAR Image, etc.



EM behavior of the terrain
in POLSAR images



Estimation of the height
in POLINSAR images



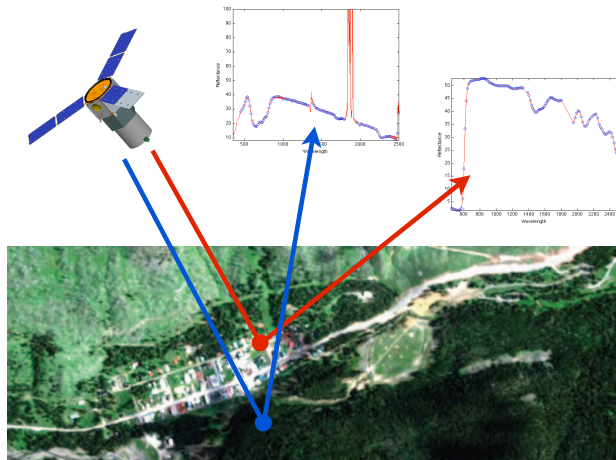
Analysis of the structures displacement in
Shanghai with multi-temporal SAR images
(@Telespazio)

Almost all the conventional techniques of detection, parameters estimation, speckle filtering techniques, classification in multi-channel SAR images (e.g. polarimetric covariance matrix, interferometric coherency matrix) are based on the **multivariate statistic**.

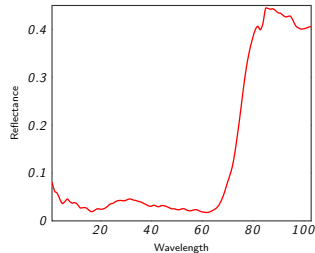
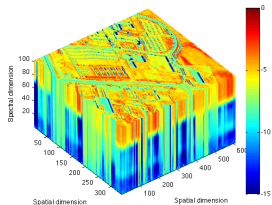
Outline

- 1 Radar basis
 - Parameter Estimation
 - Noise and Clutter in Radar
- 2 Conventional Radar and Imaging Processing
 - Range-Doppler Radar Processing
 - Array and Space-Time Adaptive Processing
 - SAR Image Processing
 - **Hyperspectral Image Processing**
- 3 Some Background on Detection Theory
 - Problem Statement
 - Modeling Homogeneous Gaussian Noise/Clutter
 - Examples of Detector Derivations
 - Synthesis of CFAR Detection Schemes Under Gaussian Noise
- 4 Motivations for more robust detection schemes
 - Examples of Gaussian Hypothesis Failure
 - Need of Better Approaches

Hyperspectral Imaging (HSI)



Hyperspectral Imaging (HSI)



- **Anomaly Detection**

To detect all that is "different" from the background (Mahalanobis distance) - No information (steering vector \mathbf{p}) about the targets of interest available. [Frontera 16].

- **"Pure" Detection**

To detect targets characterized by a given spectral signature \mathbf{p} - Regulation of False Alarm [Frontera-Pons 17].

Detection and Steering vector modeling

Many detection problems can be viewed as a "simple" problem of detection of the information vector \mathbf{p} characterizing the target:

Radar

- \mathbf{p} can model range information, Doppler frequency, direction of arrival, joint content of all these informations (STAP), multi-band, etc.

SAR

- \mathbf{p} can model polarimetric, interferometric, joint polarimetric and interferometric informations, multi-band, sub-look and sub-band behavior, etc.

Hyperspectral

- \mathbf{p} can model any material characterized by its given spectral signature (spectroscopy).
- \mathbf{p} can be also unknown (Anomaly detection)

We can then develop adaptive multivariate detection schemes using modeling the so-called target information *steering vector* \mathbf{p}

Outline

- 1 Radar basis
 - Parameter Estimation
 - Noise and Clutter in Radar
- 2 Conventional Radar and Imaging Processing
 - Range-Doppler Radar Processing
 - Array and Space-Time Adaptive Processing
 - SAR Image Processing
 - Hyperspectral Image Processing
- 3 Some Background on Detection Theory
 - **Problem Statement**
 - Modeling Homogeneous Gaussian Noise/Clutter
 - Examples of Detector Derivations
 - Synthesis of CFAR Detection Schemes Under Gaussian Noise
- 4 Motivations for more robust detection schemes
 - Examples of Gaussian Hypothesis Failure
 - Need of Better Approaches

General Formulation of All the Detection Problems

Set of two binary hypotheses

$$\begin{cases} H_0 : \mathbf{z} = \mathbf{b} \\ H_1 : \mathbf{z} = A\mathbf{p} + \mathbf{b} \end{cases}, \text{ where}$$

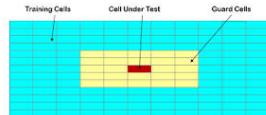
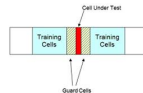
- \mathbf{z} is a m -vector of data collected in a given measurement support. It can be range support, spatial support (Imaging), etc.
- The complex amplitude A of the target to detect is considered here deterministic (no fluctuation)
- The m -vector \mathbf{b} represents the additive noise (thermal noise, photon noise, clutter, jam, etc.) characterized by a known (or unknown) PDF.
- The m -vector \mathbf{p} represents the so-called deterministic *steering vector*: it can be relative to Doppler, Polarimetry, Interferometry, Wavelength, Spatial, time, joint Angular and Spectral information (STAP).

The problem here consists in choosing between H_1 hypothesis and H_0 hypothesis.

Problem Statement

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

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



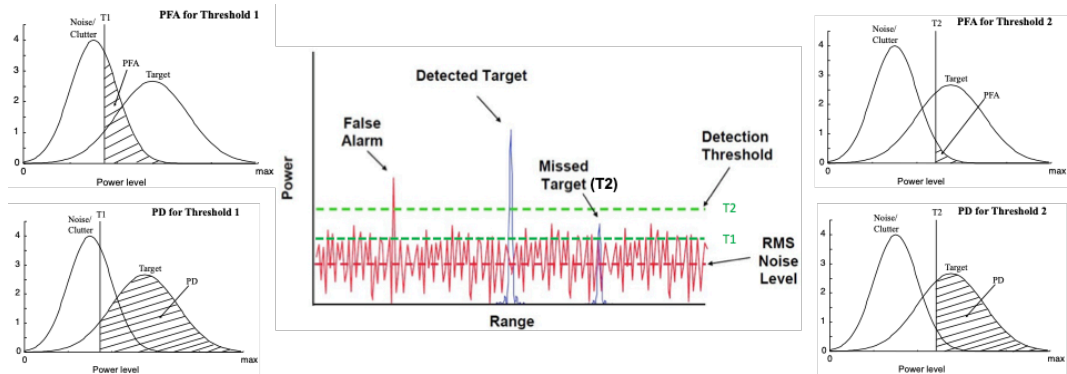
where the \mathbf{z}_i 's are n "signal-free" independent secondary data used to estimate the noise parameters. \Rightarrow **Neyman-Pearson criterion** [Kay 93, Kay 98]

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

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

- 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),

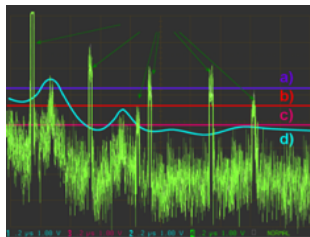
Pd/Pfa



$$P_{fa} = \mathbb{P}(\Lambda(\mathbf{z}) > \lambda/H_0),$$

$$P_d = \mathbb{P}(\Lambda(\mathbf{z}) > \lambda/H_1).$$

False Alarm Regulation Importance



- a. threshold is set too high: Probability of Detection = 20%
- b. threshold is set optimal: Probability of Detection = 80%
But one false alarm arises!
False alarm rate = $1 / 666 = 1,5 \cdot 10^{-3}$
- c. threshold is set too low: a large number of false alarms arises!
- d. threshold is set variable: constant false-alarm rate

CFAR Property

A detector is said Constant False Alarm Rate (CFAR property) if the PDF of the test is independent on the noise parameter (mean, covariance, variance, statistic) under H_0 hypothesis.

Multivariate Gaussian distribution

Definition

Let $\mathbf{x} = (x_1, \dots, x_m)^T$ be a random vector. The vector \mathbf{x} is Gaussian if and only if, for any sequence $\mathbf{a} = (a_1, \dots, a_m)^T \in \mathcal{R}^m$ of real numbers, the scalar random variable

$$z = \mathbf{a}^T \mathbf{x} = \sum_{i=1}^m a_i x_i \text{ is a Gaussian variable.}$$

We note $\boldsymbol{\mu} = \mathbb{E}(\mathbf{x})$ its mean and $\boldsymbol{\Sigma} = \mathbb{E} \left[(\mathbf{x} - \boldsymbol{\mu}) (\mathbf{x} - \boldsymbol{\mu})^T \right]$ its covariance matrix.

Its PDF that is noted $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is given by

$$p_{\mathbf{x}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^m |\boldsymbol{\Sigma}|}} \exp \left(-\frac{(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})}{2} \right).$$

The covariance matrix is modeling the correlation existing withing the components of the observation vector \mathbf{x}

Complex Gaussian distribution

Definition

A random vector $\mathbf{z} = \mathbf{x} + j\mathbf{y}$ is complex Gaussian distributed $\mathbf{z} \sim \mathbb{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}_z, \mathbf{P}_z)$ iif

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \Re(\boldsymbol{\mu}) \\ \Im(\boldsymbol{\mu}) \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_x & \boldsymbol{\Sigma}_{xy} \\ \boldsymbol{\Sigma}_{yx} & \boldsymbol{\Sigma}_y \end{bmatrix} \right)$$

with $\boldsymbol{\Sigma}_z = \mathbb{E}[(\mathbf{z} - \boldsymbol{\mu})(\mathbf{z} - \boldsymbol{\mu})^H] = \boldsymbol{\Sigma}_x + \boldsymbol{\Sigma}_y + j(\boldsymbol{\Sigma}_{yx} - \boldsymbol{\Sigma}_{xy})$ and
 $\mathbf{P}_z = \mathbb{E}[(\mathbf{z} - \boldsymbol{\mu})(\mathbf{z} - \boldsymbol{\mu})^T] = \boldsymbol{\Sigma}_x - \boldsymbol{\Sigma}_y + j(\boldsymbol{\Sigma}_{yx} + \boldsymbol{\Sigma}_{xy})$

Circularity Property

$\mathbf{z} = \mathbf{x} + j\mathbf{y} \in \mathbb{C}^m$ is circularly symmetric $\mathbf{z} \sim \mathbb{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}_z)$ iif $\mathbf{z} \stackrel{d}{=} e^{j\varphi}(\mathbf{z} - \boldsymbol{\mu}) \forall \varphi \in [0; 2\pi[$.
 Notably, $\boldsymbol{\Sigma}_x = \boldsymbol{\Sigma}_y$ and $\boldsymbol{\Sigma}_{yx} = \boldsymbol{\Sigma}_{xy} = \mathbf{0} \Leftrightarrow \boldsymbol{\Sigma}_z = 2\boldsymbol{\Sigma}_x$ and $\mathbf{P}_z = \mathbf{0}$.

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

Noise distribution

Central limit theorem

Let x_1, x_2, \dots, x_n be a sequence of random variable i.i.d. with finite mean μ and variance σ , then

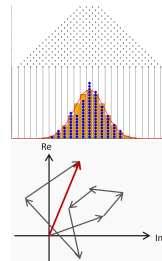
$$\sqrt{n} \bar{x}_n \xrightarrow[n \rightarrow \infty]{a.s.} \mathcal{N}(0, \sigma^2) \quad \text{with} \quad \bar{x}_n = \frac{x_1 + x_2 + \dots + x_n}{n}.$$

A Gaussian/Normal random variable has the largest entropy among all random variables of equal variance.

Speckle noise (Goodman 1976)

$$z = \sum_{i=1}^n a_i \exp j\varphi_i \Rightarrow z \sim \mathbb{CN}(0, \sigma^2), \quad p(z) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{|z|^2}{2\sigma^2}\right).$$

This explains why the **Gaussian distribution** is often used to model the in-phase return of a large number of i.i.d. backscatterers in a radar resolution cell.



The Galton board (top), Random walk (bottom)

Link Between Covariance Matrix and Power Spectral Density 1/2

The **Power Spectral Density** $\Phi(f)$ characterizes, in a given range bin, the spectral (Doppler) fluctuations of a process $\mathbf{z} = (z_0, \dots, z_{m-1})^T$ collected from pulse to pulse.

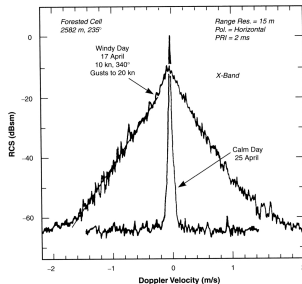


FIGURE 4.23 Power spectra of X-band radar returns from windblown trees.

[Billingsley 1993]

- Examples of some PSD models with $-1/(2 T_r) \leq f \leq 1/(2 T_r)$:

$$\Phi(f) = \Phi_0 \exp\left(-\frac{(f - f_c)^2}{2 \sigma_f^2}\right), \quad \Phi(f) = \frac{\Phi_0}{1 + \left(\frac{f}{f_c}\right)^n},$$

- Autocorrelation function (Wiener-Khintchine Theorem):

$$\rho(\tau) = \int_{-\infty}^{+\infty} \Phi(f) \exp(2i \pi f \tau) df.$$

- Covariance Matrix: $\Sigma = E[\mathbf{z} \mathbf{z}^H] = \begin{pmatrix} \rho(0) & \dots & \rho((m-1) T_r) \\ \vdots & \ddots & \vdots \\ \rho((m-1) T_r) & \dots & \rho(0) \end{pmatrix}.$

Link Between Covariance Matrix and Power Spectral Density 2/2

Examples of PSD modeling and their associated covariance matrices:

- $\Phi(f) = N_0 \in \mathbb{R}^+$ that corresponds to a white noise leads to the CM equal to $\Sigma = N_0 B \mathbf{I}$, where B is the bandwidth of the receiver and \mathbf{I} is the identity matrix.
- The exponential PSD $\Phi(f) = P_0 \exp(-\alpha|f|)$, with $\alpha \in \mathbb{R}^+$ corresponds to the CM equal to $\left\{ \rho(k T_r) = 2 \alpha P_0 (\alpha^2 + 4\pi^2(k T_r)^2)^{-1/2} \right\}_{k \in [0, m-1]}$
- For any $0 \leq |\rho_0| \leq 1$, the practical covariance model $\Sigma_{i,j} = \left\{ \rho_0^{|i-j|} \right\}_{i,j \in [1, m-1]}$ leads to the

$$\text{PSD } |\Phi(f)| = \left| \frac{1 - \rho_0 \exp(2i\pi f m T_r)}{1 - \rho_0 \exp(2i\pi f T_r)} \right|.$$

General Detection Theory

When some parameters (noise, target) are unknown:

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

$$\Lambda(\mathbf{z}) = \frac{\max_{\boldsymbol{\theta}} \max_{\boldsymbol{\mu}} p_{\mathbf{z}/H_1}(\mathbf{z}, \boldsymbol{\theta}, \boldsymbol{\mu})}{\max_{\boldsymbol{\mu}} p_{\mathbf{z}/H_0}(\mathbf{z}, \boldsymbol{\mu})} \underset{H_0}{\underset{H_1}{\gtrless}} \lambda,$$

where $\boldsymbol{\theta}$ and $\boldsymbol{\mu}$ represent respectively the unknown target parameter vector and the unknown noise parameter vector.

CFAR Property

A GLRT detector is said Constant False Alarm Rate (CFAR property) if the PDF of the GLRT test is independent on the noise parameter (mean, covariance, variance, statistic) under H_0 hypothesis.

General Estimation Theory: unknown deterministic parameters

- **Maximum Likelihood Estimation (MLE) scheme:** maximize the PDF with respect to the unknown parameter. Ex for noise parameter μ :

$$\hat{\mu} = \operatorname{argmax}_{\mu} p_{\mathbf{z}/H_0}(\mathbf{z}, \mu).$$

Example: Suppose n target-free i.i.d. m -vectors $\{\mathbf{z}_i\}_{i=1,n}$ where $\mathbf{z}_i \sim \mathcal{CN}(\mathbf{0}_m, \Sigma)$ where Σ is an unknown covariance matrix. The MLE $\hat{\Sigma}_n$ is set by solving

$$\frac{\delta}{\delta \Sigma} \log \prod_{i=1}^n p_{\mathbf{z}}(\mathbf{z}_i, \Sigma) = \frac{\delta}{\delta \Sigma^{-1}} \left(n \log |\Sigma^{-1}| - \sum_{i=1}^N \mathbf{z}_i^H \Sigma^{-1} \mathbf{z}_i \right) = \mathbf{0}.$$

Recalling that $\frac{\delta}{\delta \Sigma^{-1}} \log |\Sigma^{-1}| = \Sigma^T$ and $\frac{\delta}{\delta \Sigma^{-1}} (\mathbf{z}_i^H \Sigma^{-1} \mathbf{z}_i) = (\mathbf{z}_i \mathbf{z}_i^H)^T$, we obtain:

Sample Covariance Matrix: MLE of the Gaussian problem

$$\hat{\Sigma}_n = \frac{1}{n} \sum_{i=1}^n \mathbf{z}_i \mathbf{z}_i^H.$$

Outline

- 1 Radar basis
 - Parameter Estimation
 - Noise and Clutter in Radar
- 2 Conventional Radar and Imaging Processing
 - Range-Doppler Radar Processing
 - Array and Space-Time Adaptive Processing
 - SAR Image Processing
 - Hyperspectral Image Processing
- 3 Some Background on Detection Theory
 - Problem Statement
 - **Modeling Homogeneous Gaussian Noise/Clutter**
 - Examples of Detector Derivations
 - Synthesis of CFAR Detection Schemes Under Gaussian Noise
- 4 Motivations for more robust detection schemes
 - Examples of Gaussian Hypothesis Failure
 - Need of Better Approaches

Modeling Homogeneous Gaussian Noise/Clutter

Problem to solve in Gaussian environment

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

where $\mathbf{s} = A\mathbf{p}$, \mathbf{y} and $\mathbf{y}_i \sim \mathcal{CN}(\mathbf{0}_m, \mathbf{\Sigma})$, i.e. $p_{\mathbf{z}}(\mathbf{z}) = \frac{1}{\pi^m |\mathbf{\Sigma}|} \exp(-\mathbf{z}^H \mathbf{\Sigma}^{-1} \mathbf{z})$

Goal: to choose the best hypothesis while minimizing the risk of being wrong (False Alarm) from an observation vector \mathbf{z}

\Rightarrow **All is known for Gaussian assumption!**

Sample Covariance Matrix (SCM)

When $\mathbf{\Sigma}$ is unknown, the Gaussian environment is modeled through the SCM: $\hat{\mathbf{S}}_n = \frac{1}{n} \sum_{i=1}^n \mathbf{z}_i \mathbf{z}_i^H$.

Properties of the SCM in homogeneous Gaussian noise/clutter environment

Properties of the SCM

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

Then, $\sqrt{n} \text{vec} \left(\hat{\mathbf{S}}_n - \mathbf{\Sigma} \right) \xrightarrow{d} \mathcal{CN} \left(\mathbf{0}_{m^2}, \mathbf{C}, \mathbf{P} \right),$

$$\begin{aligned} \text{where } \mathbf{C} &= (\mathbf{\Sigma}^* \otimes \mathbf{\Sigma}) \\ \mathbf{P} &= (\mathbf{\Sigma}^* \otimes \mathbf{\Sigma}) \mathbf{K}_{m^2, m^2}. \end{aligned}$$

where $\mathbf{K}_{m,m}$ is the $m \times m$ commutation matrix transforming any m -vector $\text{vec}(\mathbf{A})$ into $\text{vec}(\mathbf{A}^T)$.

Under Gaussian assumptions $\mathcal{CN}(\mathbf{0}_m, \mathbf{\Sigma})$, 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 :

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

- This estimate is unbiased, efficient, Wishart distributed,
- n can represent any samples support called *the secondary data*: 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 **Space-Time Adaptive Processing**, it can represent the joint spatial and time information collected in a given range bin of interest, n is here the range bin 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.

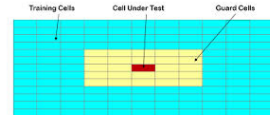
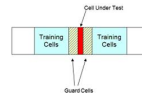
Outline

- 1 Radar basis
 - Parameter Estimation
 - Noise and Clutter in Radar
- 2 Conventional Radar and Imaging Processing
 - Range-Doppler Radar Processing
 - Array and Space-Time Adaptive Processing
 - SAR Image Processing
 - Hyperspectral Image Processing
- 3 Some Background on Detection Theory
 - Problem Statement
 - Modeling Homogeneous Gaussian Noise/Clutter
 - **Examples of Detector Derivations**
 - Synthesis of CFAR Detection Schemes Under Gaussian Noise
- 4 Motivations for more robust detection schemes
 - Examples of Gaussian Hypothesis Failure
 - Need of Better Approaches

Example 0 - Detection Schemes in Range Doppler map

- In a **scalar** measurement z , detecting an unknown complex deterministic signal s embedded in an additive noise y can be written as the following statistical test:

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



where the z_i 's are n "signal-free" independent secondary data used to estimate the noise parameters. \Rightarrow **Neyman-Pearson criterion** [Kay 93, Kay 98]

Conventional detection framework on **a mono-channel radar data** mainly consists of locally comparing the complex amplitude of pixel z . In Gaussian homogeneous environment, i.e. $y \sim \mathcal{CN}(0, \sigma^2)$:

- Known power σ^2** : global thresholding $\rightarrow \Lambda(z) = |z|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \lambda$, leads to $\lambda = -\sigma^2 \log P_{fa}$,
- Unknown power σ^2** : local thresholding $\rightarrow \Lambda(z) = \frac{|z|^2}{\frac{1}{N} \sum_{k \neq i}^N |z_k|^2} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda$ leads to $\lambda = N \left(P_{fa}^{-1/N} - 1 \right)$.

The detection scheme only consists of thresholding the intensity of each map pixel.

Example 1 - Detection Schemes in Gaussian Noise

$$\text{Problem under study: } \begin{cases} \text{Hypothesis } H_0: & \mathbf{z} = \mathbf{b} \\ \text{Hypothesis } H_1: & \mathbf{z} = A\mathbf{p} + \mathbf{b}, \end{cases}$$

where $A \neq 0$ is a **known** complex scalar amplitude, \mathbf{p} is the **known** steering vector and $\mathbf{b} \sim \mathcal{CN}(\mathbf{0}_m, \mathbf{\Sigma})$ with **known** covariance matrix $\mathbf{\Sigma}$. The probability density functions of the received m -vector \mathbf{z} under each hypothesis are given by:

$$p_{\mathbf{z}/H_0}(\mathbf{z}) = \frac{1}{\pi^m |\mathbf{\Sigma}|} \exp(-\mathbf{z}^H \mathbf{\Sigma}^{-1} \mathbf{z}) \quad p_{\mathbf{z}/H_1}(\mathbf{z}, A) = \frac{1}{\pi^m |\mathbf{\Sigma}|} \exp(-(\mathbf{z} - A\mathbf{p})^H \mathbf{\Sigma}^{-1} (\mathbf{z} - A\mathbf{p})).$$

The Log-Likelihood function $\log \frac{p_{\mathbf{z}/H_1}(\mathbf{z})}{p_{\mathbf{z}/H_0}(\mathbf{z})}$ can be simplified as:

$$\Lambda(\mathbf{z}) = \text{Re}(\mathbf{p}^H \mathbf{\Sigma}^{-1} \mathbf{z}) \underset{H_0}{\overset{H_1}{\gtrless}} \lambda.$$

The statistic of the test becomes:

$$\Lambda(\mathbf{z}) \sim \mathcal{N}(0, \mathbf{p}^H \mathbf{\Sigma}^{-1} \mathbf{p}) \text{ under } H_0 \quad \text{and} \quad \Lambda(\mathbf{z}) \sim \mathcal{N}(\text{Re}(A^H \mathbf{p}^H \mathbf{\Sigma}^{-1} \mathbf{p}), \mathbf{p}^H \mathbf{\Sigma}^{-1} \mathbf{p}) \text{ under } H_1$$

Example 2 - Matched Filter (1)

Problem under study: $\begin{cases} \text{Hypothesis } H_0: & \mathbf{z} = \mathbf{b}, \\ \text{Hypothesis } H_1: & \mathbf{z} = A\mathbf{p} + \mathbf{b}, \end{cases}$

where A is **unknown** complex scalar amplitude, \mathbf{p} is the **known** steering vector and $\mathbf{b} \sim \mathcal{CN}(\mathbf{0}_m, \mathbf{\Sigma})$ with **known** covariance matrix $\mathbf{\Sigma}$. The probability density functions of the received m -vector \mathbf{z} under each hypothesis are given by:

$$p_{\mathbf{z}/H_0}(\mathbf{z}) = \frac{1}{\pi^m |\mathbf{\Sigma}|} \exp(-\mathbf{z}^H \mathbf{\Sigma}^{-1} \mathbf{z}), \quad p_{\mathbf{z}/H_1}(\mathbf{z}, A) = \frac{1}{\pi^m |\mathbf{\Sigma}|} \exp(-(\mathbf{z} - A\mathbf{p})^H \mathbf{\Sigma}^{-1} (\mathbf{z} - A\mathbf{p})).$$

Maximizing $p_{\mathbf{z}/H_1}(\mathbf{z}, A)$ with respect to A leads to the MLE \hat{A} : $\hat{A} = \frac{\mathbf{p}^H \mathbf{\Sigma}^{-1} \mathbf{z}}{\mathbf{p}^H \mathbf{\Sigma}^{-1} \mathbf{p}}$. Replacing it in the Log-Likelihood Ratio test, we obtain the well-known *Matched Filter*:

$$\Lambda_{MF}(\mathbf{z}) = \log \frac{\max_A p_{\mathbf{z}/H_1}(\mathbf{z}, A)}{p_{\mathbf{z}/H_0}(\mathbf{z})} = \frac{|\mathbf{p}^H \mathbf{\Sigma}^{-1} \mathbf{z}|^2}{\mathbf{p}^H \mathbf{\Sigma}^{-1} \mathbf{p}} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda.$$

Example 2 - Matched Filter - Derivation of Performances (2)

Let $SNR = |A|^2 \mathbf{p}^H \boldsymbol{\Sigma}^{-1} \mathbf{p}$ be the Signal to Noise Ratio of the target to be detected.

Under H_0 hypothesis, $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}_m, \boldsymbol{\Sigma})$ and $\Lambda_{MF}(\mathbf{z}) \sim \frac{1}{2} \chi^2(2)$. We have:

$$P_{fa} = \mathbb{P}(\Lambda_{MF}(\mathbf{z}) > \lambda_{MF}/H_0) = \int_{\lambda_{MF}}^{+\infty} e^{-u} du = \exp(-\lambda_{MF}),$$

$$\lambda_{MF} = -\log P_{fa}.$$

Under H_1 hypothesis, $\mathbf{z} \sim \mathcal{CN}(A\mathbf{p}, \boldsymbol{\Sigma})$ and $\Lambda_{MF}(\mathbf{z}, \hat{A}) \sim \frac{1}{2} \chi^2(2, 2 SNR)$. We have:

$$P_d = \mathbb{P}(\Lambda_{MF}(\mathbf{z}, \hat{A}) > \lambda_{MF}/H_1) = 1 - F_{\chi^2(2, \delta)}(2\lambda_{MF}),$$

where $F_{\chi^2(2, \delta)}(\cdot)$ is the cumulative $\chi^2(2, \delta)$ density function with non-centrality parameter $\delta = 2 SNR = 2 A^2 \mathbf{p}^H \boldsymbol{\Sigma}^{-1} \mathbf{p}$.

Example 3 - Normalized Matched Filter (1)

Problem under study: $\begin{cases} \text{Hypothesis } H_0: & \mathbf{z} = \mathbf{b}, \\ \text{Hypothesis } H_1: & \mathbf{z} = A\mathbf{p} + \mathbf{b}, \end{cases}$

where A is **unknown** complex scalar amplitude, \mathbf{p} is the **known** steering vector and $\mathbf{b} \sim \mathcal{CN}(\mathbf{0}_m, \sigma^2 \mathbf{\Sigma})$ with **known covariance matrix** $\mathbf{\Sigma}$ but **unknown variance** σ^2 . The probability density functions of the received m -vector \mathbf{z} under each hypothesis are given by:

$$p_{\mathbf{z}/H_0}(\mathbf{z}, \sigma^2) = \frac{1}{\pi^m \sigma^{2m} |\mathbf{\Sigma}|} \exp\left(-\frac{\mathbf{z}^H \mathbf{\Sigma}^{-1} \mathbf{z}}{\sigma^2}\right) \quad p_{\mathbf{z}/H_1}(\mathbf{z}, A) = \frac{1}{\pi^m \sigma^{2m} |\mathbf{\Sigma}|} \exp\left(-\frac{(\mathbf{z} - A\mathbf{p})^H \mathbf{\Sigma}^{-1} (\mathbf{z} - A\mathbf{p})}{\sigma^2}\right).$$

- Maximizing $p_{\mathbf{z}/H_0}(\mathbf{z}, \sigma^2)$ with respect to σ^2 leads to the MLE: $\hat{\sigma}^2 = \frac{\mathbf{z}^H \mathbf{\Sigma}^{-1} \mathbf{z}}{m}$.
- Maximizing $p_{\mathbf{z}/H_1}(\mathbf{z}, \sigma^2, A)$ with respect to σ^2 and with respect to A leads to the MLEs:

$$\hat{\sigma}^2 = \frac{1}{m} \left(\mathbf{z}^H \mathbf{\Sigma}^{-1} \mathbf{z} - \frac{|\mathbf{p}^H \mathbf{\Sigma}^{-1} \mathbf{z}|^2}{\mathbf{p}^H \mathbf{\Sigma}^{-1} \mathbf{p}} \right) \quad \text{and} \quad \hat{A} = \frac{\mathbf{p}^H \mathbf{\Sigma}^{-1} \mathbf{z}}{\mathbf{p}^H \mathbf{\Sigma}^{-1} \mathbf{p}}.$$

Example 3 - Normalized Matched Filter (2)

Replacing it in the Log-Likelihood Ratio test, we obtain the well-known *Normalized Matched Filter*:

$$\Lambda_{NMF}(\mathbf{z}) = \log \frac{\max_A \max_{\sigma^2} p_{\mathbf{z}/H_1}(\mathbf{z}, \sigma^2, A)}{\max_{\sigma^2} p_{\mathbf{z}/H_0}(\mathbf{z}, \sigma^2)} = \frac{|\mathbf{p}^H \Sigma^{-1} \mathbf{z}|^2}{(\mathbf{p}^H \Sigma^{-1} \mathbf{p}) (\mathbf{z}^H \Sigma^{-1} \mathbf{z})} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{NMF}.$$

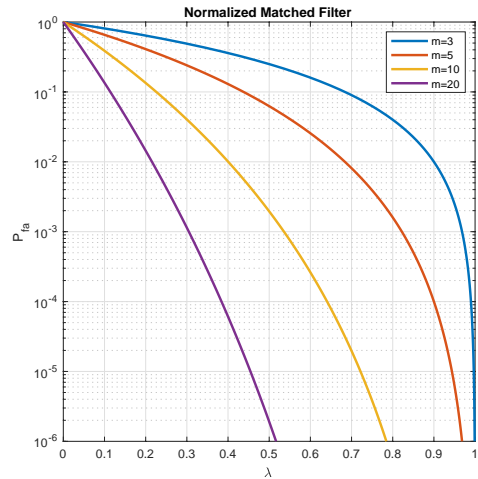
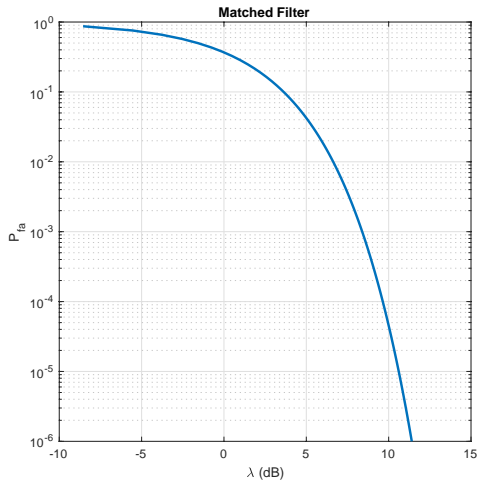
We can note that the NMF **is invariant with respect to a change scale** for \mathbf{p} , \mathbf{z} or Σ . Let $SNR = |A|^2 \mathbf{p}^H \Sigma^{-1} \mathbf{p}$ be the Signal to Noise Ratio of the target to be detected. Under H_0 hypothesis, $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}_m, \sigma^2 \Sigma)$ and $\Lambda(\mathbf{z}) \sim \beta(1, m-1)$. We have:

$$P_{fa} = \mathbb{P}(\Lambda_{NMF}(\mathbf{z}) > \lambda_{NMF}/H_0) = (1 - \lambda_{NMF})^{m-1},$$

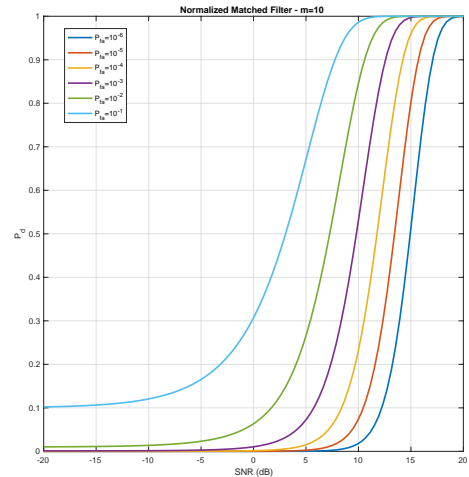
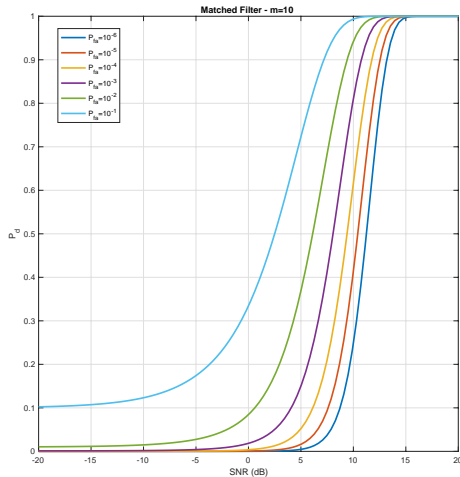
$$\lambda = 1 - P_{fa}^{1/(m-1)}.$$

We can note that the threshold λ_{NMF} does not depend on unknown variance σ^2 . The test is CFAR under H_0 hypothesis.

MF and NMF False Alarm regulation



MF and NMF Probability of Detection



Example 4 - Kelly and Adaptive Matched Filter (1)

Problem under study: $\begin{cases} \text{Hypothesis } H_0: & \mathbf{z} = \mathbf{b}, & \mathbf{z}_i = \mathbf{b}_i, & i = 1, \dots, n, \\ \text{Hypothesis } H_1: & \mathbf{z} = A\mathbf{p} + \mathbf{b}, & \mathbf{z}_i = \mathbf{b}_i, & i = 1, \dots, n. \end{cases}$

where the \mathbf{z}_i 's are n "signal-free" independent secondary data used to estimate the noise parameters, where A is **unknown** complex scalar amplitude, \mathbf{p} is the **known** steering vector and $\mathbf{b} \sim \mathcal{CN}(\mathbf{0}_m, \mathbf{\Sigma})$ with **unknown** covariance matrix $\mathbf{\Sigma}$. The probability density function of the received m -vector \mathbf{z} under hypothesis H_0 is given by:

$$p_{\mathbf{z}, \{\mathbf{z}_k\}_k, \mathbf{\Sigma} / H_0}(\mathbf{z}) = \frac{1}{\pi^m (n+1) |\mathbf{\Sigma}|^{n+1}} \exp \left(-\text{Tr} \left(\mathbf{\Sigma}^{-1} \left(\mathbf{z} \mathbf{z}^H + \sum_{k=1}^n \mathbf{z}_k \mathbf{z}_k^H \right) \right) \right).$$

With formulas $\frac{\delta \log |\mathbf{\Sigma}^{-1}|}{\delta \mathbf{\Sigma}^{-1}} = \mathbf{\Sigma}^T$ and $\frac{\delta \text{Tr}(\mathbf{\Sigma}^{-1} \mathbf{B})}{\delta \mathbf{\Sigma}^{-1}} = \mathbf{B}^T$, we obtain:

$$\arg\max_{\mathbf{\Sigma}} p_{\mathbf{z}, \{\mathbf{z}_k\}_k, \mathbf{\Sigma} / H_0}(\mathbf{z}) = \frac{1}{n+1} \left(\mathbf{z} \mathbf{z}^H + \sum_{k=1}^n \mathbf{z}_k \mathbf{z}_k^H \right).$$

Example 4 - Kelly and Adaptive Matched Filter (2)

The probability density function of the received m -vector \mathbf{z} under hypothesis H_1 is given by:

$$p_{\mathbf{z},\{\mathbf{z}_k\}_k,\boldsymbol{\Sigma},A/H_1}(\mathbf{z}) = \frac{1}{\pi^{m(n+1)} |\boldsymbol{\Sigma}|^{n+1}} \exp \left(-\text{Tr} \left(\boldsymbol{\Sigma}^{-1} \left((\mathbf{z} - A\mathbf{p}) (\mathbf{z} - A\mathbf{p})^H + \sum_{k=1}^n \mathbf{z}_k \mathbf{z}_k^H \right) \right) \right).$$

By denoting $\mathbf{S} = \sum_{k=1}^n \mathbf{z}_k \mathbf{z}_k^H$, we obtain $\arg\max_{\boldsymbol{\Sigma}} p_{\mathbf{z},\{\mathbf{z}_k\}_k,\boldsymbol{\Sigma},A/H_1}(\mathbf{z}) = \frac{(\mathbf{z} - A\mathbf{p}) (\mathbf{z} - A\mathbf{p})^H + \mathbf{S}}{n+1}$
and replacing these two expressions in the Generalized Log Likelihood Ratio leads to:

$$\Lambda(\mathbf{z}) = \frac{|\mathbf{z} \mathbf{z}^H + \mathbf{S}|}{\min_A |(\mathbf{z} - A\mathbf{p}) (\mathbf{z} - A\mathbf{p})^H + \mathbf{S}|} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda.$$

If we note $\mathbf{z}_s = \mathbf{S}^{-1/2} \mathbf{z}$ and $\mathbf{p}_s = \mathbf{S}^{-1/2} \mathbf{p}$, we have:

$$|(\mathbf{z} - A\mathbf{p}) (\mathbf{z} - A\mathbf{p})^H + \mathbf{S}| = |\mathbf{S}| \left| (\mathbf{z}_s - A\mathbf{p}_s) (\mathbf{z}_s - A\mathbf{p}_s)^H + \mathbf{I}_m \right| = |\mathbf{S}| (\|\mathbf{z}_s - A\mathbf{p}_s\|^2 + 1)$$

and $\min_A |\mathbf{S}| (\|\mathbf{z}_s - A\mathbf{p}_s\|^2 + 1) = |\mathbf{S}| (\|\mathbf{P}_{\mathbf{p}_s}^\perp \mathbf{z}_s\|^2 + 1)$ where $\mathbf{P}_{\mathbf{p}_s}^\perp = \mathbf{I}_m - \mathbf{p}_s \mathbf{p}_s^H / \mathbf{p}_s^H \mathbf{p}_s$.

Example 4 - Kelly and Adaptive Matched Filter (3)

We obtain the following Generalized Likelihood Ratio test, known as the so-called *Kelly's test* [Kelly 86]:

$$\Lambda_{\text{Kelly}}(\mathbf{z}) = \frac{|\mathbf{p}^H \mathbf{S}^{-1} \mathbf{z}|^2}{(\mathbf{p}^H \mathbf{S}^{-1} \mathbf{p}) (1 + \mathbf{z}^H \mathbf{S}^{-1} \mathbf{z})} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{\text{Kelly}} \quad \text{where} \quad \mathbf{S} = \sum_{k=1}^n \mathbf{z}_k \mathbf{z}_k^H.$$

This detector has good properties but is often (usually) replaced by a simpler one (so-called *two-step*), the *Adaptive Matched Filter* [Robey 92]:

$$\Lambda_{\text{AMF}}(\mathbf{z}) = \frac{|\mathbf{p}^H \hat{\mathbf{S}}_n^{-1} \mathbf{z}|^2}{\mathbf{p}^H \hat{\mathbf{S}}_n^{-1} \mathbf{p}} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{\text{AMF}} \quad \text{where} \quad \hat{\mathbf{S}}_n = \frac{1}{n} \sum_{k=1}^n \mathbf{z}_k \mathbf{z}_k^H.$$

The covariance matrix estimate $\hat{\mathbf{S}}_n = \frac{1}{n} \mathbf{S}$ is the *empirical* covariance matrix of the secondary data $\{\mathbf{z}_k\}_{k \in [1, n]}$ and is called *Sample Covariance Matrix* estimate.

Example 5 - Adaptive Normalized Matched Filter (1)

Detection in quasi-homogeneous Gaussian Noise: Problem under study:

$$\begin{cases} \text{Hypothesis } H_0: & \mathbf{z} = \mathbf{b}, & \mathbf{z}_i = \mathbf{b}_i, & i = 1, \dots, n, \\ \text{Hypothesis } H_1: & \mathbf{z} = A\mathbf{p} + \mathbf{b}, & \mathbf{z}_i = \mathbf{b}_i, & i = 1, \dots, n, \end{cases}$$

where the \mathbf{z}_i 's are n "signal-free" independent secondary data used to estimate the noise parameters, where A is **unknown** complex scalar amplitude, \mathbf{p} is the **known** steering vector, where $\mathbf{b}_i \sim \mathcal{CN}(\mathbf{0}_m, \mathbf{\Sigma})$ and $\mathbf{b} \sim \mathcal{CN}(\mathbf{0}_m, \sigma^2 \mathbf{\Sigma})$ with **unknown** covariance matrix $\mathbf{\Sigma}$ and **unknown** variance σ^2 . The PDF under each hypothesis is given by [Bandiera 09]:

$$p_{\mathbf{z}, \{\mathbf{z}_k\}_k, \mathbf{\Sigma} / H_0}(\mathbf{z}) = \frac{1}{\pi^{m(n+1)} |\mathbf{\Sigma}|^{n+1}} \exp \left(-\mathbf{z}^H \mathbf{\Sigma}^{-1} \mathbf{z} + \sum_{k=1}^n \mathbf{z}_k^H \mathbf{\Sigma}^{-1} \mathbf{z}_k \right),$$

$$p_{\mathbf{z}, \{\mathbf{z}_k\}_k, \mathbf{\Sigma}, \sigma^2, A / H_1}(\mathbf{z}) = \frac{1}{\pi^{m(n+1)} \sigma^{2m} |\mathbf{\Sigma}|^{n+1}} \exp \left(-\frac{(\mathbf{z} - A\mathbf{p})^H \mathbf{\Sigma}^{-1} (\mathbf{z} - A\mathbf{p})}{\sigma^2} + \sum_{k=1}^n \mathbf{z}_k^H \mathbf{\Sigma}^{-1} \mathbf{z}_k \right).$$

Example 5 - Adaptive Normalized Matched Filter (2)

The corresponding detector [Scharf 94, Kraut 99] is homogeneous of degree 0 with the variables \mathbf{p} , $\hat{\mathbf{S}}_n$ and \mathbf{z} and is named *Adaptive Normalized Matched Filter* (ANMF):

$$\Lambda_{ANMF}(\mathbf{z}) = \frac{|\mathbf{p}^H \hat{\mathbf{S}}_n^{-1} \mathbf{z}|^2}{(\mathbf{p}^H \hat{\mathbf{S}}_n^{-1} \mathbf{p}) (\mathbf{z}^H \hat{\mathbf{S}}_n^{-1} \mathbf{z})} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{ANMF} \quad \text{where} \quad \hat{\mathbf{S}}_n = \frac{1}{n} \sum_{k=1}^n \mathbf{z}_k \mathbf{z}_k^H.$$

ANMF and Cosine Estimate

This detector is often called a **Cosine Estimator** as it has the dimension of a cosine squared between the steering vector \mathbf{p} and the observation \mathbf{z} :

$$\Lambda_{ANMF}(\mathbf{z}) = \cos^2(\widehat{\mathbf{p}}, \mathbf{z}).$$

Unlike the AMF which characterizes the power of a scalar product, the ANMF measures an angle. It is so more sensible to a possible mismatch between \mathbf{p} and \mathbf{z} ([P. Develter 23]).

Example 6 - Persymmetric Adaptive Matched Filter (1)

Many applications can result in a clutter covariance matrix that exhibits **some particular structure**. For example, radars that use a symmetrically spaced linear array for spatial domain processing, or a symmetrically spaced pulse train for temporal domain processing.

- In these systems, the clutter covariance matrix Σ has the persymmetric property:

$$\Sigma = \mathbf{J}_m \Sigma^* \mathbf{J}_m ,$$

where \mathbf{J}_m is the m -dimensional antidiagonal matrix having 1 as non-zero elements.

- The signal vector is also persymmetric, i.e. it satisfies: $\mathbf{p} = \mathbf{J}_m \mathbf{p}^*$.
- The persymmetric structure of Σ can be exploited to improve its estimation accuracy compared to the SCM.

Example 6 - Persymmetric Adaptive Matched Filter (2)

We can build a two-step AMF built with the persymmetric Maximum Likelihood (ML) estimate of the clutter covariance matrix instead of the SCM. Problem under study:

$$\begin{cases} \text{Hypothesis } H_0: & \mathbf{x} = \mathbf{T} \mathbf{z} = \mathbf{T} \mathbf{b}, & \mathbf{x}_i = \mathbf{T} \mathbf{z}_i = \mathbf{T} \mathbf{b}_i, & i = 1, \dots, n, \\ \text{Hypothesis } H_1: & \mathbf{x} = \mathbf{T} \mathbf{z} = \mathbf{A} \mathbf{T} \mathbf{p} + \mathbf{T} \mathbf{b}, & \mathbf{x}_i = \mathbf{T} \mathbf{z}_i = \mathbf{T} \mathbf{b}_i, & i = 1, \dots, n, \end{cases}$$

where \mathbf{T} is the unitary matrix defined by:

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

Through this unitary transformation, secondary data $\mathbf{x}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{R})$ where $\mathbf{R} = \mathbf{T} \mathbf{\Sigma} \mathbf{T}^H$ is a **real covariance matrix**

Example 6 - Persymmetric Adaptive Matched Filter (3)

Let us now investigate the ML estimate of the **real** covariance matrix \mathbf{R} from the n transformed secondary data \mathbf{x}_k . The ML estimate $\hat{\mathbf{R}}$ of real matrix \mathbf{R} is unbiased and is given by:

$$\hat{\mathbf{R}} = \mathcal{Re}(\hat{\mathbf{R}}_n),$$

where $\mathcal{Re}(\cdot)$ stands for the real part, and where:

$$\hat{\mathbf{R}}_n = \frac{1}{n} \sum_{k=1}^n \mathbf{x}_k \mathbf{x}_k^H = \mathbf{T} \hat{\mathbf{S}}_n \mathbf{T}^H \text{ where } \hat{\mathbf{S}}_n = \frac{1}{n} \sum_{k=1}^n \mathbf{z}_k \mathbf{z}_k^H.$$

- $n\hat{\mathbf{R}}$ is real Wishart distributed with $2n$ degrees of freedom with parameter $\frac{1}{2}\mathbf{R}$,
- This result could be retrieved by the COvariance Matching Estimation Technique (COMET) procedure [Ottersen 98]!

Example 6 - Persymmetric Adaptive Matched Filter (4)

The distribution of this new detector under hypothesis H_0 can be derived. Replacing $\hat{\mathbf{R}}$ in the AMF (two-step procedure) leads to the following detection test, called the P-AMF:

$$\Lambda_{PAMF} = \frac{\left| \mathbf{s}^T \hat{\mathbf{R}}^{-1} \mathbf{x} \right|^2}{\mathbf{s}^T \hat{\mathbf{R}}^{-1} \mathbf{s}} \underset{H_0}{\overset{H_1}{\geq}} \lambda_{PAMF},$$

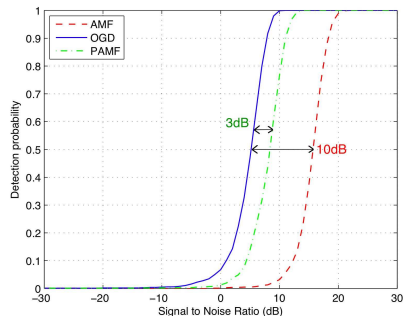
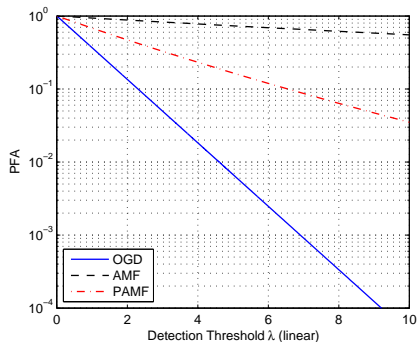
where $\mathbf{s} = \mathbf{T} \mathbf{p}$. In terms of the original data, we have, equivalently:

$$\Lambda_{PAMF} = \frac{\left| \mathbf{p}^H \mathbf{T}^H \left[\mathcal{R}e \left(\mathbf{T} \hat{\mathbf{S}}_n \mathbf{T}^H \right) \right]^{-1} \mathbf{T} \mathbf{z} \right|^2}{\mathbf{p}^H \mathbf{T}^H \left[\mathcal{R}e \left(\mathbf{T} \hat{\mathbf{S}}_n \mathbf{T}^H \right) \right]^{-1} \mathbf{T} \mathbf{p}} \underset{H_0}{\overset{H_1}{\geq}} \lambda_{PAMF}.$$

In the ML estimation procedure, **taking into account the real structure of \mathbf{R}** , or equivalently, the persymmetric structure of $\mathbf{\Sigma}$, **virtually doubles the amount of secondary data**.

Example 6 - Persymmetric Adaptive Matched Filter (5)

Theoretical λ/P_{fa} relationship: $P_{fa} = {}_2F_1\left(\frac{2n-m+1}{2}, \frac{2n-m+2}{2}, \frac{2n+1}{2}; -\frac{\lambda_{PAMF}}{n}\right)$.



- Left figure: Threshold decreasing brought by the P-AMF compared to the AMF for $n = 25$ and $m = 20$.
- Right figure: Improvement of about 7dB in terms of detection for the PAMF compared to the AMF for this set of parameters.

Example 7 - Anomaly Detector (1)

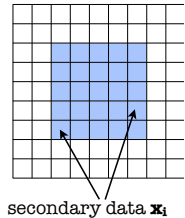
Model : $\begin{cases} \text{Hypothesis } H_0 : \mathbf{x}_i = \mathbf{b}_i, & i = 1, \dots, n \\ \text{Hypothesis } H_1 : \mathbf{x}_i = \alpha_i \mathbf{p} + \mathbf{b}_i, & i = 1, \dots, n \end{cases}$ where \mathbf{p} , $\{\alpha_i\}_{i \in [1, n]}$ are unknown and $\{\mathbf{b}_i\}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$. If we note $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)^T$ and $\mathbf{X} = \begin{bmatrix} \mathbf{x}_1(1) & \dots & \mathbf{x}_n(1) \\ \vdots & \ddots & \vdots \\ \mathbf{x}_1(m) & \dots & \mathbf{x}_n(m) \end{bmatrix}$. The **RXD GLRT**

test (Reed and Yu, 90) is defined as:

$$\Lambda_{RXD}(\mathbf{X}) = \frac{(\mathbf{X}\boldsymbol{\alpha}^T)^H (\mathbf{X}\mathbf{X}^H)^{-1} (\mathbf{X}\boldsymbol{\alpha}^T)}{\boldsymbol{\alpha}\boldsymbol{\alpha}^H}$$

Taking a particular $\boldsymbol{\alpha} = [0, \dots, 0, 1, 0, \dots, 0]^T$, a more simple and well-known RXD version yields (the signal under test \mathbf{x}_i is present in the covariance estimation!):

$$\Lambda_{RXD}(\mathbf{x}_i) = \mathbf{x}_i^H \hat{\mathbf{S}}_n^{-1} \mathbf{x}_i \underset{H_0}{\overset{H_1}{\geq}} \lambda$$



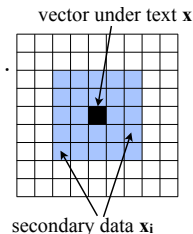
Example 7 - Anomaly Detector (2)

Model : $\begin{cases} \mathcal{H}_0 : \mathbf{x} = \mathbf{b}, & , \{\mathbf{x}_i = \mathbf{b}_i\}_i, i = 1, \dots, n \\ \mathcal{H}_1 : \mathbf{x} = \alpha \mathbf{p} + \mathbf{b}, & , \{\mathbf{x}_i = \mathbf{b}_i\}_i, i = 1, \dots, n \end{cases}$ where $\{\mathbf{b}_i\}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$, α and \mathbf{p} are unknown. The **Kelly GLRT test** (Frontera, 14) is defined as:

$$\Lambda_{RXD}(\mathbf{x}) = \mathbf{x}^H \hat{\mathbf{S}}_n^{-1} \mathbf{x} \underset{H_0}{\overset{H_1}{\geq}} \lambda$$

that corresponds to the **Mahalanobis distance**.

The Kelly test is Hotelling T^2 distributed: $\frac{n-m}{m(n+1)} RXD_{SCM}(\mathbf{c}/H_0) \sim F_{m,n-m}$.



Outline

- 1 Radar basis
 - Parameter Estimation
 - Noise and Clutter in Radar
- 2 Conventional Radar and Imaging Processing
 - Range-Doppler Radar Processing
 - Array and Space-Time Adaptive Processing
 - SAR Image Processing
 - Hyperspectral Image Processing
- 3 Some Background on Detection Theory
 - Problem Statement
 - Modeling Homogeneous Gaussian Noise/Clutter
 - Examples of Detector Derivations
 - **Synthesis of CFAR Detection Schemes Under Gaussian Noise**
- 4 Motivations for more robust detection schemes
 - Examples of Gaussian Hypothesis Failure
 - Need of Better Approaches

Synthesis of CFAR Detection Schemes Under Gaussian Noise (1)

- Adaptive Matched Filter [Robey 92]:

$$\Lambda_{AMF}(\mathbf{z}) = \frac{|\mathbf{p}^H \hat{\mathbf{S}}_n^{-1} \mathbf{z}|^2}{\mathbf{p}^H \hat{\mathbf{S}}_n^{-1} \mathbf{p}} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{AMF}$$

$$P_{fa} = {}_2F_1 \left(n - m + 1, n - m + 2; n + 1; -\frac{\lambda_{AMF}}{n} \right),$$

- Adaptive Kelly Filter [Kelly 86]:

$$\Lambda_{Kelly}(\mathbf{z}) = \frac{|\mathbf{p}^H \hat{\mathbf{S}}_n^{-1} \mathbf{z}|^2}{(\mathbf{p}^H \hat{\mathbf{S}}_n^{-1} \mathbf{p}) (n + \mathbf{z}^H \hat{\mathbf{S}}_n^{-1} \mathbf{z})} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{Kelly}$$

$$P_{fa} = \left(\frac{1}{\lambda_{Kelly}} - 1 \right)^{n+1-m},$$

Synthesis of CFAR Detection Schemes Under Gaussian Noise (2)

- Adaptive Normalized Matched Filter [*Scharf 94, Kraut 99*]:

$$\Lambda_{ANMF}(\mathbf{z}) = \frac{|\mathbf{p}^H \hat{\mathbf{S}}_n^{-1} \mathbf{z}|^2}{(\mathbf{p}^H \hat{\mathbf{S}}_n^{-1} \mathbf{p}) (\mathbf{z}^H \hat{\mathbf{S}}_n^{-1} \mathbf{z})} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{ANMF} :$$

$$P_{fa} = (1 - \lambda_{ANMF})^{n-m+1} {}_2F_1(n-m+2, n-m+1; n+1; \lambda_{ANMF}) .$$

- Persymmetric Adaptive Matched Filter [*Pailloux 09*]:

$$\Lambda_{PAMF} = \frac{\left| \mathbf{p}^H \mathbf{T}^H \left[\mathcal{R}e \left(\mathbf{T} \hat{\mathbf{S}}_n \mathbf{T}^H \right) \right]^{-1} \mathbf{T} \mathbf{z} \right|^2}{\mathbf{p}^H \mathbf{T}^H \left[\mathcal{R}e \left(\mathbf{T} \hat{\mathbf{S}}_n \mathbf{T}^H \right) \right]^{-1} \mathbf{T} \mathbf{p}} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{PAMF} .$$

$$P_{fa} = {}_2F_1 \left(\frac{2n-m+1}{2}, \frac{2n-m+2}{2}, \frac{2n+1}{2}; -\frac{\lambda_{PAMF}}{n} \right) .$$

The particular case of conventional Range Doppler 1/2

If we assume the noise is **white Gaussian** with **known covariance matrix** $\Sigma = \sigma^2 \mathbf{I}$, then the conventional detection scheme $\Lambda_{MF}(\mathbf{z}) = \frac{|\mathbf{p}^H \Sigma^{-1} \mathbf{z}|^2}{\mathbf{p}^H \Sigma^{-1} \mathbf{p}} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{MF}$ leads to the well known simplified test:

$$\Lambda_{MF}(\mathbf{z}) = \frac{|\mathbf{p} \mathbf{z}|^2}{\mathbf{p}^H \mathbf{p}} \underset{H_0}{\overset{H_1}{\gtrless}} \sigma^2 \lambda_{MF},$$

The test consists, for each range bin, in comparing the normalized Discrete Fourier Transform of the vector \mathbf{z} acquired for m pulses to a threshold. The corresponding PFA/threshold relationship is defined as:

$$\lambda_{MF} = -\sigma^2 \log P_{fa}.$$

The conventional Range Doppler algorithm makes implicitly assumption that the noise is white. In clutter environment, this processing is not optimal and we generally do not know the power σ^2 of the noise.

The particular case of conventional Range Doppler 2/2

If we assume the noise is **Gaussian** with **unknown covariance matrix**, then we have to use the conventional detection scheme $\Lambda_{AMF}(\mathbf{z}) = \frac{|\mathbf{p}^H \hat{\mathbf{S}}_n^{-1} \mathbf{z}|^2}{\mathbf{p}^H \hat{\mathbf{S}}_n^{-1} \mathbf{p}} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{AMF}$. For particular white noise with unknown power σ^2 , we can build a *simplified* two-step detection scheme, assuming that $\hat{\mathbf{S}}_n = \hat{\sigma}^2 \mathbf{I}$ where $\hat{\sigma}^2 = \frac{1}{m} \sum_{k=1}^m |\mathbf{p}^H \mathbf{z}_k|^2$. The new detection test becomes:

$$\Lambda_{AMF}(\mathbf{z}) = |\mathbf{p} \mathbf{z}|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \hat{\sigma}^2 \lambda_{AMF},$$

The test consists, for each range bin, in comparing the normalized Discrete Fourier Transform of the vector \mathbf{z} acquired for m pulses to a adaptive threshold built with secondary data $\{\mathbf{z}_k\}_{k \in [1, m]}$. The corresponding PFA/threshold relationship is defined as:

$$\lambda_{AMF} = m \left(P_{fa}^{-1/m} - 1 \right).$$

This threshold tends to λ_{MF} for large value m of secondary data.

Outline

- 1 Radar basis
 - Parameter Estimation
 - Noise and Clutter in Radar
- 2 Conventional Radar and Imaging Processing
 - Range-Doppler Radar Processing
 - Array and Space-Time Adaptive Processing
 - SAR Image Processing
 - Hyperspectral Image Processing
- 3 Some Background on Detection Theory
 - Problem Statement
 - Modeling Homogeneous Gaussian Noise/Clutter
 - Examples of Detector Derivations
 - Synthesis of CFAR Detection Schemes Under Gaussian Noise
- 4 Motivations for more robust detection schemes
 - **Examples of Gaussian Hypothesis Failure**
 - Need of Better Approaches

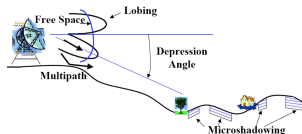
Examples of Gaussian Hypothesis Failure

High Resolution Radars

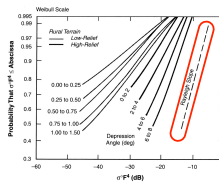
- Small number of scatterers in the cell under test - Varying number of scatterers from cell to cell - *Central Limit Theorem* non valid \Rightarrow non-Gaussianity [Jakeman 80]
- No validity of conventional tools based on Gaussian statistics [Farina 87, Gini 00, Jay 02].

Low-Grazing angles Illumination Radar

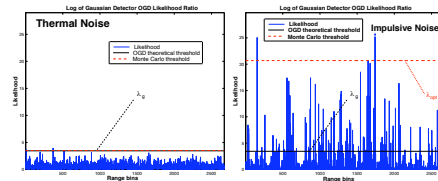
- Microshadowing \Rightarrow impulsive clutter [Billingsley 93]
- Transitions of clutter areas, heterogeneity of spatial area under test \Rightarrow difficulty to set up the detection test λ_{opt} and the Probability of False Alarm depending on the area.



Low-Grazing angle surveillance
Please refer to [F. Gini, A. Farina and M. S. Greco 2001]



Non-Gaussian behavior



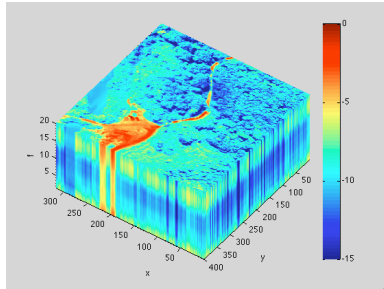
False Alarm regulation problem

Examples of Gaussian Hypothesis Failure

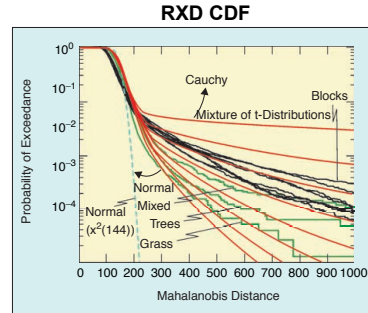


- The SAR images are more and more complex, detailed, heterogeneous. The spatial statistic of SAR images is not at all Gaussian,
- In polarimetry research field, almost all Non-Coherent Polarimetric Decomposition and classification techniques [Lee 09, Formont 2012] are generally based on conventional covariance matrix estimate (covariance or coherency matrix), typically the Sample Covariance Matrix (SCM),

Examples of Gaussian Hypothesis Failure



DSO data 2010



[Manolakis 2002]

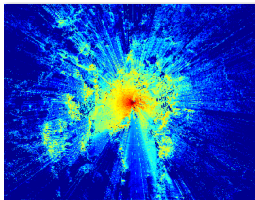
Bad regulation of False Alarm rate for Anomaly Detector [Reed 1990, Manolakis 2002, Ovarlez 2011, Frontera-Pons 2016] and detectors of targets [Frontera-Pons 2017] in Hyperspectral Images when they are based on conventional SCM estimate.

Outline

- 1 Radar basis
 - Parameter Estimation
 - Noise and Clutter in Radar
- 2 Conventional Radar and Imaging Processing
 - Range-Doppler Radar Processing
 - Array and Space-Time Adaptive Processing
 - SAR Image Processing
 - Hyperspectral Image Processing
- 3 Some Background on Detection Theory
 - Problem Statement
 - Modeling Homogeneous Gaussian Noise/Clutter
 - Examples of Detector Derivations
 - Synthesis of CFAR Detection Schemes Under Gaussian Noise
- 4 Motivations for more robust detection schemes
 - Examples of Gaussian Hypothesis Failure
 - **Need of Better Approaches**

Need of Better Approaches

Need to build alternatives to conventional approaches :



⇒ Better Covariance Matrix Estimation

Requirements:

- Background modeling: Compound 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.

Some solutions will be proposed in Part B

End of Part A

Questions?

Bibliography

- F. Bandiera, D. Orlando, G. Ricci, Advanced radar detection schemes under mismatched signal models. Morgan & Claypool, 2009.
- P. Woodward, Probability and information theory with applications to radar. London (UK): Pergamon, 1953.
- E. J. Kelly, An adaptive detection algorithm, *Aerospace and Electronic Systems, IEEE Transactions on*, **23**(1):115–127, 1986.
- A. W. Rihaczek, Principles of high resolution radar. Mac-Graw-Hill, 1969.
- J. B. Billingsley, Ground Clutter Measurements for Surface-Sited Radar, *Technical Report*, 780, MIT, February 1993.
- J. Ward, Space-time adaptive processing for airborne radar. *IET Conference Proceedings*, (1), pp.2–2, January 1998.
- M. Soumekh, Fourier array imaging. Englewood Cliffs: Prentice Hall, 1994.
- M. Soumekh, Synthetic Aperture Radar signal processing with Matlab algorithms. New York: John Wiley and Sons, 1999.
- D. Mensa High resolution radar imaging. USA: Artech House, 1981.
- J. Bertrand, P. Bertrand and J.-P. Ovarlez, Frequency directivity scanning in laboratory radar imaging. *International Journal of Imaging Systems and Technology*, 5(1), 39-51, March 1994.
- J.-P. Ovarlez, G. Ginolhac and A. M. Atto, Multivariate linear time-frequency modeling and adaptive robust target detection in highly textured monovariate SAR image, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4029–4033, 2017.

Bibliography

- A. Farina, A. Russo and F. Scannapieco, Radar detection in coherent Weibull clutter, *Acoustics, Speech, and Signal Processing, IEEE Transactions on*, **35**(6), June 1987.
- P. Formont, *Statistical and Geometric Tools for the Classification of Highly Textured Polarimetric SAR Images*, Ph.D. Thesis, Paris Saclay University / ONERA, 2012.
- A. Mian, *Robust Change Detection in Time Series SAR Images*, PhD Thesis, Université de Paris Saclay / ONERA, 2019.
- A. Mian, J.-P. Ovarlez, A. M. Atto and G. Ginolhac, Design of New Wavelet Packets Adapted to High-Resolution SAR Images With an Application to Target Detection, *Geoscience and Remote Sensing, IEEE Transactions on*, **57**(6), pp.3919-3932, June 2019
- I. Reed and X. Yu, Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution, *Acoustics, Speech and Signal Processing, IEEE Transactions on*, **38**(10):1760–1770, 1990.
- D. Manolakis and G. Shaw, Detection algorithms for hyperspectral imaging applications. *IEEE Signal Processing Magazine*, **19**(1), pp.2943, Jan. 2002.
- J.-P. Ovarlez, S. K. Pang, F. Pascal, V. Achard and T.K. Ng, Robust Detection using the SIRV Background Modelling for Hyperspectral Imaging, *Proc IEEE-IGARSS*, Vancouver, Canada, July 2011.
- J. Frontera-Pons, *Robust Detection and Classification for Hyperspectral Imaging*, PhD Thesis, Paris Saclay University / ONERA, 2014.
- J. Frontera-Pons, M. A. Veganzones, S. Velasco-Forero, F. Pascal, J.-P. Ovarlez, J. Chanussot, Robust Anomaly Detection in Hyperspectral Imaging, *Proc. IEEE-IGARSS*, Quebec, Canada, July 2014.

Bibliography

- J. Frontera-Pons, M. A. Veganzones, F. Pascal and J.-P. Ovarlez, Hyperspectral Anomaly Detectors using Robust Estimators, *Selected Topics in Applied Earth Observation and Remote Sensing (IEEE-JSTARS)*, *IEEE Journal of*, **9**(2), pp.720-731, 2016.
- J. Frontera-Pons, F. Pascal, J.-P. Ovarlez, Adaptive Nonzero-Mean Gaussian Detection, *Geoscience and Remote Sensing, IEEE Transactions on*, **55**(2), pp.1117-1124, 2017.
- J. Frontera-Pons, J.-P. Ovarlez and F. Pascal, Robust ANMF Detection in Noncentered Impulsive Background, *IEEE Signal Processing Letters*, **24**(12), pp.1891-1895, 2017.
- S. M. Kay. Fundamentals of Statistical Signal Processing: Estimation Theory, volume 1 of Prentice Hall Signal Processing Series. Prentice-Hall, Inc., 1993.
- S. M. Kay. Fundamentals of Statistical Signal Processing: Detection theory, volume 2 of Prentice Hall Signal Processing Series. Prentice-Hall PTR, 1998.
- F.C. Robey, D. R. Fuhrmann, E. J. Kelly and R. Nitzberg, A CFAR adaptive matched filter detector, *Aerospace and Electronic Systems, IEEE Transactions on*, **28**(1):208–216, 1992.
- E. J. Kelly, An adaptive detection algorithm. *Aerospace and Electronic Systems, IEEE Transactions on*, **23**(1), 115-127, November 1986.
- F. Bandiera, D. Orlando, G. Ricci, Advanced radar detection schemes under mismatched signal models. Morgan & Claypool, 2009.
- L. L. Scharf and B. Friedlander, Matched subspace detectors, *Signal Processing, IEEE Transactions on*, **42**(8):2146–2157, 1994.

Bibliography

- S. Kraut and L. L. Scharf, The CFAR adaptive subspace detector is a scale-invariant GLRT, *Signal Processing, IEEE Transactions on*, **47**(9):2538–2541, 1999.
- A. Farina, A. Russo and F. Scannapieco, Radar detection in coherent Weibull clutter, *Acoustics, Speech, and Signal Processing, IEEE Transactions on*, **35**(6), June 1987.
- E. Jay, *Detection in non-Gaussian noise*, PhD thesis, University of Cergy-Pontoise / ONERA, France, 2002.
- F. Gini, M. Greco, M. Diani and L. Verrazzani, Performance analysis of two adaptive radar detectors against non-Gaussian real sea clutter data. *Aerospace and Electronic Systems, IEEE Transactions on*, **36**(4), 1429-1439, Oct. 2000.
- E. Jakeman, On the statistics of K-distributed noise. *Journal of Physics A: Mathematical and General*, **13**(1), 31, 1980.
- J.-S. Lee and E. Pottier, *Polarimetric Radar Imaging, From Basics to Applications*, CRC Press, 2009.
- P. Formont, *Statistical and Geometric Tools for the Classification of Highly Textured Polarimetric SAR Images*, PhD Thesis, Paris Saclay University / ONERA, 2012.
- F. Gini, A. Farina, M. S. Greco, Selected List of References on Radar Signal Processing, *Aerospace and Electronic Systems, IEEE Transactions on*, **37**(1), Jan. 2001.
- B. Ottersten, P. Stoica, R. Roy, Covariance Matching Estimation Techniques for Array Signal Processing Applications, *Digital Signal Processing*, Vol. 8, Issue 3, pp.185-210, 1998.

A Tutorial on the Estimation and Detection Theory with Applications to Radar, Array Processing, and Imaging

Jean-Philippe Ovarlez^{1,2}

¹SONDRA, CentraleSupélec, France

²ONERA, The French Aerospace Lab, DEMR/MATS, France

jeanphilippe.ovarlez@centralesupelec.fr
<https://www.jeanphilippeovarlez.com>

T03 Tutorial - Conference Radar 2024 - 21 October 2024



Contents

- **Part A:**
Background on Statistical Processing for Radar, Array Processing, SAR and Hyperspectral Imaging,
- **Part B:**
Recent Methodologies on Robust Estimation and Detection in non-Gaussian Environment
- Applications and Results in Radar, STAP and Array Processing, SAR Imaging, Hyperspectral Imaging.

Part B

Recent Methodologies on Robust Estimation and
Detection in non-Gaussian Environment

Applications and Results in Radar, STAP and
Array Processing, SAR Imaging, Hyperspectral
Imaging

Part B: Contents

- 1 Robust Estimation and Detection
- 2 Other Refinements
- 3 Applications and Results in Radar, STAP, SAR imaging, Hyperspectral Imaging

Outline

1 Robust Estimation and Detection

■ Going to Robust Adaptive Detection

- Modeling the Background
- Robust Estimation
- Robust Detection
- Robustness of M-estimators

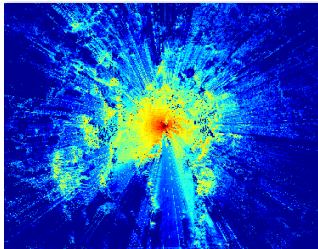
2 Other Refinements

- Exploiting Prior Information: Covariance Structure
- Low Rank Detectors
- Shrinkage of M -estimator
- Riemannian manifolds
- RMT Theory and M -Estimator based Detectors
 - RMT key ideas
 - Radar Detection Schemes for Joint Time and Spatial Correlated Clutter

3 Applications and Results in Radar, STAP, SAR imaging, Hyperspectral Imaging

- Surveillance Radar against Ground and Sea Clutter
- Detection Performance on STAP Data
- Detection Performance on SAR Image
- Hyperspectral Imaging: Detection and Anomaly Detection

Going to Robust Adaptive Detection



Generally, some parameters (e.g. second order statistic Σ) are unknown and cannot be estimated through Gaussian methodology

$$\underbrace{\mathbf{y}}_{\text{Heavy tailed}} = \underbrace{\mathbf{y}_{\mathcal{N}}}_{\text{Gaussian}} + \underbrace{\mathbf{o}}_{\text{Non Gaussian}}$$

\Rightarrow Robust Covariance Matrix Estimation

Requirements:

- Background modeling: Spherically Invariant Random Vectors (K-distribution, Weibull, etc.) [Conte 87, Barnard 96], Compound Gaussian [Conte 98, Sangston 12, 15], Complex Elliptically Symmetric (Multidimensional Generalized Gaussian Distributions, etc.) [Kelker 70, Frahm 04],
- Estimation procedure: ML-based approaches, M -estimation, etc.
- Adaptive detectors derivation and adaptive performance evaluation.

Outline

1 Robust Estimation and Detection

- Going to Robust Adaptive Detection
- **Modeling the Background**
- Robust Estimation
- Robust Detection
- Robustness of M-estimators

2 Other Refinements

- Exploiting Prior Information: Covariance Structure
- Low Rank Detectors
- Shrinkage of M -estimator
- Riemannian manifolds
- RMT Theory and M -Estimator based Detectors
 - RMT key ideas
 - Radar Detection Schemes for Joint Time and Spatial Correlated Clutter

3 Applications and Results in Radar, STAP, SAR imaging, Hyperspectral Imaging

- Surveillance Radar against Ground and Sea Clutter
- Detection Performance on STAP Data
- Detection Performance on SAR Image
- Hyperspectral Imaging: Detection and Anomaly Detection

Modeling the Background

Complex Elliptically Symmetric (CES) distributions:

Let \mathbf{z} be a complex circular random vector of length m . \mathbf{z} has a Complex Elliptically Symmetric (CES) distribution ($\mathcal{CES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_z)$) if its PDF is [Mahot 12, Ollila 12]:

$$f_{\mathbf{z}}(\mathbf{z}) = \delta_{m,g}^{-1} \frac{\Gamma(m)}{\pi^m |\boldsymbol{\Sigma}|} g_z \left((\mathbf{z} - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{z} - \boldsymbol{\mu}) \right) \text{ with } \delta_{m,g} = \int_0^{+\infty} t^{m-1} g_z(t) dt,$$

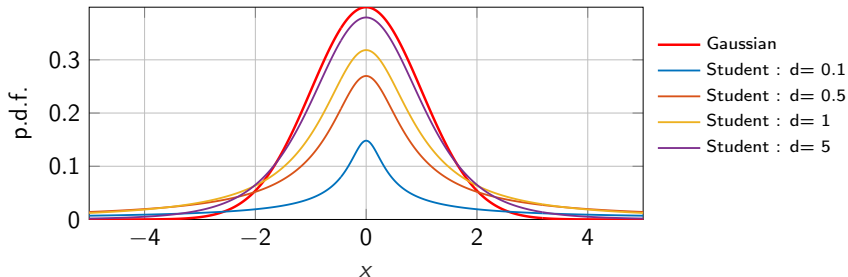
where $g_z : [0, \infty) \rightarrow [0, \infty)$ is the density generator, where $\boldsymbol{\mu}$ is the statistical mean (generally known or $\mathbf{0}_m$) and $\boldsymbol{\Sigma}$ is the scatter matrix. In general, $E[\mathbf{z} \mathbf{z}^H] = \alpha \boldsymbol{\Sigma}$ where α is known.

- **Large class of distributions:** Gaussian ($g_z(z) = \exp(-z)$), SIRV, MGGD ($g_z(z) = \exp(-z^\alpha)$), etc. **Validated through several experiments** [Billingsley 93, Ovarlez 95, Ovarlez 96],
- **Closed under affine transformations** (e.g. matched filter),
- **Stochastic representation theorem:** $\mathbf{z} =_d \boldsymbol{\mu} + \mathcal{R} \mathbf{A} \mathbf{u}^{(k)}$,

where the m -vector $\mathbf{u}^{(k)}$ is uniformly distributed on the sphere of radius 1, where $\mathcal{R} \geq 0$, independent of $\mathbf{u}^{(k)}$ and $\boldsymbol{\Sigma} = \mathbf{A} \mathbf{A}^H$ is a factorization of $\boldsymbol{\Sigma}$, where $\mathbf{A} \in \mathbb{C}^{m \times k}$ with $k = \text{rank}(\boldsymbol{\Sigma})$.

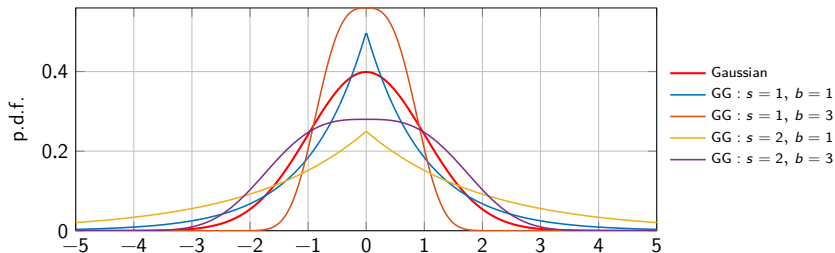
Example of CES

	Gaussian	Generalized Gaussian	t-distribution	W-distribution	K-distribution
	\mathbb{CN}_m	$\mathbb{CGN}_{m,s,b}$	$\mathbb{C}t_{m,d}$	$\mathbb{CW}_{m,s,b}$	$\mathbb{CK}_{m,\nu}$
$g(t)$	$\exp(-t)$	$\exp(-t^s/b) \quad s, b > 0$	$(1+t/d)^{-(d+m)} \quad d > 0$	$t^{s-1} \exp(-t^s/b) \quad s, b > 0$	$\sqrt{t}^{\nu-m} K_{\nu-m}(2\sqrt{\nu t}) \quad \nu > 0$
$\delta_{m,g}$	π^{-m}	$\frac{s\Gamma(m)b^{-m/s}}{\pi^m\Gamma(m/s)}$	$\frac{\Gamma(m+d)}{\pi^m d^m \Gamma(d)}$	$\frac{s\Gamma(m)b^{-(m+s-1)/s}}{\pi^m \Gamma((m+s-1)/s)}$	$2 \frac{\nu^{(\nu+m)/2}}{\pi^m \Gamma(\nu)}$



Example of CES

	Gaussian \mathcal{CN}_m	Generalized Gaussian $\mathcal{CGN}_{m,s,b}$	t -distribution $\mathcal{Ct}_{m,d}$	W -distribution $\mathcal{CW}_{m,s,b}$	K -distribution $\mathcal{CK}_{m,\nu}$
$g(t)$	$\exp(-t)$	$\exp(-t^s/b) \quad s, b > 0$	$(1 + t/d)^{-(d+m)} \quad d > 0$	$t^{s-1} \exp(-t^s/b) \quad s, b > 0$	$\sqrt{t}^{\nu-m} K_{\nu-m}(2\sqrt{\nu t}) \quad \nu > 0$
$\delta_{m,g}$	π^{-m}	$\frac{s\Gamma(m)b^{-m/s}}{\pi^m\Gamma(m/s)}$	$\frac{\Gamma(m+d)}{\pi^m d^m \Gamma(d)}$	$\frac{s\Gamma(m)b^{-(m+s-1)/s}}{\pi^m \Gamma((m+s-1)/s)}$	$2 \frac{\nu^{(\nu+m)/2}}{\pi^m \Gamma(\nu)}$



Modeling the Background

Spherically Invariant Random Vector: a CES subclass

The m -vector \mathbf{z} is a complex Spherically Invariant Random Vector [Yao 73, Jay 02] if its PDF can be put in the following form:

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

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 τ 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, spectral, spatial domain, etc.),
- Exploitation of the spectral information using the covariance matrix (*scatter matrix*) Σ ,
- **Stochastic representation theorem:** $\mathbf{z} = \mu + \sqrt{\tau} \mathbf{A} \mathbf{x}$, where $\tau \geq 0$ is the texture, independent of \mathbf{x} and $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}_m, \mathbf{I})$.

Modeling the Background

Compound-Gaussian Distribution

It can be assumed here that the n available secondary data are such that $\mathbf{z}_k = \sqrt{\tau_k} \mathbf{x}_k$ where $\mathbf{x}_k \sim \mathcal{CN}(\mathbf{0}_m, \mathbf{\Sigma})$ and where the textures $\{\tau_k\}_{k \in [1, n]}$ are **deterministic and unknown scalar variables** to be estimated.

$$g_{\mathbf{z}_k}(\mathbf{z}) = \frac{1}{\pi^m \tau_k^m |\mathbf{\Sigma}|} \exp \left(\frac{\mathbf{z}^H \mathbf{\Sigma}^{-1} \mathbf{z}}{\tau_k} \right).$$

- Conditionally to the bin k , the observed vector \mathbf{x}_k is Gaussian-distributed, i.e. $\mathbf{z}_k \sim \mathcal{CN}(\mathbf{0}_m, \tau_k \mathbf{\Sigma})$,
- The covariance matrix represents the spectral distribution of the noise through the support k ,
- The deterministic texture scalar τ 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, spectral, spatial domain, etc.).

Modeling the Background

Normalization of a CES vector \rightsquigarrow same underlying distribution

$$\mathbf{z} \sim \mathbb{CES}_m(\mathbf{0}, \Sigma, g) \Rightarrow \mathbf{z}_a = \frac{\mathbf{z}}{\|\mathbf{z}\|} \sim \mathbb{CAE}_m(\Sigma) \quad \text{for } \mathbf{z} \neq \mathbf{0}.$$

Complex Angular Elliptical (CAE) distribution:

👉 Probability density function [Greco & Gini 2013]: $f(\mathbf{z}_a; \Sigma) \propto |\Sigma|^{-1} (\mathbf{z}_a^H \Sigma^{-1} \mathbf{z}_a)^{-m}$

✓ Free from unknown density generator \rightsquigarrow robustness

✗ Scale ambiguity on Σ \rightsquigarrow additional constraint required
 \rightsquigarrow shape matrix: $\mathbf{V} = \frac{m}{\text{Tr}(\Sigma)} \Sigma$

Outline

1 Robust Estimation and Detection

- Going to Robust Adaptive Detection
- Modeling the Background
- **Robust Estimation**
- Robust Detection
- Robustness of M-estimators

2 Other Refinements

- Exploiting Prior Information: Covariance Structure
- Low Rank Detectors
- Shrinkage of M -estimator
- Riemannian manifolds
- RMT Theory and M -Estimator based Detectors
 - RMT key ideas
 - Radar Detection Schemes for Joint Time and Spatial Correlated Clutter

3 Applications and Results in Radar, STAP, SAR imaging, Hyperspectral Imaging

- Surveillance Radar against Ground and Sea Clutter
- Detection Performance on STAP Data
- Detection Performance on SAR Image
- Hyperspectral Imaging: Detection and Anomaly Detection

Estimating the Covariance/Scatter Matrix: Conventional Estimators

Assuming n available SIRV secondary data $\mathbf{z}_k = \sqrt{\tau_k} \mathbf{x}_k$ where $\mathbf{x}_k \sim \mathcal{CN}(\mathbf{0}_m, \mathbf{\Sigma})$ 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:

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

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

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

This estimate does not depend on the texture τ_k , but **it is biased** and shares the same eigenvectors but has different eigenvalues, with the same ordering [Bausson 07].

Maximum Likelihood Estimate of the Covariance/Scatter Matrix

MLE-estimators:

Example: Suppose n target-free i.i.d. m -vectors $\{\mathbf{z}_i\}_{i=1,n}$ where $\mathbf{z}_i \sim CE_m(\mathbf{0}_m, \Sigma, g_z)$ where $g_z(\cdot)$ is known and where Σ is an unknown scatter matrix. The MLE $\hat{\Sigma}$ is set by solving

$$\frac{\delta}{\delta \Sigma} \log \prod_{i=1}^n g_z(\mathbf{z}_i) = \frac{\delta}{\delta \Sigma^{-1}} \left(n \log |\Sigma^{-1}| + \sum_{i=1}^n \log h_z(\mathbf{z}_i^H \Sigma^{-1} \mathbf{z}_i) \right) = \mathbf{0}.$$

Recalling that $\frac{\delta}{\delta \Sigma^{-1}} \log |\Sigma^{-1}| = \Sigma^T$ and $\frac{\delta}{\delta \Sigma^{-1}} \log h_z(\mathbf{z}_i^H \Sigma^{-1} \mathbf{z}_i) = \frac{g'_z(\mathbf{z}_i \Sigma^{-1} \mathbf{z}_i^H)}{g_z(\mathbf{z}_i \Sigma^{-1} \mathbf{z}_i^H)} (\mathbf{z}_i \mathbf{z}_i^H)^T$, we obtain:

M-Estimator as MLE of the CES problem

$$\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^n \frac{-g'_z(\mathbf{z}_i^H \hat{\Sigma}^{-1} \mathbf{z}_i)}{g_z(\mathbf{z}_i^H \hat{\Sigma}^{-1} \mathbf{z}_i)} \mathbf{z}_i \mathbf{z}_i^H.$$

Estimating the Covariance/Scatter Matrix

M-estimators:

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

PDF $g_z(\cdot)$ specified: Maximum Likelihood-estimator of Σ : $\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^n \frac{-g_z' \left(\mathbf{z}_i^H \hat{\Sigma}^{-1} \mathbf{z}_i \right)}{g_z \left(\mathbf{z}_i^H \hat{\Sigma}^{-1} \mathbf{z}_i \right)} \mathbf{z}_i \mathbf{z}_i^H$,

PDF $g_z(\cdot)$ not specified: M-estimator of Σ : $\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^n u \left(\mathbf{z}_i^H \hat{\Sigma}^{-1} \mathbf{z}_i \right) \mathbf{z}_i \mathbf{z}_i^H$,

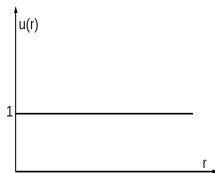
[Maronna 76, Kent 91, Maronna 06, Pascal 08, Mahot 13]

- Existence, Uniqueness, Asymptotic Properties,
- Convergence of the recursive algorithm, etc.
- Several PhD ONERA thesis: [Jay 02, Pascal 06, Mahot 12, Terreaux 18].

Examples of M -Estimators

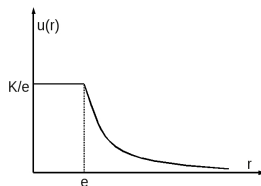
SCM:

$$u(r) = 1$$



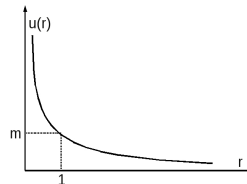
Huber's M -estimator:

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



Tyler:

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



- Huber = mix between SCM and Tyler [Huber 64],
- Tyler and SCM are “not” (theoretically) M -estimators,
- Tyler 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 $\sim CE_m(\mathbf{0}_m, \Sigma, g_z(\cdot))$ (Secondary data).

Tyler Estimator ([Tyler 87, Gini 02, Pascal 08])

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

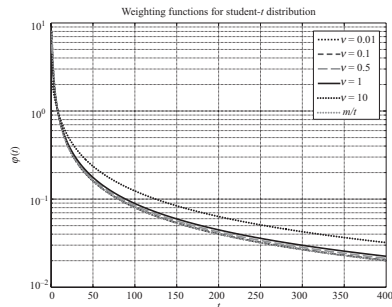
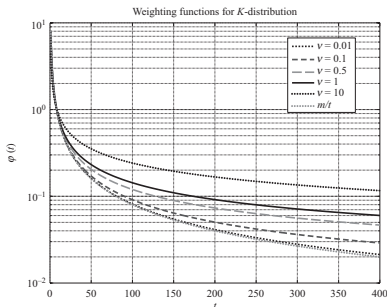
- The Tyler M-estimator **does not depend on the texture** (SIRV or CES distributions),
- Convergence of the algorithm: $\hat{\Sigma}_{n+1} = f(\hat{\Sigma}_n)$ with $f(\hat{\Sigma}) = \frac{m}{n} \sum_{k=1}^n \frac{\mathbf{z}_k \mathbf{z}_k^H}{\mathbf{z}_k^H \hat{\Sigma}^{-1} \mathbf{z}_k}$ and $\hat{\Sigma}_0 = \mathbf{I}_m$.

Existence, Uniqueness,

- $\hat{\Sigma}_{FPE}$ is the true Maximum Likelihood Estimate when considering textures $\{\tau_k\}_{k \in [1, n]}$ as unknown deterministic parameters. In that case, the joint texture estimation leads to

$$\hat{\tau}_k = \frac{\mathbf{z}_k^H \hat{\Sigma}_{FPE}^{-1} \mathbf{z}_k}{m}$$

Some Weighting Functions of M -estimators



$$u(t) = \varphi(t) = \frac{\sqrt{\nu}}{t} \frac{K_{\nu-m-1}(4\sqrt{\nu}t)}{K_{\nu-m}(4\sqrt{\nu}t)},$$

$$u(t) = \varphi(t) = \frac{\nu + 2m}{\nu + 2t}.$$

We have $\lim_{\nu \rightarrow 0} \hat{\Sigma} = \hat{\Sigma}_{FPE}$ and $\lim_{\nu \rightarrow \infty} \hat{\Sigma} = \hat{\Sigma}_n$.

Asymptotic distribution of complex M -estimators

Using the results of Tyler, we derived the following results [Mahot 2012, Mahot 2013]:

Theorem 1: Asymptotic distribution of $\hat{\Sigma}$

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

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

$$\begin{aligned} \mathbf{C} &= \sigma_1 (\Sigma^* \otimes \Sigma) + \sigma_2 \operatorname{vec}(\Sigma) \operatorname{vec}(\Sigma)^H, \\ \mathbf{P} &= \sigma_1 (\Sigma^* \otimes \Sigma) \mathbf{K}_{m^2, m^2} + \sigma_2 \operatorname{vec}(\Sigma) \operatorname{vec}(\Sigma)^T, \end{aligned}$$

where $\mathbf{K}_{m, m}$ is the $m \times m$ commutation matrix transforming any m -vector $\operatorname{vec}(\mathbf{A})$ into $\operatorname{vec}(\mathbf{A}^T)$ and where the constant σ_1 and σ_1 are completely defined.

An important property of complex M -estimators

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

$$\sqrt{n} \left(\text{vec} \left(\hat{\Sigma} - \Sigma \right) \right) \xrightarrow{d} \mathcal{CN} \left(\mathbf{0}_m, \mathbf{C}, \mathbf{P} \right), \quad (3)$$

with $\begin{cases} \mathbf{C} = \mathbf{v}_1 \Sigma^* \otimes \Sigma + \mathbf{v}_2 \text{vec}(\Sigma) \text{vec}(\Sigma)^H, \\ \mathbf{P} = \mathbf{v}_1 (\Sigma^* \otimes \Sigma) \mathbf{K}_{m^2, m^2} + \mathbf{v}_2 \text{vec}(\Sigma) \text{vec}(\Sigma)^T \end{cases}$, where \mathbf{v}_1 and \mathbf{v}_2 are any real numbers.

e.g.

	SCM	M -estimators	FPE
\mathbf{v}_1	1	σ_1	$(m+1)/m$
\mathbf{v}_2	0	σ_2	$-(m+1)/m^2$
...	More accurate		More robust

Known asymptotic behavior: **Any M -estimator behaves exactly as SCM but with σ_1 more secondary data ($\sigma_1 = (m+1)/m$ times more for Tyler)**: It implies that, in Gaussian case, SCM can be replaced by any M -estimate in previous detectors without changing performance (finite distance).

An important property of Tyler estimator

Tyler M-estimator:
$$\hat{\Sigma}_{FPE} = \frac{m}{n} \sum_{k=1}^n \frac{\mathbf{z}_k \mathbf{z}_k^H}{\mathbf{z}_k^H \hat{\Sigma}_{FPE}^{-1} \mathbf{z}_k}.$$

Theorem 2: Asymptotic distribution of $\hat{\Sigma}_{FPE} - \hat{\mathbf{S}}_n$

$$\sqrt{n} \left(\hat{\Sigma}_{FPE} - \hat{\mathbf{S}}_n \right) \xrightarrow{d} \mathcal{CN}(\mathbf{0}_m, \mathbf{C}_{FP}, \mathbf{P}_{FP}),$$

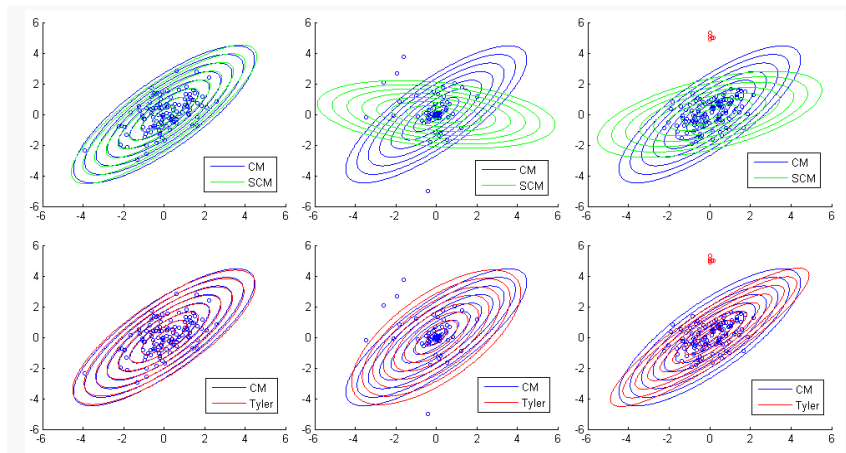
where \mathbf{C}_{FP} and \mathbf{P}_{FP} are defined as

$$\mathbf{C}_{FP} = \frac{1}{m} \left(\Sigma^T \otimes \Sigma \right) + \frac{m-1}{m^2} \text{vec}(\Sigma) \text{vec}(\Sigma)^H,$$

$$\mathbf{P}_{FP} = \frac{1}{m} \left(\Sigma^T \otimes \Sigma \right) \mathbf{K}_{m^2, m^2} + \frac{m-1}{m^2} \text{vec}(\Sigma) \text{vec}(\Sigma)^H.$$

Conclusion: $\left(\hat{\Sigma}_{FPE} - \hat{\mathbf{S}}_n \right)$ goes **faster** to **0** than $\left(\hat{\Sigma}_{FPE} - \Sigma \right)$ and then $\hat{\Sigma}_{FPE}$ behavior is **better approximated by the Wishart distribution** than by its asymptotic properties! [Draskovic 2019].

Tyler covariance matrix and SCM comparison in the presence of outliers



Outline

1 Robust Estimation and Detection

- Going to Robust Adaptive Detection
- Modeling the Background
- Robust Estimation
- **Robust Detection**
- Robustness of M-estimators

2 Other Refinements

- Exploiting Prior Information: Covariance Structure
- Low Rank Detectors
- Shrinkage of M -estimator
- Riemannian manifolds
- RMT Theory and M -Estimator based Detectors
 - RMT key ideas
 - Radar Detection Schemes for Joint Time and Spatial Correlated Clutter

3 Applications and Results in Radar, STAP, SAR imaging, Hyperspectral Imaging

- Surveillance Radar against Ground and Sea Clutter
- Detection Performance on STAP Data
- Detection Performance on SAR Image
- Hyperspectral Imaging: Detection and Anomaly Detection

An important property of complex M -estimators

- 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 [Mahot 13, Ovarlez 15]:

Theorem 3: (Asymptotic distribution of $H(\hat{\Sigma})$)

$$\sqrt{n} \left(H(\hat{\Sigma}) - H(\Sigma) \right) \xrightarrow{d} \mathcal{CN}(\mathbf{0}_r, \mathbf{C}_H, \mathbf{P}_H), \quad (4)$$

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

$$\begin{aligned} \mathbf{C}_H &= \nu_1 H'(\Sigma) \left(\Sigma^T \otimes \Sigma \right) H'(\Sigma)^H, \\ \mathbf{P}_H &= \nu_1 H'(\Sigma) \left(\Sigma^T \otimes \Sigma \right) \mathbf{K}_{m^2, m^2} H'(\Sigma)^T, \end{aligned}$$

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

$H(\text{SCM})$ and $H(\text{M-estimators})$ share the same asymptotic distribution (differs from ν_1)

Illustration with the two-step GLRT ANMF

Adaptive Normalized Matched Filter detector

$$H(\hat{\Sigma}) = \Lambda_{ANMF}(\mathbf{z}, \hat{\Sigma}) = \frac{\left| \mathbf{p}^H \hat{\Sigma}^{-1} \mathbf{z} \right|^2}{\left(\mathbf{p}^H \hat{\Sigma}^{-1} \mathbf{p} \right) \left(\mathbf{z}^H \hat{\Sigma}^{-1} \mathbf{z} \right)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{ANMF},$$

where $\hat{\Sigma}$ stands for any M -estimators [Conte 95, Kraut 99].

- The ANMF is **scale-invariant (homogeneous of degree 0)**, i.e.

$$\forall \alpha, \beta \in \mathbb{R}, \Lambda_{ANMF}(\alpha \mathbf{z}, \beta \hat{\Sigma}) = \Lambda_{ANMF}(\mathbf{z}, \hat{\Sigma}).$$

- Its **asymptotic distribution** (conditionally to $\mathbf{z}!$) is known [Pascal 15, Ovarlez 15].

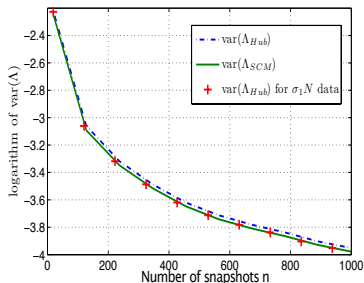
$$\sqrt{n} \left(H(\hat{\Sigma}) - H(\Sigma) \right) \xrightarrow{d} \mathcal{CN} \left(0, 2 \nu_1 H(\Sigma) (H(\Sigma) - 1)^2 \right).$$

$$\text{Recall for SCM: } \sqrt{n} \left(H(\hat{\mathbf{S}}) - H(\Sigma) \right) \xrightarrow{d} \mathcal{CN} \left(0, 2 H(\Sigma) (H(\Sigma) - 1)^2 \right).$$

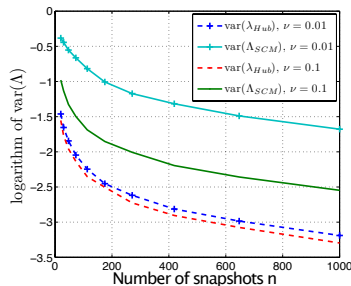
- It is **CFAR w.r.t the covariance/scatter matrix**,
- It is **CFAR w.r.t the texture**.

Illustrations of the Result on the ANMF

- $\Lambda = \text{var} \left(H \left(\hat{\Sigma} \right) - H(\Sigma) \right)$. Here $\hat{\Sigma}$ = complex Huber's M -estimator.
- Figure 1: Gaussian context, here $\sigma_1 = 1.066$.
- Figure 2: K-distributed clutter (shape parameter: $\nu = 0.1$ and 0.01).



Validation of theorem (even for small n)



Interest of the M -estimators

Performances are slightly the same in the Gaussian case but are better in the non-Gaussian case.

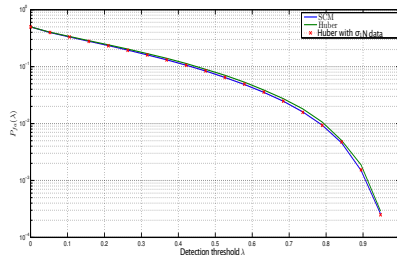
Illustrations of the Result on P_{fa}

- Figure 1: Gaussian context :

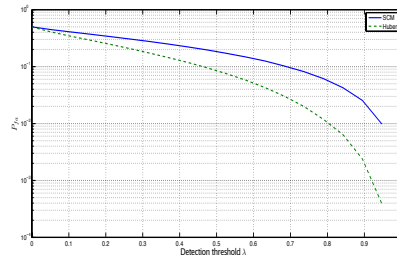
$$P_{fa} = (1 - \lambda_{ANMF})^{n-m+1} {}_2F_1(n-m+2, n-m+1; n+1; \lambda_{ANMF}) .$$

- Figure 2: K-distributed clutter (shape parameter: $\nu = 0.1$), here $\sigma_1 = 1.066$:

$$P_{fa} = (1 - \lambda_{ANMF})^{n/\sigma_1-m+1} {}_2F_1(n/\sigma_1-m+2, n/\sigma_1-m+1; n/\sigma_1+1; \lambda_{ANMF}) .$$



Validation of theorem (even for small n)



Interest of the M -estimators for False Alarm
regulation

Illustration of the Results on Multiple Signal Classification (MUSIC) method

- K (**known**) direction of arrival θ_k on m antennas
- Gaussian stationary narrowband signal with additive noise.
- the DoA [*Bienvenu 1979, Schmidt 1986*] is estimated from n snapshots, using the SCM, the Huber's M -estimator and the Tyler's estimator.

$$\mathbf{y}(t) = \mathbf{A}(\boldsymbol{\theta}_0) \mathbf{s}(t) + \mathbf{w}(t).$$

- $\boldsymbol{\theta}_0 = (\theta_1, \theta_2, \dots, \theta_K)^T$,
- the steering matrix $\mathbf{A}(\boldsymbol{\theta}) = (\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \dots, \mathbf{a}(\theta_K))$,
- $\mathbf{s}(t) = (s_1(t), s_2(t), \dots, s_K(t))^T$ signal vector,
- $\mathbf{w}(t)$ stationary additive noise.

$$\boldsymbol{\Sigma} = E[\mathbf{y} \mathbf{y}^H] = \mathbf{A}(\boldsymbol{\theta}_0) E[\mathbf{s} \mathbf{s}^H] \mathbf{A}^H(\boldsymbol{\theta}_0) + \sigma^2 \mathbf{I} = \mathbf{E}_S \mathbf{D}_S \mathbf{E}_S^H + \sigma^2 \mathbf{E}_W \mathbf{E}_W^H,$$

where \mathbf{E}_S (resp. \mathbf{E}_W) are the signal (resp. noise) subspace eigenvectors.

The MUSIC statistic is

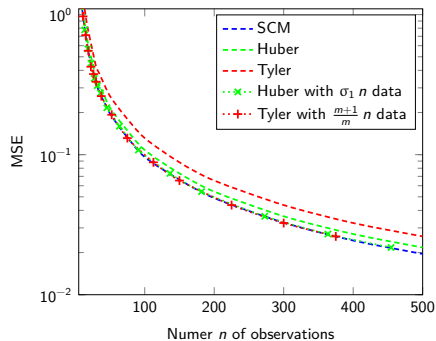
$$\begin{cases} H(\hat{\Sigma}) = \operatorname{argmax}_{\theta} \gamma(\theta) & \text{where } \gamma(\theta) = \mathbf{s}(\theta)^H \mathbf{E}_W \mathbf{E}_W^H \mathbf{s}(\theta), \\ \textcolor{red}{H}(\hat{\Sigma}) = \operatorname{argmax}_{\theta} \hat{\gamma}(\theta) & \text{where } \hat{\gamma}(\theta) = \sum_{i=1}^{m-K} \mathbf{s}(\theta)^H \hat{\mathbf{e}}_i \hat{\mathbf{e}}_i^H \mathbf{s}(\theta), \end{cases}$$

where $\hat{\mathbf{e}}_i$ are the eigenvectors of $\hat{\Sigma}$.

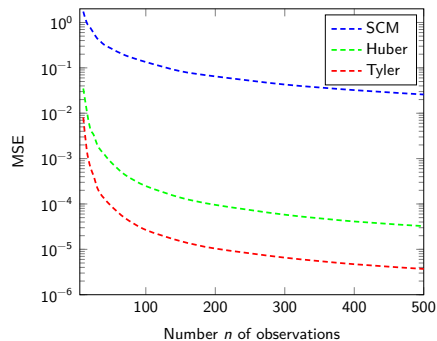
This function respects the assumptions of Theorem 3!

The Mean Square Error (MSE) between the estimated angle $\hat{\theta}$ and the exact angles θ can then be computed (case of one source).

- A $m = 3$ uniform linear array (ULA) with half wavelength sensors spacing is used,
- Gaussian stationary narrowband signal with DoA 20° plus additive noise.
- the DoA is estimated from n snapshots, using the SCM, the Huber's M -estimator and the Tyler's estimator.



(a) White additive Gaussian noise



(b) K-distributed additive noise ($\nu = 0.1$)

Figure: MSE of $\hat{\theta}$ for a number n of observations, with $m = 3$.

Similar conclusions for detection can be drawn...

Illustration of the ANMF CFAR Properties For CES Noise

False Alarm regulation for ANMF built with Tyler's estimate

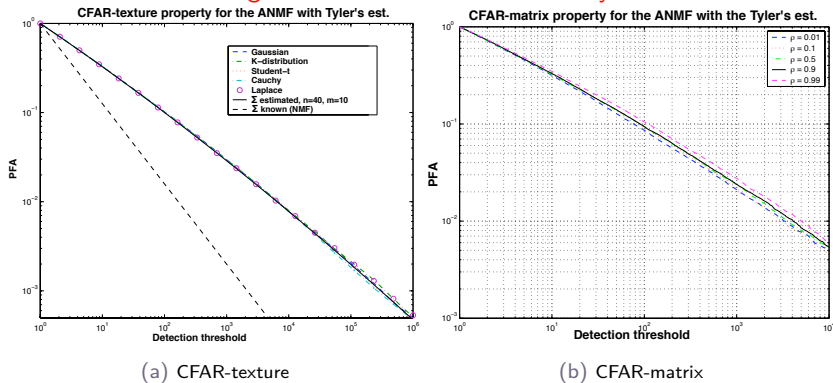
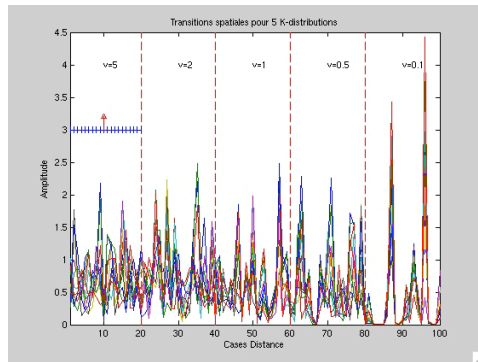
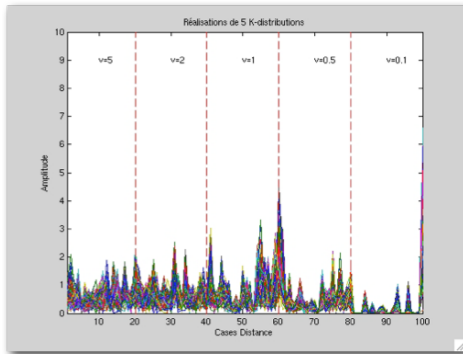


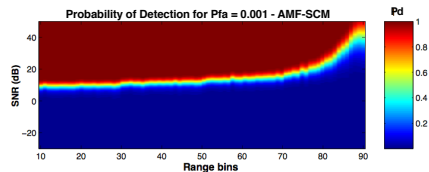
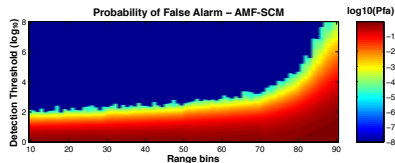
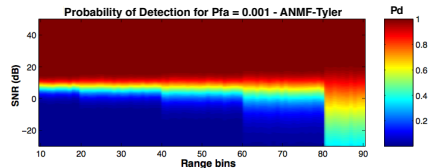
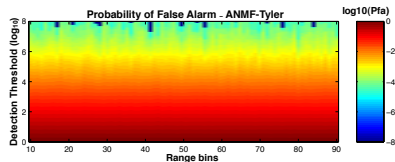
Figure: Illustration of the CFAR properties of the ANMF built with Tyler's estimator, for a Toeplitz CM whose (i, j) -entries are $\rho^{|i-j|}$.

Properties of ANMF-Tyler Detector on Clutter Transitions



- Five K-distributed clutter range transitions: from Gaussian to impulsive noise,
- Estimating the covariance matrix with secondary data in a 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 terms of detection is kept for moderate non-Gaussian clutter and improved for spiky clutter.

Outline

1 Robust Estimation and Detection

- Going to Robust Adaptive Detection
- Modeling the Background
- Robust Estimation
- Robust Detection
- **Robustness of M-estimators**

2 Other Refinements

- Exploiting Prior Information: Covariance Structure
- Low Rank Detectors
- Shrinkage of M -estimator
- Riemannian manifolds
- RMT Theory and M -Estimator based Detectors
 - RMT key ideas
 - Radar Detection Schemes for Joint Time and Spatial Correlated Clutter

3 Applications and Results in Radar, STAP, SAR imaging, Hyperspectral Imaging

- Surveillance Radar against Ground and Sea Clutter
- Detection Performance on STAP Data
- Detection Performance on SAR Image
- Hyperspectral Imaging: Detection and Anomaly Detection

Robustness of the M-estimators

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

- Sample Covariance Matrix case:

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

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

- Tyler (or FP) Covariance Matrix case:

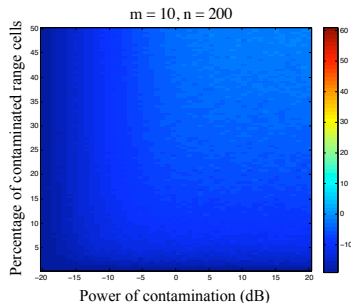
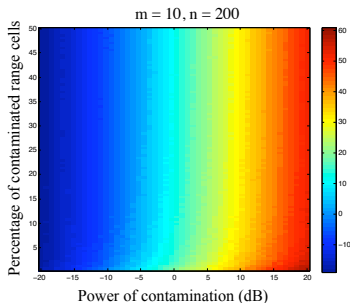
$$\hat{\mathbf{\Sigma}}_{FPEPol} = \frac{m}{n} \sum_{k=1}^n \frac{\mathbf{y}_k \mathbf{y}_k^H}{\mathbf{y}_k^H \hat{\mathbf{\Sigma}}_{FPEPol}^{-1} \mathbf{y}_k}, \quad E[\hat{\mathbf{\Sigma}}_{FPEPol}] = \mathbf{\Sigma} + \frac{m+1}{n} \left[E \left[\frac{\mathbf{p}_0 \mathbf{p}_0^H}{\mathbf{p}_0^H \mathbf{\Sigma}^{-1}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

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

Robustness of M-estimators

Gaussian vectors \mathbf{y}_k polluted by outliers

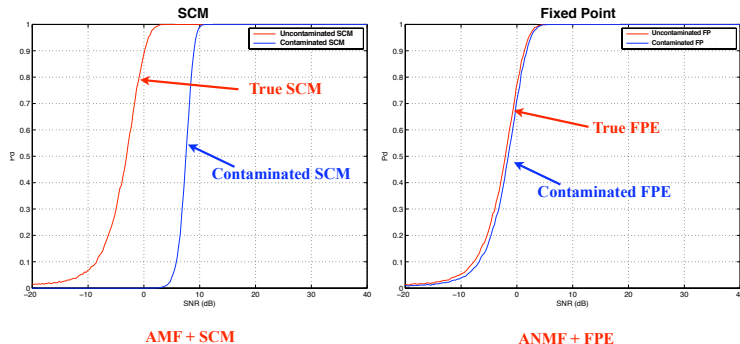
$$\hat{\mathbf{S}}_n = \frac{1}{n} \sum_{k=1}^n \mathbf{y}_k \mathbf{y}_k^H, \quad \hat{\mathbf{\Sigma}}_{FPE} = \frac{m}{n} \sum_{k=1}^n \frac{\mathbf{y}_k \mathbf{y}_k^H}{\mathbf{y}_k^H \hat{\mathbf{\Sigma}}_{FPE}^{-1} \mathbf{y}_k}.$$



Plot of the error (dB) 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)



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

Outline

1 Robust Estimation and Detection

- Going to Robust Adaptive Detection
- Modeling the Background
- Robust Estimation
- Robust Detection
- Robustness of M -estimators

2 Other Refinements

- Exploiting Prior Information: Covariance Structure
- Low Rank Detectors
- Shrinkage of M -estimator
- Riemannian manifolds
- RMT Theory and M -Estimator based Detectors
 - RMT key ideas
 - Radar Detection Schemes for Joint Time and Spatial Correlated Clutter

3 Applications and Results in Radar, STAP, SAR imaging, Hyperspectral Imaging

- Surveillance Radar against Ground and Sea Clutter
- Detection Performance on STAP Data
- Detection Performance on SAR Image
- Hyperspectral Imaging: Detection and Anomaly Detection

Motivations

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

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

- Toeplitz: *[Burg 82]* for estimation,
- known rank $r < m$ (ex: subspace detector) *[Kirsteins 94, Haimovich 96, Rangaswamy 03]*,
- Persymmetry: *[Nitzberg 80]* for estimation, *[Cai 92]* for detection in Gaussian case, *[De Maio 03, Conte 03, Pailloux 11]* in non-Gaussian noise.
- Shrinkage: when the number n of available secondary data does guarantee the inversion of the covariance matrix estimate ($n < m$). *[Abramovich 07, Chen 11, Abramovich 13, Besson 13, Couillet 14, Wiesel 14, Pascal 14]*
- In high dimension regime, some RMT-based results *[Couillet 11, 14, 15]* for detection schemes
- COvariance Matching Estimation techniques (COMET) *[Ottersten 98]* and its robust versions SESAME and RCOMET *[Meriaux 19, 20]*

Outline

1 Robust Estimation and Detection

- Going to Robust Adaptive Detection
- Modeling the Background
- Robust Estimation
- Robust Detection
- Robustness of M -estimators

2 Other Refinements

- **Exploiting Prior Information: Covariance Structure**
- Low Rank Detectors
- Shrinkage of M -estimator
- Riemannian manifolds
- RMT Theory and M -Estimator based Detectors
 - RMT key ideas
 - Radar Detection Schemes for Joint Time and Spatial Correlated Clutter

3 Applications and Results in Radar, STAP, SAR imaging, Hyperspectral Imaging

- Surveillance Radar against Ground and Sea Clutter
- Detection Performance on STAP Data
- Detection Performance on SAR Image
- Hyperspectral Imaging: Detection and Anomaly Detection

Covariance Matrix Convex Structure

Problem setup:

- ➡ N i.i.d., m -dimensional, centered CES distributed observations:

$$\forall n = 1, \dots, N \quad \mathbf{z}_n \sim \mathbb{CES}_m(\mathbf{0}, \mathbf{R}_e, g) \quad \rightsquigarrow \quad p_{\mathbf{Z}}(\mathbf{z}_n; \mathbf{R}_e)$$
- ➡ $\mathbf{R}_e \in \mathcal{S}$: convex subset of Hermitian matrices
- ➡ There exists a **one-to-one differentiable mapping** $\theta \mapsto \mathcal{R}(\theta)$ from $\mathcal{S}_\theta \subset \mathbb{R}^P$ to \mathcal{S}
- ➡ Unknown parameter of interest: θ with exact value θ_e and $\mathbf{R}_e = \mathcal{R}(\theta_e)$

Working hypothesis:

- ➡ *True* distribution unknown in practice $\Rightarrow g$ unknown
- ➡ *Assumed* model: $\mathbb{CES}_m(\mathbf{0}, \mathcal{R}(\theta), g_{\text{mod}}) \rightsquigarrow f_{\mathbf{Z}}(\mathbf{z}_n; \theta)$
 with $g_{\text{mod}}(t)$ possibly different from $g(t)$ for all $t \in \mathbb{R}^+$

SESAME Algorithm

SESAME (StructurEd ScAtter Matrix Estimator): two-step procedure

1 Unstructured estimate of \mathbf{R}

$$\hat{\mathbf{R}}_m = \frac{1}{N} \sum_{n=1}^N u_{\text{mod}} \left(\mathbf{z}_n^H \hat{\mathbf{R}}_m^{-1} \mathbf{z}_n \right) \mathbf{z}_n \mathbf{z}_n^H,$$

2 Projection on the subset \mathcal{S}

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\text{argmin}} \mathcal{J}_{\hat{\mathbf{R}}_m, \hat{\mathbf{R}}}(\boldsymbol{\theta}) \quad \text{with,}$$

$$\mathcal{J}_{\hat{\mathbf{R}}_m, \hat{\mathbf{R}}}(\boldsymbol{\theta}) = \kappa_1 \text{Tr} \left(\left[\hat{\mathbf{R}}^{-1} \left(\hat{\mathbf{R}}_m - \mathcal{R}(\boldsymbol{\theta}) \right) \right]^2 \right) + \kappa_2 \left[\text{Tr} \left(\hat{\mathbf{R}}^{-1} \left(\hat{\mathbf{R}}_m - \mathcal{R}(\boldsymbol{\theta}) \right) \right) \right]^2,$$

$$\text{where } \begin{cases} \kappa_1 = \kappa_2 + 1 = \frac{\mathbb{E} \left[\psi_{\text{mod}}^2 \left(\|\zeta\|_2^2 \right) \right]}{m(m+1)} \text{ with } \zeta \sim \mathbb{CES}_m(\mathbf{0}, \mathbf{I}, g_{\text{mod}}) \text{ and } \psi_{\text{mod}}(s) = s u_{\text{mod}}(s) \\ \hat{\mathbf{R}} \text{ any consistent estimator of } \mathbf{R}_e, \text{ up to a scale factor, e.g., } \hat{\mathbf{R}}_m \end{cases}$$

Strictly convex in $\mathcal{R}(\boldsymbol{\theta})$ + one-to-one mapping \Rightarrow unique solution for $\boldsymbol{\theta}$.

SESAME asymptotic performance

Theorem

[Mériaux et al., 2019] Let $\hat{\theta}_N$ be the SESAME estimate based on N i.i.d. observations, $\mathbf{z}_n \sim \mathbb{CES}_m(\mathbf{0}, \mathbf{R}_e, g)$ but with an assumed model of $\mathbb{CES}_m(\mathbf{0}, \mathcal{R}(\mu), g_{mod})$. Then, we obtain

👉 the consistency: $\hat{\theta}_N \xrightarrow{\mathcal{P}} \theta_c$ such that $\mathbf{R}_c \triangleq \mathcal{R}(\theta_c) = \sigma^{-1} \mathcal{R}(\theta_e)$

👉 the asymptotic distribution: $\sqrt{N}(\hat{\theta}_N - \theta_c) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \Gamma_\theta)$

$$\text{with } \begin{cases} \Gamma_\theta = (\kappa_1 \mathbf{C} + \kappa_2 \mathbf{D})^{-1} (\beta_1 \mathbf{C} + \beta_2 \mathbf{D}) (\kappa_1 \mathbf{C} + \kappa_2 \mathbf{D})^{-1}, \\ \mathbf{C} = \mathcal{J}(\theta_c)^H (\mathbf{R}_c^{-T} \otimes \mathbf{R}_c^{-1}) \mathcal{J}(\theta_c), \mathbf{D} = \mathcal{J}(\theta_c)^H \text{vec}(\mathbf{R}_c^{-1}) \text{vec}(\mathbf{R}_c^{-1})^H \mathcal{J}(\theta_c) \\ \beta_1 = \sigma_1 \kappa_1^2, \beta_2 = \sigma_1 \kappa_2 (2\kappa_1 + m\kappa_2) + \sigma_2 (\kappa_1 + m\kappa_2)^2 \end{cases}$$

Structured shape matrix estimator in CAE framework

RCOMET: Robust COvariance Matching Estimation Technique

$$\left(\hat{\alpha}, \hat{\theta} \right) = \underset{\alpha, \theta}{\operatorname{argmin}} \operatorname{Tr} \left(\left(\hat{\mathbf{R}}_{\text{FPE}} - \alpha \mathbf{R}(\theta) \right) \hat{\mathbf{R}}^{-1} \left(\hat{\mathbf{R}}_{\text{FPE}} - \alpha \mathbf{R}(\theta) \right) \hat{\mathbf{R}}^{-1} \right),$$

where $\alpha > 0$ and $\hat{\mathbf{R}}$ refers to any consistent estimator of \mathbf{R}_e up to a scale factor, e.g., $\hat{\mathbf{R}} = \hat{\mathbf{R}}_{\text{FPE}}$.

Convex problem w.r.t. $\alpha \mathbf{R}(\theta)$ + one-to-one mapping \Rightarrow unique solution for θ

Theorem (*Mériaux et al., 2017, Mériaux et al., 2019*)

Let $\hat{\theta}$ the RCOMET estimate of θ_e based on N i.i.d. observations, $\mathbf{y}_n \sim \mathcal{U}_m(\mathbf{R}(\theta_e))$. $\hat{\theta}$ is consistent, asymptotically efficient, and Gaussian:

$$\sqrt{N} \left(\hat{\theta} - \theta_e \right) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{CRB}_{\text{CAE}})$$

with $\mathbf{CRB}_{\text{CAE}} = \frac{m+1}{m} \mathbf{CRB}_G$, where $\mathbf{CRB}_G = \text{CRB on } \theta \text{ of the problem } \mathbf{z}_n \sim \mathcal{CN}(\mathbf{0}, \alpha_e \mathbf{R}(\theta_e))$

Recursive RCOMET

In the same way, we can define **Recursive RCOMET**

- A recursive procedure naturally follows

$$\text{for } k = 1, \dots, N_{\text{it}}, \quad \hat{\boldsymbol{\theta}}^{(k)} = \underset{\alpha, \boldsymbol{\theta}}{\operatorname{argmin}} \operatorname{Tr} \left[\left\{ \left(\hat{\mathbf{R}}_{\text{FPE}} - \alpha \mathbf{R}(\boldsymbol{\theta}) \right) \mathbf{R}(\hat{\boldsymbol{\theta}}_{k-1})^{-1} \right\}^2 \right]$$

$$\text{such that } \operatorname{Tr} \left[\mathbf{R}(\hat{\boldsymbol{\theta}}_k) \right] = m$$

- Leading to R-RCOMET estimate: $\hat{\boldsymbol{\theta}}_{\text{R-RCOMET}} = \hat{\boldsymbol{\theta}}^{(N_{\text{it}})}$.
- Same asymptotic properties as RCOMET, but the asymptotic regime is reached faster (numerical ascertainment).
- A more elaborated stopping rule in practice, e.g., a combination of $k \leq N_{\text{it}}$ and $\left\| \hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_{k-1} \right\| \leq \varepsilon_{\text{tol}} \left\| \hat{\boldsymbol{\theta}}_{k-1} \right\|$.

Using Persymmetry Property

Under persymmetric considerations (ex: symmetrically spaced linear array, symmetrically spaced pulse train, etc.), 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 one as non-zero elements. If the unitary matrix \mathbf{T} is defined by:

$$\mathbf{T} = \begin{cases} \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{I}_{m/2} & \mathbf{J}_{m/2} \\ i \mathbf{I}_{m/2} & -i \mathbf{J}_{m/2} \end{pmatrix} & \text{for } m \text{ even} \\ \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{I}_{(m-1)/2} & 0 & \mathbf{J}_{(m-1)/2} \\ 0 & \sqrt{2} & 0 \\ i \mathbf{I}_{(m-1)/2} & 0 & -i \mathbf{J}_{(m-1)/2} \end{pmatrix} & \text{for } m \text{ odd,} \end{cases} \quad (5)$$

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

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 $\mathbf{x}, \mathbf{x}_k \sim \mathcal{CN}(\mathbf{0}_m, \mathbf{R})$ 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 $\mathbf{\Sigma}$ is transformed through \mathbf{T} onto a real covariance matrix \mathbf{R} .

The Persymmetric FP Covariance Matrix Estimate

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

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

The persymmetric Fixed-Point Covariance Matrix Estimate has been first empirically defined as:

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

Statistical performance of $\hat{\mathbf{R}}_{PFP}$ [Pailloux 08, 10 and 11]:

- $\hat{\mathbf{R}}_{PFP}$ is a consistent and unbiased estimate of \mathbf{R} when n tends to infinity,
- Its asymptotic distribution is the same as the asymptotic distribution of a real Wishart matrix with $2nm/(m+1)$ degrees of freedom,
- RCOMET technique [Mériaux 19 and 20] gives exactly the same result for persymmetric structure.

$\hat{\mathbf{R}}_{PFP}$ estimate can be considered as the true Maximum Likelihood Estimate.

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:

$$\Lambda(\hat{\mathbf{R}}_{PFP}) = \frac{|s^T \hat{\mathbf{R}}_{PFP}^{-1} \mathbf{z}|^2}{(s^T \hat{\mathbf{R}}_{PFP}^{-1} s) (\mathbf{z}^H \hat{\mathbf{R}}_{PFP}^{-1} \mathbf{z})} \underset{H_0}{\overset{H_1}{\geq}} \lambda.$$

Properties:

- $\Lambda(\hat{\mathbf{R}}_{PFP})$ is texture-CFAR,
- $\Lambda(\hat{\mathbf{R}}_{PFP})$ 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.

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

More recent works can be found in [Mériaux 19 and 20]

Outline

1 Robust Estimation and Detection

- Going to Robust Adaptive Detection
- Modeling the Background
- Robust Estimation
- Robust Detection
- Robustness of M -estimators

2 Other Refinements

- Exploiting Prior Information: Covariance Structure
- **Low Rank Detectors**
- Shrinkage of M -estimator
- Riemannian manifolds
- RMT Theory and M -Estimator based Detectors
 - RMT key ideas
 - Radar Detection Schemes for Joint Time and Spatial Correlated Clutter

3 Applications and Results in Radar, STAP, SAR imaging, Hyperspectral Imaging

- Surveillance Radar against Ground and Sea Clutter
- Detection Performance on STAP Data
- Detection Performance on SAR Image
- Hyperspectral Imaging: Detection and Anomaly Detection

Conventional Low-Rank Detectors

Principle of Low-Rank Matched Filter approaches found, for example, in [Kirsteins 94] (Principal Component Inverse) and [Haimovich 96] (Eigencanceler) and [Rangaswamy 04].

Let suppose that 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**.

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

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

$$\Lambda_{LR-ANMF-SCM}(\mathbf{z}) = \frac{\left| \mathbf{p}^H \left(\mathbf{I} - \hat{\Pi}_{SCM} \right) \mathbf{z} \right|^2}{\left(\mathbf{p}^H \left(\mathbf{I} - \hat{\Pi}_{SCM} \right) \mathbf{p} \right) \left(\mathbf{z}^H \left(\mathbf{I} - \hat{\Pi}_{SCM} \right) \mathbf{z} \right)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda.$$

Extended Low-Rank Detectors

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

$$\mathbf{\Sigma}_{NSCM} = \frac{m}{n} \sum_{k=1}^n \frac{\mathbf{y}_k \mathbf{y}_k^H}{\mathbf{y}_k^H \mathbf{y}_k} = (\mathbf{U}_r \mathbf{U}_0) \begin{pmatrix} \mathbf{\Sigma}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_0 \end{pmatrix} (\mathbf{U}_r \mathbf{U}_0)^H$$

[Ginolhac 12 and 13] proved that $\mathbf{\Pi}_{NSCM} = \mathbf{U}_r \mathbf{U}_r^H$ is a consistent estimate projector onto the clutter subspace. We can define the extended Low-Rank ANMF-NSCM:

$$\Lambda_{LR-ANMF-NSCM}(\mathbf{y}) = \frac{\left| \mathbf{p}^H \left(\mathbf{I} - \hat{\mathbf{\Pi}}_{NSCM} \right) \mathbf{z} \right|^2}{\left(\mathbf{p}^H \left(\mathbf{I} - \hat{\mathbf{\Pi}}_{NSCM} \right) \mathbf{p} \right) \left(\mathbf{z}^H \left(\mathbf{I} - \hat{\mathbf{\Pi}}_{NSCM} \right) \mathbf{z} \right)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda.$$

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 the adaptive version of Rangaswami's Low-Rank Matched Filter and Low-Rank Normalized Matched Filter.

Outline

1 Robust Estimation and Detection

- Going to Robust Adaptive Detection
- Modeling the Background
- Robust Estimation
- Robust Detection
- Robustness of M -estimators

2 Other Refinements

- Exploiting Prior Information: Covariance Structure
- Low Rank Detectors
- **Shrinkage of M -estimator**
- Riemannian manifolds
- RMT Theory and M -Estimator based Detectors
 - RMT key ideas
 - Radar Detection Schemes for Joint Time and Spatial Correlated Clutter

3 Applications and Results in Radar, STAP, SAR imaging, Hyperspectral Imaging

- Surveillance Radar against Ground and Sea Clutter
- Detection Performance on STAP Data
- Detection Performance on SAR Image
- Hyperspectral Imaging: Detection and Anomaly Detection

Shrinkage of Tyler's estimators

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

Chen estimator [Chen 11]

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

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

- Originally introduced in [Abramovich 07],
- Existence, uniqueness and algorithm convergence proved in [Chen 11],
- Active research [Abramovich 13, Besson 13, Couillet 14, Wiesel 14, Pascal 14]

Shrinkage Tyler's estimators

Pascal estimator [Pascal 14]

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

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

- Σ_P (naturally) verifies $\text{Tr}(\Sigma_P^{-1}) = m$ for all $\beta \in (0, 1]$,
- Existence, uniqueness and algorithm convergence proved,
- The main challenge is to find the optimal β ! [Couillet 14].

Outline

1 Robust Estimation and Detection

- Going to Robust Adaptive Detection
- Modeling the Background
- Robust Estimation
- Robust Detection
- Robustness of M -estimators

2 Other Refinements

- Exploiting Prior Information: Covariance Structure
- Low Rank Detectors
- Shrinkage of M -estimator
- **Riemannian manifolds**
- RMT Theory and M -Estimator based Detectors
 - RMT key ideas
 - Radar Detection Schemes for Joint Time and Spatial Correlated Clutter

3 Applications and Results in Radar, STAP, SAR imaging, Hyperspectral Imaging

- Surveillance Radar against Ground and Sea Clutter
- Detection Performance on STAP Data
- Detection Performance on SAR Image
- Hyperspectral Imaging: Detection and Anomaly Detection

Riemannian Geometry

- Hermitian covariance matrices are definite-positive and **belong to a Riemannian manifold: conventional Euclidean space is not at all adapted to this space**,
- for characterizing the barycenter of covariance matrices; for example, in K-means classification, the arithmetical mean is not recommended. Please see the H/α SAR polarimetric classification application.

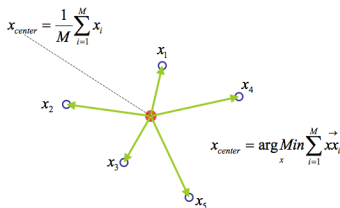
Euclidean mean (arithmetic)

$$\arg \min_{\mathbf{M} \in \mathcal{P}(m)} \sum_{i=1}^n d(\mathbf{M}, \mathbf{M}_i)^2, \text{ where } d(\mathbf{M}, \mathbf{M}_i) = \|\mathbf{M} - \mathbf{M}_i\|_F$$

Riemannian mean (geometric)

$$\arg \min_{\mathbf{M} \in \mathcal{P}(m)} \sum_{i=1}^n d(\mathbf{M}, \mathbf{M}_i)^2, \text{ where } d(\mathbf{M}, \mathbf{M}_i) = ?$$

Riemannian Geometry

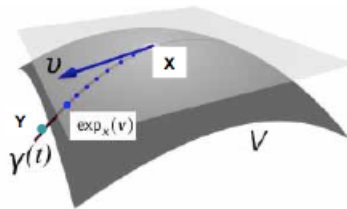


Euclidean Geometry

The mean M_ω in the **Euclidean** sense of n given positive-definite Hermitian matrices M_1, \dots, M_n in $\mathcal{P}(p)$ is defined as:

$$M_\omega = \operatorname{argmin}_{M_\omega \in \mathcal{P}(m)} \sum_{k=1}^n \|M_k - M_\omega\|_F^2$$

leading to $M_\omega = \frac{1}{n} \sum_{k=1}^n M_k$.



Riemannian Geometry

The mean M_ω in the **Riemannian** sense of n given positive-definite Hermitian matrices M_1, \dots, M_n in $\mathcal{P}(p)$ is defined as:

$$M_\omega = \operatorname{argmin}_{M_\omega \in \mathcal{P}(m)} \sum_{k=1}^n \left\| \log(M_k^{-1} M_\omega) \right\|_F^2$$

leading to M_ω such that $\frac{1}{n} \sum_{k=1}^n \log(M_k^{-1} M_\omega) = 0$.

Outline

1 Robust Estimation and Detection

- Going to Robust Adaptive Detection
- Modeling the Background
- Robust Estimation
- Robust Detection
- Robustness of M -estimators

2 Other Refinements

- Exploiting Prior Information: Covariance Structure
- Low Rank Detectors
- Shrinkage of M -estimator
- Riemannian manifolds
- **RMT Theory and M -Estimator based Detectors**
 - RMT key ideas
 - Radar Detection Schemes for Joint Time and Spatial Correlated Clutter

3 Applications and Results in Radar, STAP, SAR imaging, Hyperspectral Imaging

- Surveillance Radar against Ground and Sea Clutter
- Detection Performance on STAP Data
- Detection Performance on SAR Image
- Hyperspectral Imaging: Detection and Anomaly Detection

Some RMT-based results for detection schemes

The RMT (ex: [Couillet]) allows 1) to understand the statistical behavior of expressions involving estimate of large covariance matrices (ex: quadratic forms, ratios of the quadratic forms, SNIR Loss, performances of detection tests as ANMF, LR-ANMF, etc.) and 2) to correct it. The corrected results are often valid at a finite distance (practical m, N values).

- **Sources localisation applications** [F. Pascal, R. Couillet, etc.]: the based-RMT Music algorithm (G-Music) is known to have higher performance than conventional algorithms when using all the eigenvalues of the covariance matrix.
- **MIMO-STAP**: the goal of A. Combernoux's PHD thesis [Combernoux] was to analyze/improve the detection and filtering performances of low-rank detectors.
- **Adaptive Radar Detection**: when secondary data are correlated [Couillet].
- **Hyperspectral Anomaly Detection - Unmixing**: the goal of E. Terreaux PhD thesis [Terreaux] is to better analyse the rank of the anomalies space (model order selection) in Hyperspectral Imaging (high dimensional problem) for heterogeneous, correlated non-Gaussian environment.

RMT key ideas (1)

Let $\{\mathbf{y}_i\}_{i \in [1, N]}$ be independent and distributed according to $\mathcal{CN}(\mathbf{0}_m, \mathbf{M})$. The Maximum Likelihood Estimate of \mathbf{M} is the Sample Covariance Matrix given by

$$\hat{\mathbf{M}} = \frac{1}{N} \sum_{i=1}^N \mathbf{y}_i \mathbf{y}_i^H = \frac{1}{N} \mathbf{Y} \mathbf{Y}^H.$$

Asymptotic Regime

If $N \rightarrow \infty$, then the strong law of large numbers says (or equivalently, in spectral norm):

$$\|\hat{\mathbf{M}} - \mathbf{M}\| \xrightarrow{a.s.} 0.$$

Random Matrix Regime

- No longer valid if $m, N \rightarrow \infty$ with $m/N \rightarrow c \in [0, \infty[$: $\|\hat{\mathbf{M}} - \mathbf{M}\| \not\rightarrow 0$,
- For practical large m, N with $m \simeq N$, it can lead to dramatically wrong conclusions (even $m = N/100$).

RMT key ideas (2)

Let $\{\mathbf{n}_i\}_{i \in [1, N]}$ be distributed according to $\mathcal{CN}(\mathbf{0}_m, \mathbf{C} = \sigma^2 \mathbf{I}_m)$. We analyze the eigenvalues distribution of $\hat{\mathbf{C}} = \frac{1}{N} \sum_{i=1}^N \mathbf{n}_i \mathbf{n}_i^H = \frac{1}{N} \mathbf{N} \mathbf{N}^H$ where $c = m/N \in [0, \infty[$

Random Matrix Regime

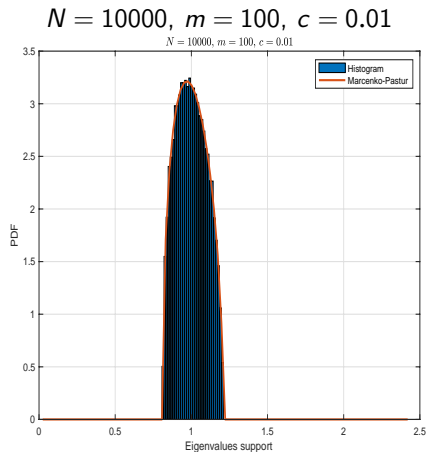
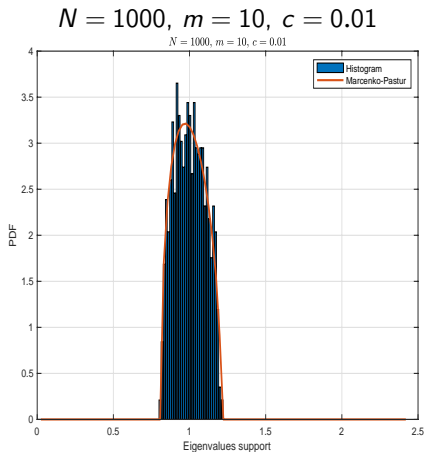
The distribution of the eigenvalues of $\hat{\mathbf{C}}$ tends almost surely toward the Marcenko-Pastur distribution

$$p(x) = \left(1 - \frac{1}{c}\right)_+ \delta(x) + \frac{1}{2\pi c x} \sqrt{(x - \lambda_-)_+ (\lambda_+ - x)_+},$$

where $\lambda_- = \sigma^2 (1 - \sqrt{c})^2$ and $\lambda_+ = \sigma^2 (1 + \sqrt{c})^2$.

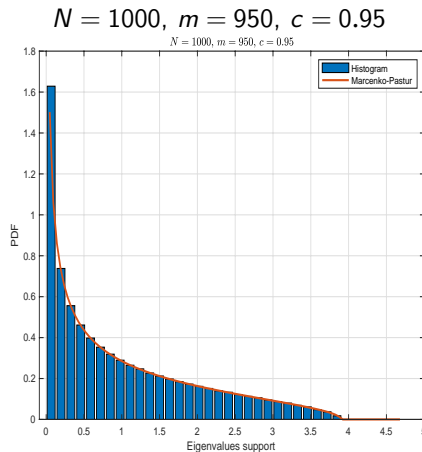
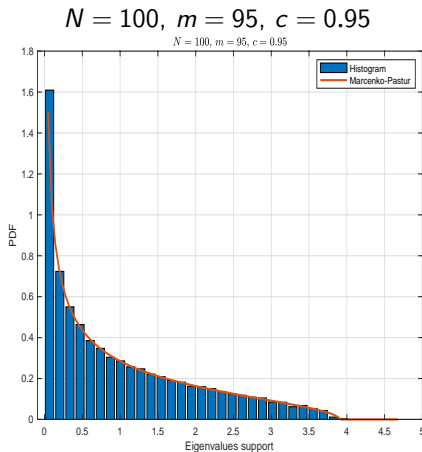
Not restricted to Gaussian statistics !

RMT Examples (1): classical asymptotic regime



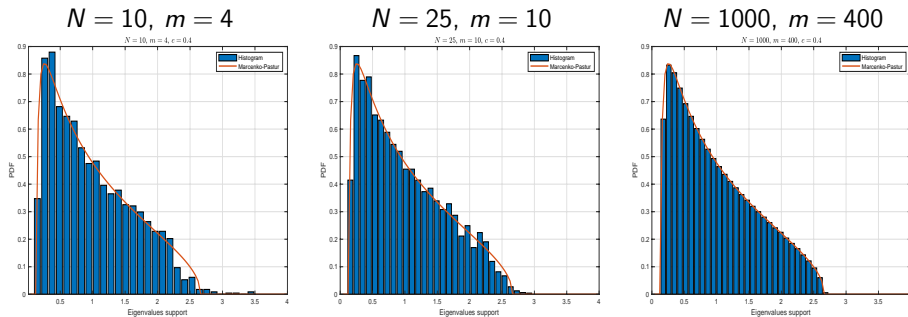
Eigenvalues support for **white** Gaussian noise ($\sigma^2 = 1, \mathbf{C} = \sigma^2 \mathbf{I}_m$).

RMT Examples (2): same RMT regime



Eigenvalues support for **white** Gaussian noise ($\sigma^2 = 1, \mathbf{C} = \sigma^2 \mathbf{I}_m$).

RMT Examples (3): from where does the RMT regime start?

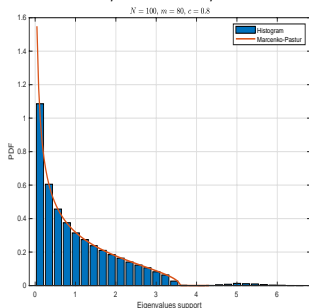


Eigenvalues support for **white** Gaussian noise ($\sigma^2 = 1$, $\mathbf{C} = \sigma^2 \mathbf{I}_m$) and $c = 0.4$.

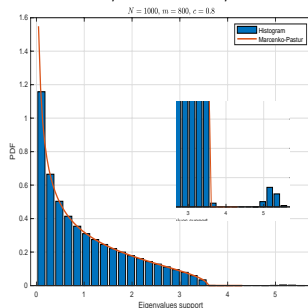
Key ideas (3)

The behavior of the spectral measure brings information about the vast majority of the eigenvalues but is not affected by some individual eigenvalues behavior (like sources !). Whatever the perturbations (sources), the spectral measure converges toward the Marcenko-Pastur distribution.

$N = 100, m = 80, c = 0.8$



$N = 1000, m = 800, c = 0.8$



SCM eigenvalues support for **white** Gaussian noise ($\sigma^2 = 1, \mathbf{C} = \sigma^2 \mathbf{I}_m$) and sources.

Source Detection with RMT

We consider N observations $\left\{ \mathbf{y}_k = \sqrt{\theta} \mathbf{u} + \mathbf{n}_k \right\}_{k \in [1, N]}$ with $\|\mathbf{u}\| = 1$. If the power θ of the source is **large enough**, then the limit of $\lambda_{\max} \left(\frac{1}{N} \mathbf{Y} \mathbf{Y}^H \right)$ is strictly larger than the right edge of the bulk.

■ if $\theta \leq \sigma^2 \sqrt{c}$, then

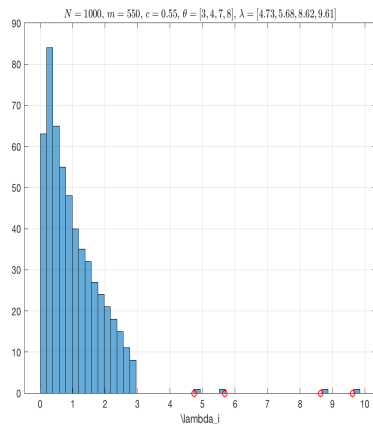
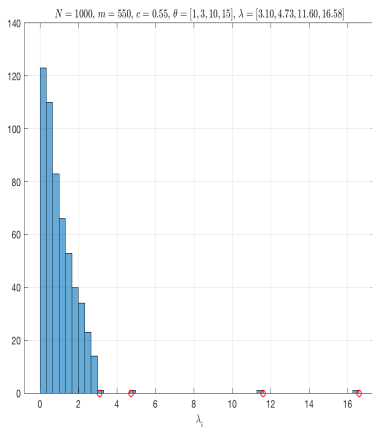
$$\lambda_{\max} \left(\frac{1}{N} \mathbf{Y} \mathbf{Y}^H \right) \xrightarrow[N, m \rightarrow \infty]{a.s.} \sigma^2 (1 + \sqrt{c})^2 ,$$

■ if $\theta \geq \sigma^2 \sqrt{c}$, then

$$\lambda_{\max} \left(\frac{1}{N} \mathbf{Y} \mathbf{Y}^H \right) \xrightarrow[N, m \rightarrow \infty]{a.s.} \sigma^2 (1 + \theta) \left(1 + \frac{c}{\theta} \right) \geq \sigma^2 (1 + \sqrt{c})^2 .$$

Above the threshold $\sigma^2 \sqrt{c}$, $\lambda_{\max} \left(\frac{1}{N} \mathbf{Y} \mathbf{Y}^H \right)$ asymptotically separates from the bulk.

Source Detection with RMT



Radar Detection Schemes for Joint Time and Spatial Correlated Clutter

Motivations: Adaptive radar detection and estimation schemes are often based on **the independence** of the secondary data used for building estimators and detectors. This independence allows to build Likelihood functions.

Example: estimating a covariance matrix \mathbf{M}

With a given set of n independent m -dimensional vectors $\{\mathbf{y}_i\}_{i \in [1, n]}$ distributed according to $\mathcal{CN}(\mathbf{0}_m, \mathbf{M})$, the corresponding Likelihood function Λ can be built as

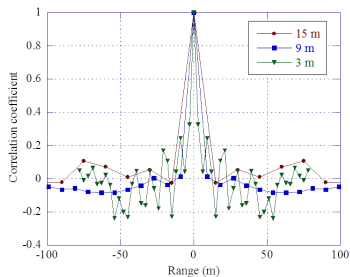
$$\Lambda(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n | \mathbf{M}) = \prod_{i=1}^n p(\mathbf{y}_i) = \prod_{i=1}^n \frac{1}{\pi^m |\mathbf{M}|} \exp(-\mathbf{y}_i^H \mathbf{M}^{-1} \mathbf{y}_i).$$

The Maximum Likelihood Estimate $\hat{\mathbf{M}}$ of \mathbf{M} is the zero of the partial derivative of $\Lambda(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n | \mathbf{M})$ with respect to \mathbf{M} leading to the well known SCM.

Motivations

In many radar and imagery applications, data $\{\mathbf{y}_i\}_{i \in [1, n]}$ can be viewed as a joint spatial and temporal process:

- For high-resolution radar, the sea clutter is jointly spatially and temporally correlated,



Sea clutter spatial correlation, IPIX radar [Greco].

Motivations

- In multichannel (polarimetric, interferometric or multi-temporal) SAR imaging, the multivariate vector characterizing each spatial pixel of the image is correlated over the channels. Still, it can also be strongly correlated with those of neighborhood pixels,
- When a radar signal with bandwidth B is oversampled ($F_e = k B$, $k > 1$), the associated range bins can be spatially correlated and the measurements are not independent anymore.

In the radar community, one generally supposes that the vectors of information collected over spatial support are **identically and independently distributed**.

This problem could be, for example, addressed using Multidimensional Space-time ARMA modeling.

This work aims to relax this hypothesis through the use of recent Random Matrix Theory results.

Problem formulation

Detection of a complex signal corrupted by an additive Gaussian noise $\mathbf{c} \sim \mathcal{CN}(\mathbf{0}_m, \mathbf{M})$ in a N -dimensional complex observation vector \mathbf{y} :

$$\begin{cases} H_0 : \mathbf{y} = \mathbf{c} & \mathbf{y}_i = \mathbf{c}_i \quad i = 1, \dots, n \\ H_1 : \mathbf{y} = \alpha \mathbf{p} + \mathbf{c} & \mathbf{y}_i = \mathbf{c}_i \quad i = 1, \dots, n \end{cases},$$

where \mathbf{p} is a perfectly known complex steering vector, α is the unknown signal amplitude and where the $\mathbf{c}_i \sim \mathcal{CN}(\mathbf{0}_m, \mathbf{M})$ are n signal-free non independent measurements. The covariance matrix \mathbf{M} characterizes the temporal or spectral correlation within the components of the noise vectors.

To model the spatial dependency between the secondary data, from the Gaussian assumption on \mathbf{c}_i , we may write the $m \times n$ -matrix $\mathbf{C} = [\mathbf{c}_1, \dots, \mathbf{c}_n]$ under the following form:

$$\mathbf{C} = \mathbf{M}^{1/2} \mathbf{X} \mathbf{T}^{1/2},$$

where $\mathbf{M} \in \mathbb{C}^{m \times m}$ and $\mathbf{T} \in \mathbb{C}^{m \times n}$ are both nonnegative definite, \mathbf{X} is standard Gaussian $\mathcal{CN}(\mathbf{0}_m, \mathbf{I}_m)$, and where \mathbf{T} satisfies the normalization $\frac{1}{n} \text{tr}(\mathbf{T}) = 1$.

Problem formulation

The matrix \mathbf{T} is considered Toeplitz, i.e., for all i, j , $\mathbf{T}_{i,j} = t_{|i-j|}$ for $t_0 = 1$ and $t_k \in \mathbb{C}$, and positive definite. Besides, $\sum_{k=0}^{n-1} |t_k| < \infty$.

Example: $m = 2$, $n = 3$

$$\mathbf{C} = \underbrace{\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}}_{\text{Temporal correlation}}^{1/2} \underbrace{\begin{pmatrix} x_{1,1} & x_{1,2} & x_{1,3} \\ x_{2,1} & x_{2,2} & x_{2,3} \end{pmatrix}}_{\text{Temporal or Spectral Measurements}} \underbrace{\begin{pmatrix} t_0 & t_1 & t_2 \\ t_1 & t_0 & t_1 \\ t_2 & t_1 & t_0 \end{pmatrix}}_{\text{Spatial correlation}}^{1/2}.$$

Some RMT results

Proposition: Consistent Estimation for \mathbf{T} [Couillet, 15]

As $m, n \rightarrow \infty$ such that $m/n \rightarrow c \in [0, \infty[$, and for every $\beta < 1$,

$$m^\beta \left\| \mathcal{T} \left[\frac{1}{m} \mathbf{C}^H \mathbf{C} \right] - \left(\frac{1}{m} \text{tr } \mathbf{M} \right) \mathbf{T} \right\|_F \xrightarrow{\text{a.s.}} 0,$$

where $\mathcal{T}[\cdot]$ is the Toeplitzification operator: $(\mathcal{T}[\mathbf{X}])_{ij} = \frac{1}{n} \sum_{k=1}^n \mathbf{X}_{k, k+|i-j|}$.

Up to a constant, a consistent estimator $\hat{\mathbf{T}}$ of the spatial covariance \mathbf{T} characterizing data $\{\mathbf{c}_i\}_{i \in [1, n]}$ is therefore defined as $\hat{\mathbf{T}} \propto \mathcal{T} \left[\frac{1}{m} \mathbf{C}^H \mathbf{C} \right]$ and the associated time whitened sample covariance matrix estimate $\hat{\mathbf{M}}$ of \mathbf{M} is defined as $\hat{\mathbf{M}} \propto \frac{1}{n} \mathbf{C} \hat{\mathbf{T}}^{-1} \mathbf{C}^H$.

This technique has been extended in the framework of robust M -estimators.

Gaussian and non-Gaussian scenarios

Simulated Data: joint spatial and time correlated Gaussian or K-distributed ($\nu = 0.5$) data characterized by $m = 10$ pulses, $n = 20$ secondary data where:

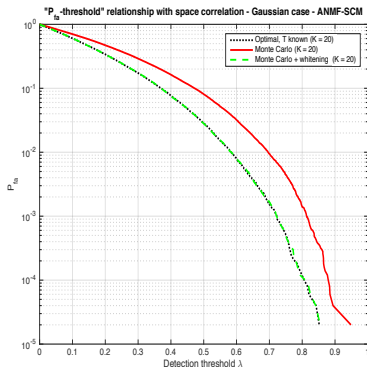
$$\mathbf{M} = \left(\rho_{\mathbf{M}}^{|i-j|} \right)_{i,j \in [1,m]}, \mathbf{T} = \left(\rho_{\mathbf{T}}^{|i-j|} \right)_{i,j \in [1,n]} \text{ with } \rho_{\mathbf{M}} = 0.5, \rho_{\mathbf{T}} = 0.9.$$

To evaluate the detection performance of the Λ_{ANMF} test statistic, we have compared three approaches:

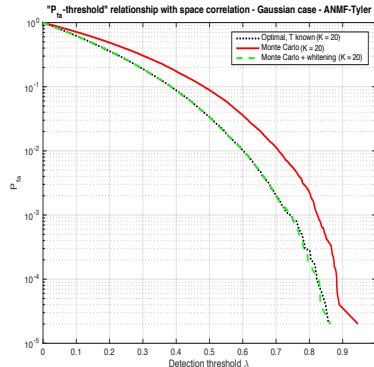
- \mathbf{M} is unknown but \mathbf{T} is assumed to be known: the covariance estimate $\hat{\mathbf{M}}$ is either given by $\frac{1}{n} \mathbf{C} \mathbf{T}^{-1} \mathbf{C}^H$ (SCM) or the Tyler's estimate of the true spatial-whitened data $\mathbf{C} \mathbf{T}^{-1/2}$,
- \mathbf{T} is assumed to be unknown and is estimated through $\hat{\mathbf{T}} \propto \mathcal{T} \left[\frac{1}{m} \mathbf{C}^H \mathbf{C} \right]$: the covariance estimate $\hat{\mathbf{M}}$ is either given by $\frac{1}{n} \mathbf{C} \hat{\mathbf{T}}^{-1} \mathbf{C}^H$ (SCM) or the Tyler's estimate of the spatial-whitened data $\mathbf{C} \hat{\mathbf{T}}^{-1/2}$,
- the classical approach that does not take into account the space correlation: the covariance estimate $\hat{\mathbf{M}}$ is either given by $\frac{1}{n} \mathbf{C} \mathbf{C}^H$ (SCM) or Tyler's estimate of the data \mathbf{C} .

False Alarm Regulation - Gaussian Case

ANMF-SCM



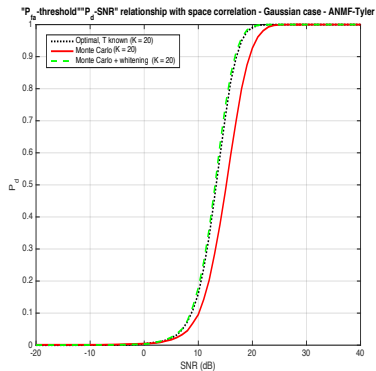
ANMF-Tyler



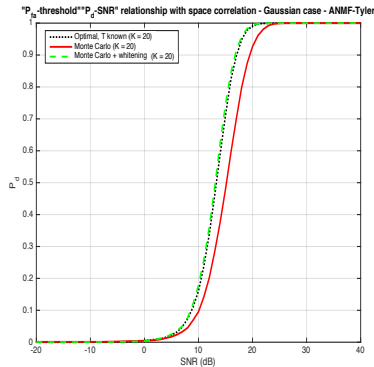
Same False Alarm Regulation performance for ANMF-SCM and ANMF-Tyler (Gaussian case)

Associated Detection Performance - Gaussian Case

ANMF-SCM



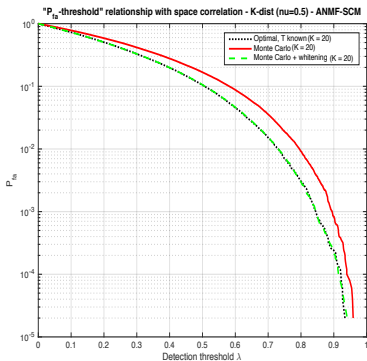
ANMF-Tyler



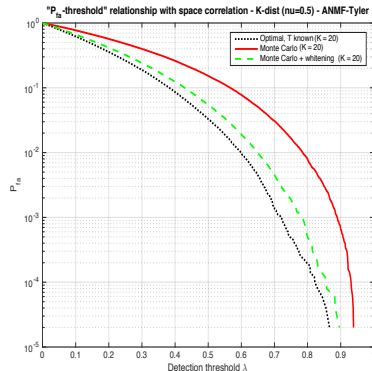
- Same Probability of Detection performance.
- Around 3dB gain improvement with RMT whitening procedure

False Alarm Regulation - K-distributed Case

ANMF-SCM



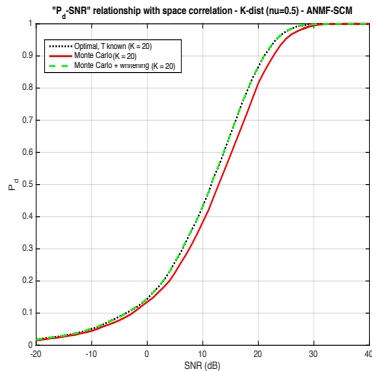
ANMF-Tyler



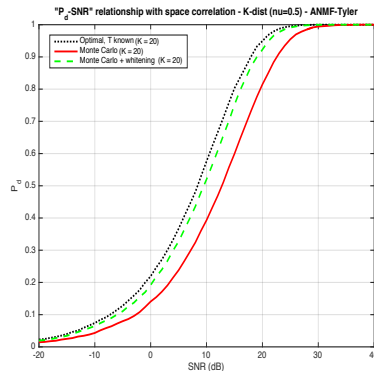
- Better False Alarm regulation performance for ANMF-FP (Non-Gaussian case).
- Better False Alarm regulation with RMT whitening procedure

Associated Detection Performance - K-distributed Case

ANMF-SCM



ANMF-Tyler



- Better performances in terms of Probability of Detection performance for ANMF-Tyler.
- Around 3dB gain improvement with RMT whitening procedure

Outline

1 Robust Estimation and Detection

- Going to Robust Adaptive Detection
- Modeling the Background
- Robust Estimation
- Robust Detection
- Robustness of M-estimators

2 Other Refinements

- Exploiting Prior Information: Covariance Structure
- Low Rank Detectors
- Shrinkage of M -estimator
- Riemannian manifolds
- RMT Theory and M -Estimator based Detectors
 - RMT key ideas
 - Radar Detection Schemes for Joint Time and Spatial Correlated Clutter

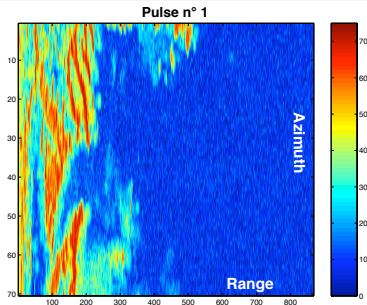
3 Applications and Results in Radar, STAP, SAR imaging, Hyperspectral Imaging

- **Surveillance Radar against Ground and Sea Clutter**
- Detection Performance on STAP Data
- Detection Performance on SAR Image
- Hyperspectral Imaging: Detection and Anomaly Detection

False Alarm Regulation on THALES Ground Clutter

Data Description

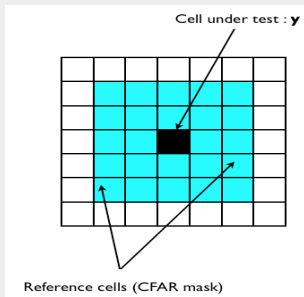
- "Range-azimuth" map from ground clutter data collected by radar from THALES Air Defense, placed 13 meters above ground and illuminating area at a low grazing angle.
- Ground clutter complex echoes collected in 868 range bins for 70 different azimuth angles and for $m = 8$ pulses.



False Alarm Regulation on THALES Ground Clutter

Data processing

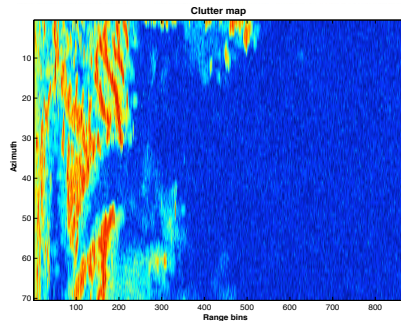
- Rectangular CFAR mask 5×5 for $0 \leq k \leq m$ different steering vectors \mathbf{p}_k .



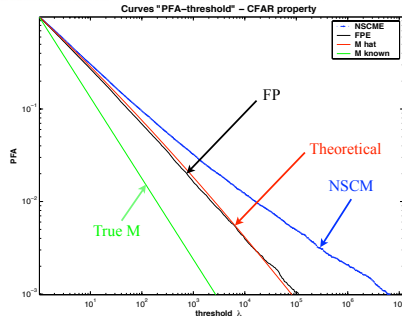
$$\mathbf{p}_k = \begin{pmatrix} 1 \\ \exp\left(\frac{2i\pi(k-1)}{m}\right) \\ \exp\left(\frac{2i\pi(k-1)2}{m}\right) \\ \vdots \\ \exp\left(\frac{2i\pi(k-1)(m-1)}{m}\right) \end{pmatrix}$$

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

False Alarm Regulation on THALES Ground Clutter



Azimuth/range bins map

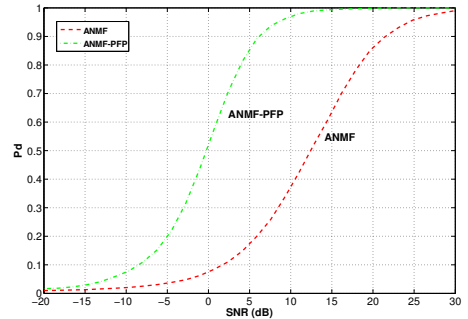
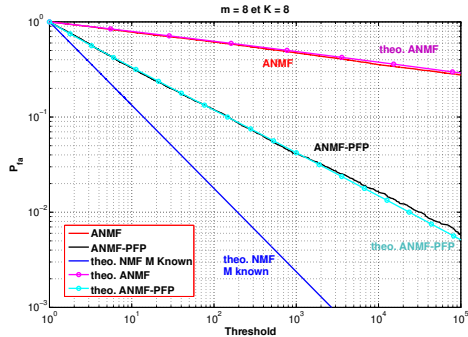


Relationship " P_{fa} -threshold"

Figure: ANMF with Tyler's M-estimate - False alarm regulation for $p_0 = (1 \dots 1)^T$.

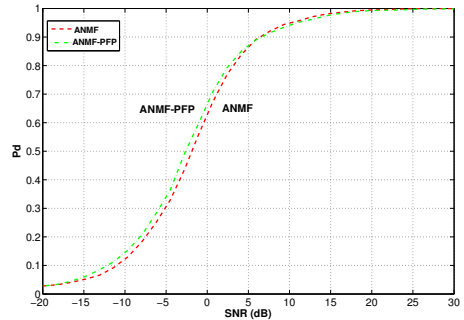
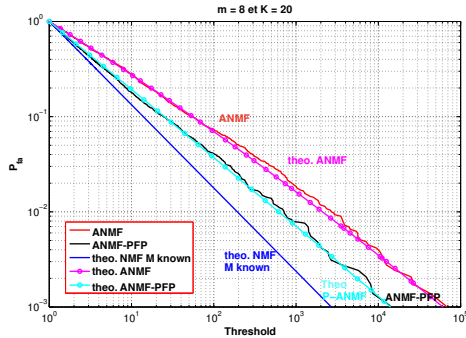
Black curve fits red curve until $PFA = 10^{-3}$ [Ovarlez et al. 16].

False Alarm Regulation on THALES Ground Clutter



Persymmetric Tyler-ANMF and Tyler ANMF on THALES dataset - $m = 8$, $n = 8$

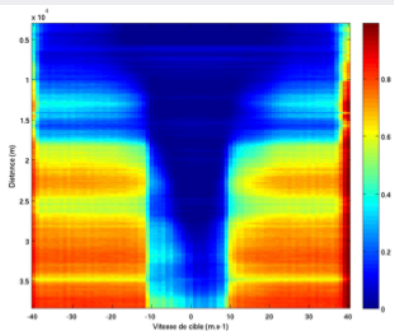
False Alarm Regulation on THALES Ground Clutter



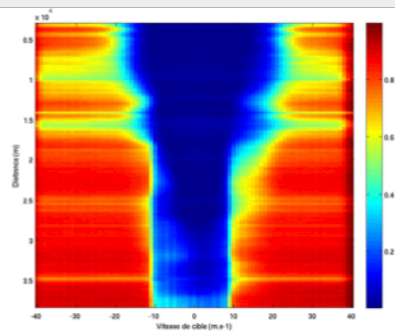
Persymmetric Tyler-ANMF and Tyler ANMF on THALES dataset - $m = 8$, $n = 20$

Detection Performance on THALES Sea Clutter

Non-Stationary and Heterogeneous THALES Sea clutter



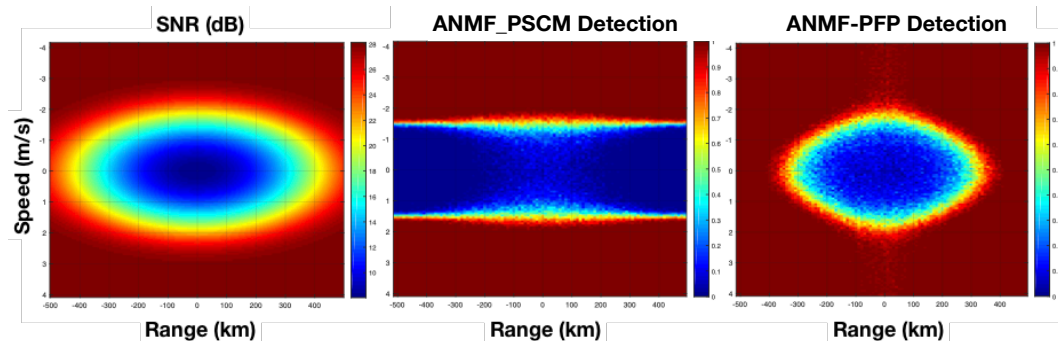
X Detector on Dieppe sea clutter



ANMF-FP Detector on Dieppe sea clutter

Detection in very Heterogeneous Environment

Simulation: Spatially and Spectrally Heterogeneous Strong Clutter

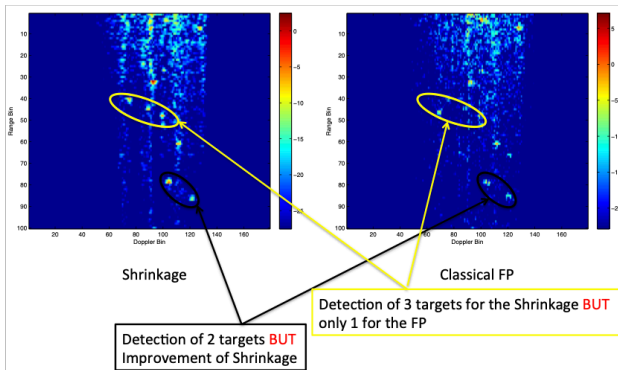


Application of Shrinkage to HFSWR: Detection

Singapore NTU HFSWR dataset

Shrinkage FP Estimator

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



Outline

1 Robust Estimation and Detection

- Going to Robust Adaptive Detection
- Modeling the Background
- Robust Estimation
- Robust Detection
- Robustness of M-estimators

2 Other Refinements

- Exploiting Prior Information: Covariance Structure
- Low Rank Detectors
- Shrinkage of M -estimator
- Riemannian manifolds
- RMT Theory and M -Estimator based Detectors
 - RMT key ideas
 - Radar Detection Schemes for Joint Time and Spatial Correlated Clutter

3 Applications and Results in Radar, STAP, SAR imaging, Hyperspectral Imaging

- Surveillance Radar against Ground and Sea Clutter
- **Detection Performance on STAP Data**
- Detection Performance on SAR Image
- Hyperspectral Imaging: Detection and Anomaly Detection

Detection Performance on STAP Data

Problem

Using joint spatial and time measurements, estimate the position (angle) and the Doppler frequency (speed) of the target

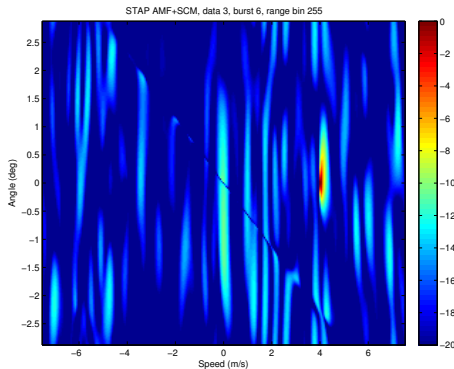
⇒ use of the ANMF with a particular steering vector [Ovarlez 2011]

Data parameters: experimental clutter with synthetic target

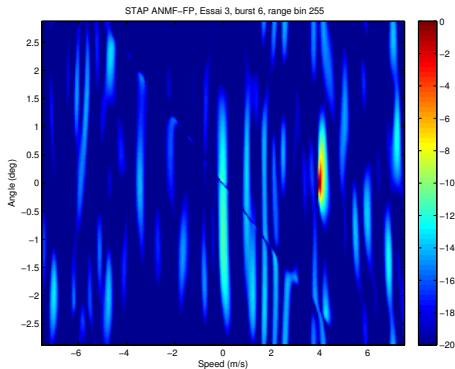
X-Band $\simeq 10^9$ Hz, wavelength $\lambda = 0.03\text{m}$, flight speed $v = 100\text{m/s}$, distance to the scene 30km, 5 deg of incidence, PRF (Pulse Repetition Frequency) of 1 kHz, inter-sensor distance $d = 0.3\text{m}$, 12 trials with $n = 410$ range bins, $M = 64$ pulses and $N = 4$ sensors.

- This means observations of size $m = N M = 256$ 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



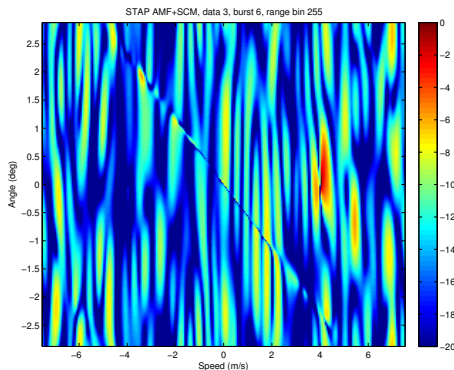
AMF detector with the SCM



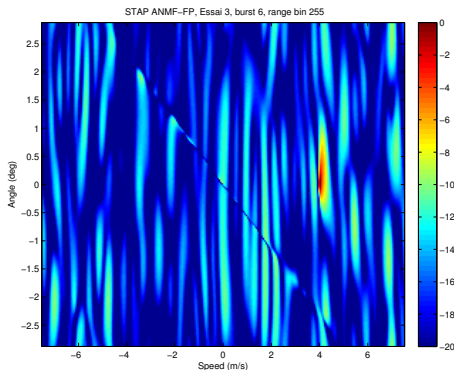
ANMF detector with Tyler's estimate.

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

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



AMF detector with the SCM



ANMF detector with Tyler's estimate.

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

Detection Performance on STAP Data

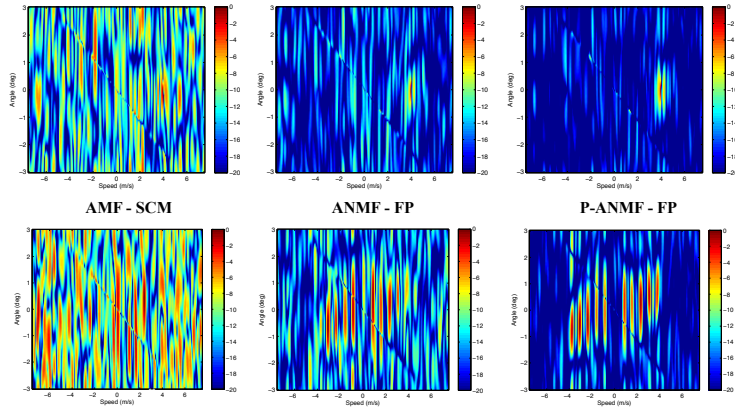


Figure: Doppler-angle map for the range bin 255 with $n = 404$ secondary data, $m = 256$ [Pailloux 10].

Extended Low Rank Detectors [*Ginolhac 11, 12 and 13*]

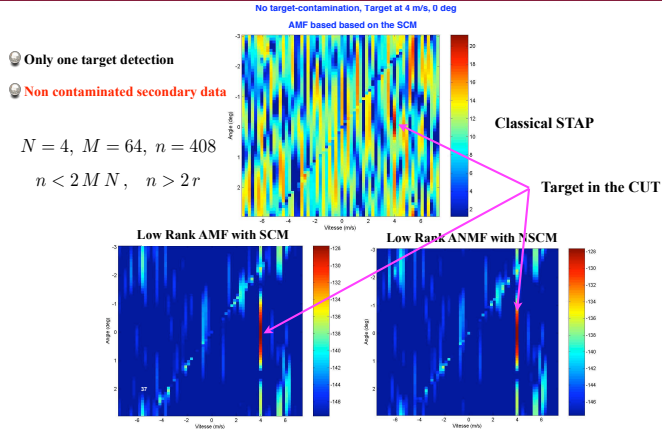


Figure: Doppler-angle map for the range bin 255 with $n = 100 < m$ secondary data and $m = 256$.
(guard cells are removed)

Extended Low Rank Detectors [*Ginolhac 11, 12 and 13*]

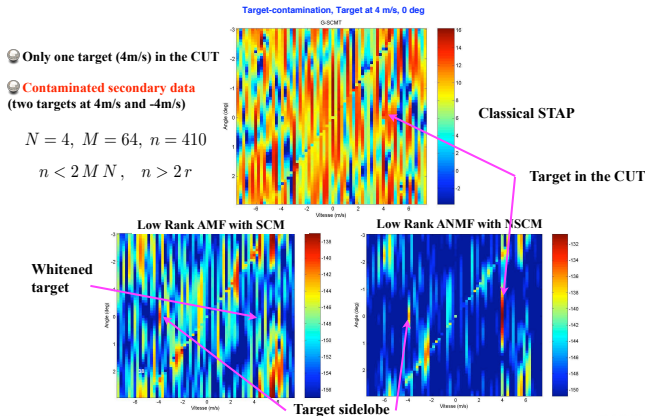


Figure: Doppler-angle map for the range bin 255 with $n = 100 < m$ secondary data and $m = 256$. (guard cells are removed)

Outline

1 Robust Estimation and Detection

- Going to Robust Adaptive Detection
- Modeling the Background
- Robust Estimation
- Robust Detection
- Robustness of M-estimators

2 Other Refinements

- Exploiting Prior Information: Covariance Structure
- Low Rank Detectors
- Shrinkage of M -estimator
- Riemannian manifolds
- RMT Theory and M -Estimator based Detectors
 - RMT key ideas
 - Radar Detection Schemes for Joint Time and Spatial Correlated Clutter

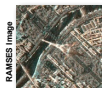
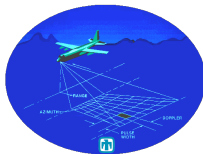
3 Applications and Results in Radar, STAP, SAR imaging, Hyperspectral Imaging

- Surveillance Radar against Ground and Sea Clutter
- Detection Performance on STAP Data
- **Detection Performance on SAR Image**
- Hyperspectral Imaging: Detection and Anomaly Detection

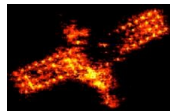
Background on SAR and Radar Imaging



ONERA RAMSES Image



RAMSES Image



ONERA ISAR Image



ONERA RAMSES Image

Radar Imaging allows to build more and more precise images:

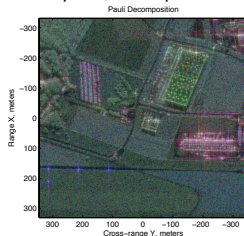
- Current use of **very high spectral bandwidth** and **very high angular bandwidth** leading to very high spatial resolution,
- Application to monitoring (detection, change detection), classification, 3D reconstruction, EM analysis, etc.

These applications require some physical diversity to achieve good performance.

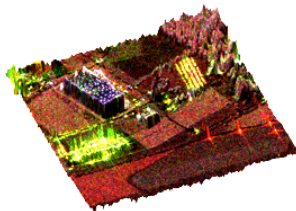
Multi-Channel SAR Images

Multi-channel SAR images automatically propose this diversity through:

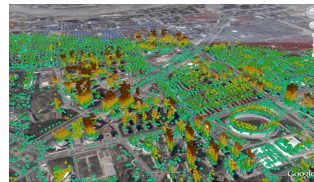
- polarimetric channels (POLSAR), interferometric channels (INSAR), polarimetric and interferometric channels (POLINSAR),
- multi-temporal, multi-passes SAR Image, etc.



EM behavior of the terrain
in POLSAR images



Estimation of the height
in POLINSAR images



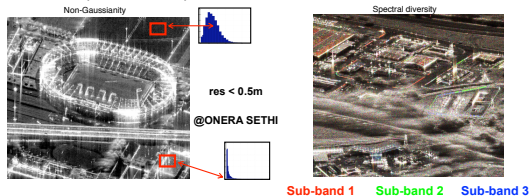
Analysis of the structures displacement in
Shanghai with multi-temporal SAR images
(@Telespazio)

Almost all the conventional techniques of detection, parameters estimation, speckle filtering techniques, and classification in multi-channel SAR images (e.g., polarimetric covariance matrix, interferometric coherency matrix) are based on the **multivariate statistic**.

Mono-Channel SAR Images

For **mono-channel SAR Images**, each pixel of the spatial image is **only** characterized by a complex amplitude, and we don't have direct access to this diversity. Moreover,

- very high-resolution SAR images are more and more complex, detailed, heterogeneous,
- the spatial statistic of SAR images may be **not at all Gaussian** !
- SAR pixels may be **dispersive** (or colored) and **anisotropic**.



Challenging Problems

- How to retrieve how to exploit this diversity (dispersive and anisotropic information) from mono-channel SAR image?
- How to derive Multivariate Adaptive Detectors on a mono-channel complex SAR image?

Conventional Principle of Radar/SAR Imaging

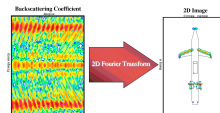
Conventional Fourier Imaging (laboratory, SAR, ISAR):

- Assumptions of white and isotropic bright points
- It does not exploit the potential non-stationarities or diversities of the scatterers
- Hypothesis of bright points modeling: all the scatterers localized in \mathbf{x} and characterized by the complex spatial amplitude distribution $I(\mathbf{x})$ have **the same behavior** for any wave vector $\mathbf{k} = \frac{2f}{c} (\cos \theta, \sin \theta)^T$. After some processing, the backscattering coefficient $H(\mathbf{k})$ acquired by the radar is related to the SAR image $I(\mathbf{x})$ through:

$$H(\mathbf{k}) = \int_{\mathcal{D}_{\mathbf{x}}} I(\mathbf{x}) \exp(-2i\pi \mathbf{k}^T \mathbf{x}) d\mathbf{x}$$

- The SAR image $I(\mathbf{x})$ is then obtained through the Inverse Fourier Transform:

$$I(\mathbf{x}) = \int_{\mathcal{D}_{\mathbf{k}}} H(\mathbf{k}) \exp(2i\pi \mathbf{k}^T \mathbf{x}) d\mathbf{k}$$



This model loses all information relative to frequency f and angle θ . Hence, spectral and angular diversities are lost.

Time-Frequency Distributions for SAR Imaging - Key Idea

Time-frequency distributions are generally devoted to non-stationary time signals analysis (e.g., spectral components varying with time). They can be easily extended in 2D.

Key idea: In the context of SAR Imaging, Time-Frequency Analysis allows:

- to highlight the coloration and anisotropy properties of monodimensional SAR scatterers,
- to characterize each pixel of the complex SAR image with a vector of information related to angular or/and frequency behaviors.

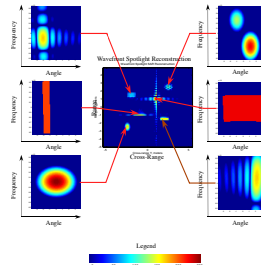
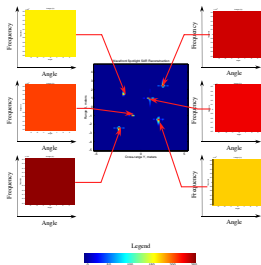
LTFD analysis and the physical group theory (Heisenberg or affine group) allow us to construct **hyperimages** [Bertrand 91, Bertrand 94, Bertrand 96] through:

$$\tilde{I}(\mathbf{r}_0, \mathbf{k}_0) = \langle H(\cdot), \Psi_{\mathbf{r}_0, \mathbf{k}_0}(\cdot) \rangle = \int_{\mathcal{D}_{\mathbf{k}}} H(\mathbf{k}) \Psi_{\mathbf{r}_0, \mathbf{k}_0}^*(\mathbf{k}) d\mathbf{k},$$

where $\Psi_{\mathbf{r}_0, \mathbf{k}_0}(\mathbf{k})$ is a family of wavelet bases (Gabor, wavelet) generated from a mother wavelet $\phi(f, \theta)$ through the chosen physical group of transformation (translations, scale in frequency, etc.) and where $\mathcal{D}_{\mathbf{k}}$ is the spectral/angular support of the wavelet Ψ .

Highlighting the Spectral and Angular Behaviors of Scatterers

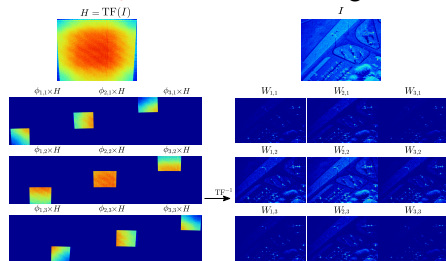
Some examples of synthetic hyperimages $\tilde{I}(\mathbf{r}_0, \mathbf{k}_0)$:



- An isotropic and white scatterer is mainly located on a pixel of SAR image,
- An anisotropic and colored scatterer may naturally spread out in the spatial domain.

From Mono-Channel to Multi-Channel SAR Image

Example of $N_f = 3$ sub-bands and $N_\theta = 3$ sub-looks image decomposition:



Exploitation of the diversity

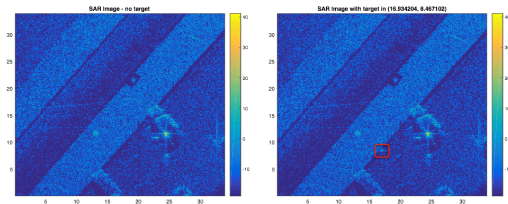
Each pixel i of the mono-channel SAR image can now be characterized by a N -vector $\mathbf{x}_i = [W_{1,1}^i, \dots, W_{N_f, N_\theta}^i]^T$ of information ($N = N_f N_\theta$) related to **dispersion** in frequency domain and **anisotropy** in angular domain. Which multivariate statistic can characterize the vector \mathbf{x}_i ?

Detection Performance on SAR Image

Analysis of Performance

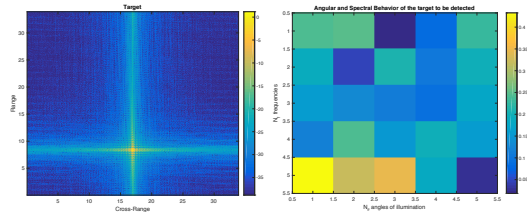
- Evaluation of the CFAR property of the AMF and ANMF detectors,
- Comparison of the target detection performance between AMF and ANMF.

Dataset from SANDIA National Laboratories



Left: Original SAR Image without target. Right: SAR image with the specific embedded target.

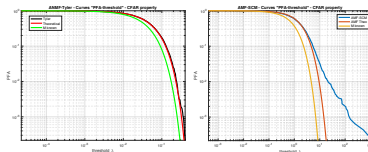
Artificial embedded target



Left: SAR Image of the target. Right: True target response \mathbf{p} in angular and spectral spaces ($N_\theta = 5$ sub-looks, $N_f = 5$ sub-bands).

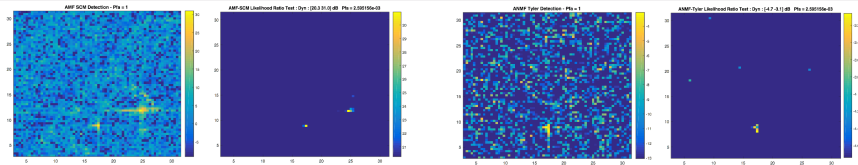
Detection Performance on SAR Image

Perfect PFA regulation with ANMF-Tyler but poor PFA regulation for AMF-SCM



Left: FA Regulation with ANMF-Tyler. Right: FA Regulation with AMF-SCM. $N_0 = 5$, $N_f = 5$, $K = 88$.

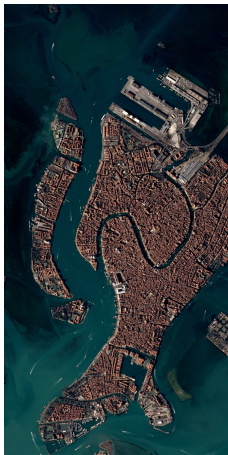
Better target detection for ANMF-Tyler [Ovarlez 17, Mian 19]



Left: Full AMF-SCM detection test, $P_{fa} = 1$. Right: AMF-SCM detection test, $P_{fa} = 2.6 \cdot 10^{-3}$.

Left: ANMF-Tyler detection test, $P_{fa} = 1$. Right: ANMF-Tyler detection test, $P_{fa} = 2.6 \cdot 10^{-3}$.

Change Detection on SAR Image Time Series



Contributions to SAR Image Time Series Analysis

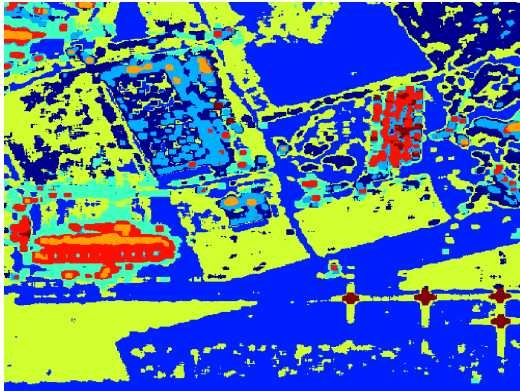
Guillaume Ginolhac and Arnaud Breloy

RadarConf Tutorial - 26 September 2020

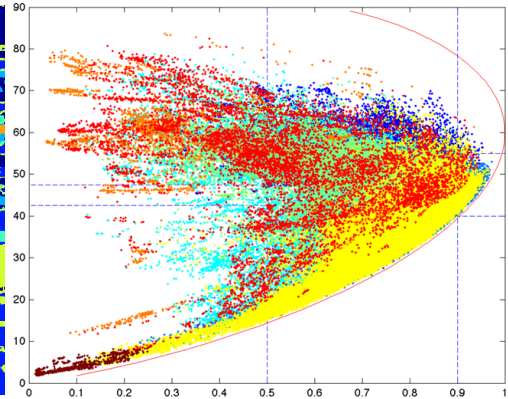
Joint work with: Ammar Mian (Univ. of Aalto),
Jean-Philippe Ovarlez (ONERA & SONDR),
and Abdourrahmane M. Atto (LISTIC)



Example of Polarimetric Repartition in the $H - \alpha$ plane



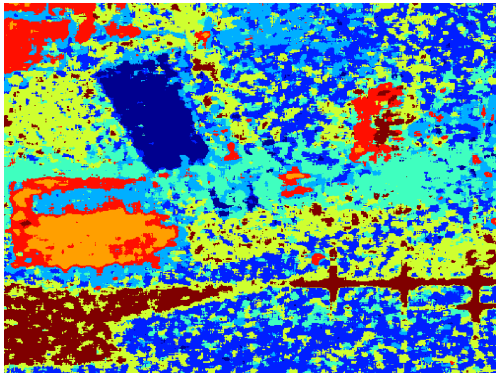
(a) Classification results



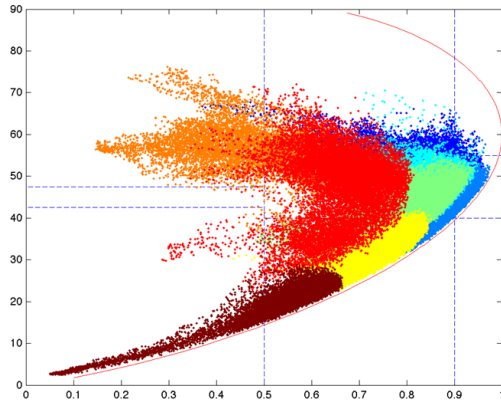
(b) Repartition

Figure: SCM, Euclidean mean, Wishart distance

Example of Polarimetric Repartition in the $H - \alpha$ plane



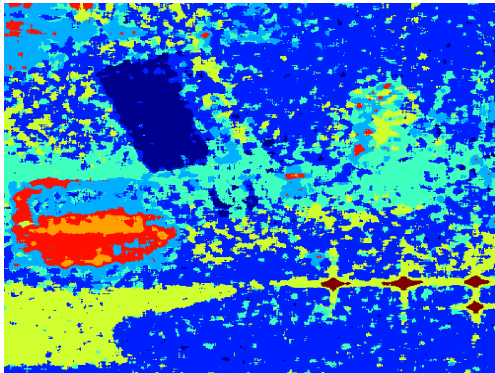
(a) Classification results



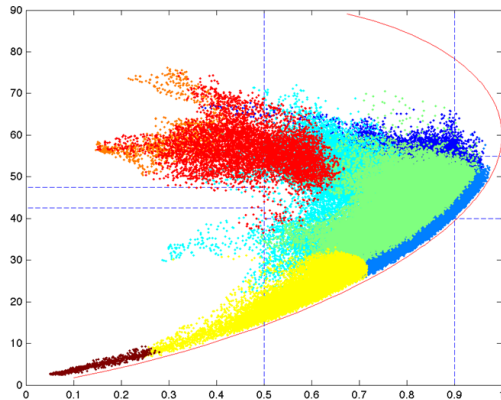
(b) Repartition

Figure: Tyler, Euclidean mean, Wishart distance

Example of Polarimetric Repartition in the $H - \alpha$ plane



(a) Classification results



(b) Repartition

Figure: Tyler, Riemannian mean, Wishart distance

Outline

1 Robust Estimation and Detection

- Going to Robust Adaptive Detection
- Modeling the Background
- Robust Estimation
- Robust Detection
- Robustness of M-estimators

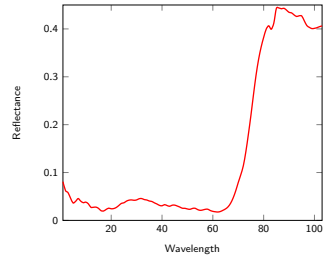
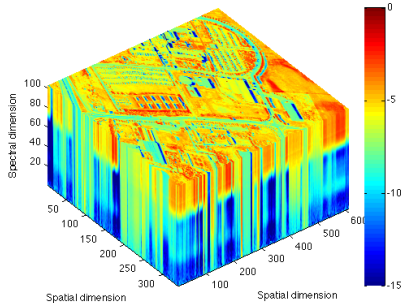
2 Other Refinements

- Exploiting Prior Information: Covariance Structure
- Low Rank Detectors
- Shrinkage of M -estimator
- Riemannian manifolds
- RMT Theory and M -Estimator based Detectors
 - RMT key ideas
 - Radar Detection Schemes for Joint Time and Spatial Correlated Clutter

3 Applications and Results in Radar, STAP, SAR imaging, Hyperspectral Imaging

- Surveillance Radar against Ground and Sea Clutter
- Detection Performance on STAP Data
- Detection Performance on SAR Image
- Hyperspectral Imaging: Detection and Anomaly Detection

Hyperspectral Imaging



- **Anomaly Detection**

To detect all that is "different" from the background (Mahalanobis distance) -
No information about the targets of interest available [Frontera 16].

- **"Pure" Detection**

To detect targets characterized by a given spectral signature \mathbf{p} - Regulation of False Alarm [Ovarlez 11, Frontera 17].

Problem: the statistical mean is non null \Rightarrow M -estimator of the mean is required

$$\hat{\mu} = \frac{\sum_{i=1}^n u_1(t_i) \mathbf{z}_i}{\sum_{i=1}^n u_1(t_i)} \text{ and } \hat{\Sigma} = \frac{1}{n} \sum_{i=1}^n u_2(t_i^2) (\mathbf{z}_i - \hat{\mu}) (\mathbf{z}_i - \hat{\mu})^H,$$

where $t_i = \left((\mathbf{z}_i - \hat{\mu})^H \hat{\Sigma}^{-1} (\mathbf{z}_i - \hat{\mu}) \right)^{1/2}$ and $u_1(\cdot), u_2(\cdot)$ denote any real-valued *weight functions* (following the conditions of Maronna).

Joint estimation of location and scale [Bilodeau 08]

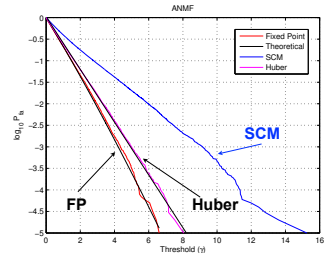
Hyperspectral Imaging

ANMF and M-estimates for Hyperspectral target detection [Frontera 14]

$$\Lambda(\mathbf{c}) = \frac{\left| \mathbf{p}^H \hat{\Sigma}^{-1} (\mathbf{c} - \hat{\mu}) \right|^2}{(\mathbf{p}^H \hat{\Sigma}^{-1} \mathbf{p}) \left((\mathbf{c} - \hat{\mu})^H \hat{\Sigma}^{-1} (\mathbf{c} - \hat{\mu}) \right)} \underset{H_0}{\overset{H_1}{\approx}} \lambda$$

$$P_{fa} = (1 - \lambda)^{\frac{n-1}{\sigma_1} - m + 1} {}_2F_1 \left(\frac{n-1}{\sigma_1} - m + 2, \frac{n-1}{\sigma_1} - m + 1; \frac{n-1}{\sigma_1} - 1; \lambda \right), \text{ where } \sigma_1 = (m + 1)/m.$$

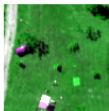
- This two-step GLRT test is homogeneous of degree 0: it is independent of any particular Elliptical distribution: CFAR texture and CFAR Matrix properties,
- Under homogeneous Gaussian region, it achieves the same performance as the detector built with the SCM estimate.



Hyperspectral Imaging



Original data set

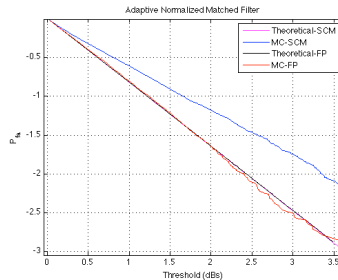
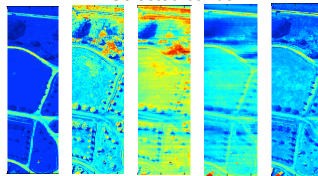


Extracted region :

- ▶ 100 x 100 pixels,
- ▶ 5 bands,
- ▶ Sliding Window: 19x19

Non-Gaussian region

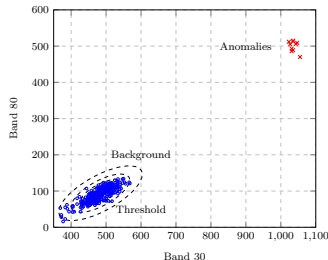
Selected bands



Hyperspectral Imaging

GLRT RX Anomaly Detector: Mahalanobis Distance [Reed 90]

Binary Hypotheses test: $\begin{cases} H_0 & : & \mathbf{c} = \mathbf{b} \\ H_1 & : & \mathbf{c} = A\mathbf{p} + \mathbf{b} \end{cases} \quad \mathbf{c}_1, \dots, \mathbf{c}_n \quad \text{where } \mathbf{b} \sim \mathcal{CN}(\mathbf{0}_m, \Sigma) \text{ and } \mathbf{c}_i \sim \mathcal{CN}(\mathbf{0}_m, \Sigma), A \text{ known and } \mathbf{p} \text{ unknown}$



$$\text{denoting } \hat{\boldsymbol{\mu}} = \frac{1}{n} \sum_{i=1}^n \mathbf{c}_i$$

$$RXD_{SCM}(\mathbf{c}) = (\mathbf{c} - \hat{\boldsymbol{\mu}})^H \hat{\mathbf{S}}_n^{-1} (\mathbf{c} - \hat{\boldsymbol{\mu}}) \underset{H_0}{\overset{H_1}{\gtrless}} \lambda$$

(Hotelling T^2 distributed)

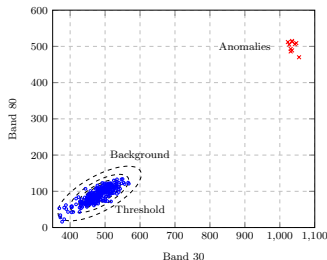
$$\frac{n-m}{m(n+1)} RXD_{SCM}(\mathbf{c}) \sim F_{m, n-m}$$

- Derived and valid only under Gaussian hypotheses,
- Its false alarm rate is independent of the covariance matrix: CFAR-matrix property in homogeneous Gaussian data.

Hyperspectral Imaging

Extended GLRT RX Anomaly Detector: Mahalanobis Distance [Frontera 14]

Binary Hypotheses test: $\begin{cases} H_0 &: \mathbf{c} = \mathbf{b} \\ H_1 &: \mathbf{c} = A\mathbf{p} + \mathbf{b} \end{cases} \quad \mathbf{c}_1, \dots, \mathbf{c}_n \quad \text{where } \mathbf{b} \sim CE(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_z) \text{ and } \mathbf{c}_i \sim CE(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_z), A$
known and \mathbf{p} unknown

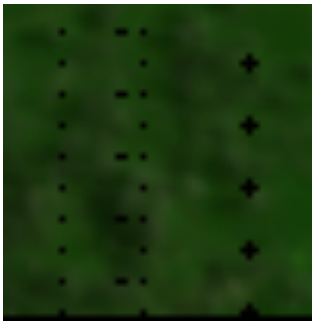


$$RXD_{M-est}(\mathbf{c}) = (\mathbf{c} - \hat{\boldsymbol{\mu}})^H \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{c} - \hat{\boldsymbol{\mu}}) \underset{H_0}{\overset{H_1}{\gtrless}} \lambda$$

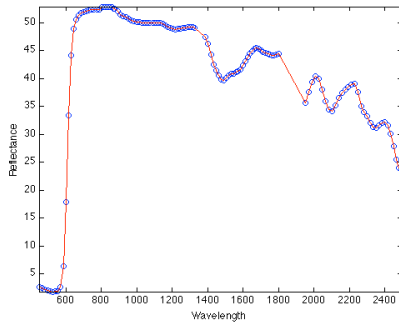
where $\hat{\boldsymbol{\Sigma}}$ and $\hat{\boldsymbol{\mu}}$ are M-estimates
of the location and scale

- Derived and valid for any Elliptical Contoured Distributions,
- Its false alarm rate depends on the texture statistic of the data.

Anomaly Detection Results on Artificial Targets



Original image (Forest Region)

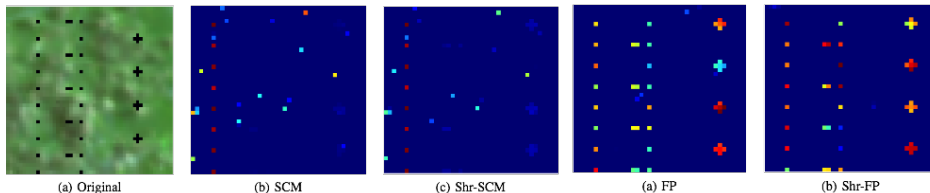


Target Spectrum

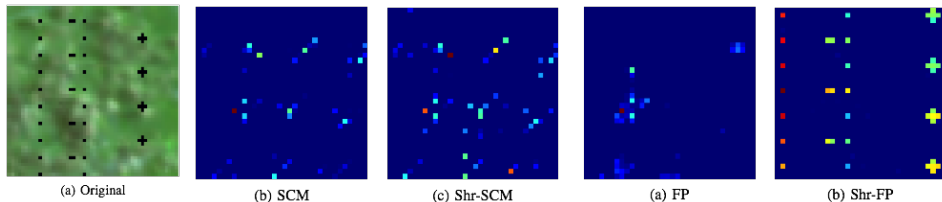
50 x 50 pixels, 126 spectral bands



Anomaly Detection Results on Artificial Targets



Extended Kelly AD built with conventional and robust estimates for artificial targets in real HSI ($m = 9$, $n = 80$, $PFA = 0.03$).

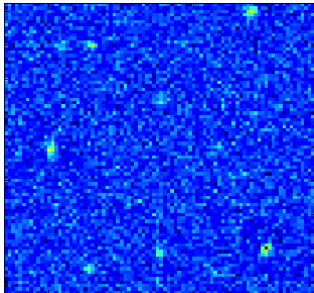


Extended Kelly AD built with conventional and robust estimates for artificial targets in real HSI ($m = 126$, $n = 288$, $PFA = 0.03$).

Galaxies Anomaly Detection Results on MUSE data

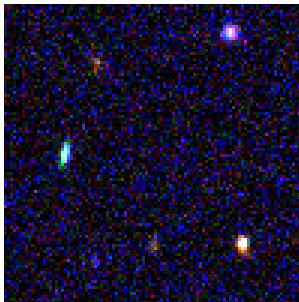
Problem of detecting galaxies in HS MUSE (Multi Unit Spectroscopic Explorer) data (465-930 nm)

Classical RXD



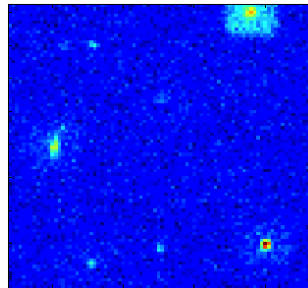
$RXD_{SCM}(c)$

Muse Image



300 x 300 pixels
3578 spectral bands

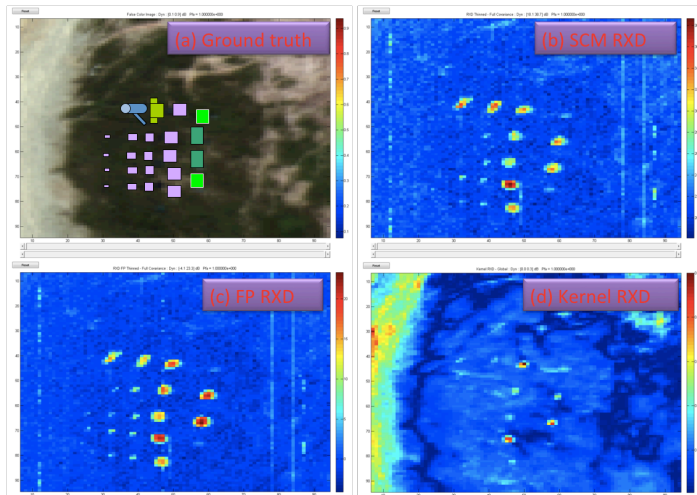
Extended RXD



$RXD_{Tyler}(c)$

Better detection and False Alarm regulation with Tyler estimate (same Pfa).

Experimental Data from DSO Singapore



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 us 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 excellent detection performance,
- Taking into account additional *a priori* properties on the covariance matrix structure (low rank, persymmetry, Toeplitz, etc.) can lead to an 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, Financial Portfolio Optimization, Hyperspectral Anomaly detection, Hyperspectral detection.

On-going works and Perspectives

- Link with **Random Matrix Theory**: for high dimensionality data (ex: hyperspectral, STAP), strong statistical connection with Robust Estimation theory [R. Couillet, and F. Pascal, J.-P. Ovarlez, E. Terreaux],
- Robust estimation of structured covariance matrices, Low-Rank covariance matrices [Y. Sun, D. P. Palomar, A. Breloy, G. Ginolhac, F. Pascal, J.-P. Ovarlez, C. Ren, P. Forster, B. Mériaux 2020]: persymmetric, Toeplitz, Bloc Toeplitz, Low-Rank matrices, etc.,
- 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],
- No secondary data: [C. Ren, N. El-Korso, P. Forster, A. Breloy, J.-P. Ovarlez, B. Mériaux],
- M-Estimators and Riemannian Geometry: [F. Barbaresco], [P. Formont, F. Pascal, G. Ginolhac, A. Renaux, A. Collas, J.-P. Ovarlez, F. Bouchard],
- Shrinkage of M-Estimators: [A. Wiesel, Y. Abramovitch, O. Besson, F. Pascal, E. Ollila, etc.], [Q. Hoarau, G. Ginolhac],
- Sparsity and high dimension: [A. Bitar, J.-P. Ovarlez].
- Performance of Estimation : [A. Renaux, B. Mériaux, S. Fortunati]
- Neural Network for improving classification: [A. Barrachina 22] for complex-valued Neural Networks, Spectral Clustering and Transfer Learning in high dimension [C. Doz], Neural Network for detection [A. Rouzoumka] and Anomaly Detection [M. Muzeau], Neural Network for SAR: Complex GAN for SAR [Q. Gabot].

Former ONERA/SONDRA Ph.D. students linked to this tutorial (1)

- **A. Rouzoumka**, Deep Learning Applied to the Robust Detection of Radar Targets, Ph.D. Thesis 2023-2026, Paris Saclay University,
- **Q. Gabot**, Generative Adversarial Networks for SAR Imaging, Ph.D. Thesis 2023-2026, Paris Saclay University,
- **M. Muzeau**, Anomaly Detection Schemes for SAR Imaging, Ph.D. Thesis 2021-2024, Paris Saclay University,
- **H. Brehier**, Detection and Classification for Radar Through The Wall from subspaces model, Ph.D. Thesis 2021-2024, Paris Saclay University,
- **O. Lerda**, Sonar Detection, Estimation and Classification for Targets in Complex Environnement, Ph.D. Thesis 2020-2024, Annecy University,
- **P. Develter**, New Radar Processing Robust to Mismatch Models: Case of off-grid targets, Ph.D. Thesis 2020-2023, Paris Saclay University,
- **C. Doz**, Spectral clustering Based Methods for Unsupervised Classification in Radar Imaging Applications, Ph.D. Thesis 2019-2023, Paris Saclay University,
- **J. A. Barrachina**, Complex Valued Neural Networks for Radar Applications, Ph.D. Thesis 2019-2022, Paris Saclay University,
- **A. Mian**, Exploitation of SAR and Hyperspectral Time Series Analysis, Ph.D. Thesis 2016-2019, Paris Saclay University,
- **B. Mériaux**, Contributions to robust signal processing for multi-sensor systems, Ph.D. Thesis 2017-2020, Paris Saclay University,
- **E. Terreaux**, Robust model order selection using Random Matrix Theory, Ph.D. Thesis 2015-2018, Paris Saclay University,
- **A. Bitar**, Exploitation of Sparsity for Hyperspectral Target Detection, Ph.D. Thesis 2015-2018, Paris Saclay University,

Former Ph.D. students linked to this tutorial (2)

- **U. H. Tan**, Colocated MIMO Radar Waveform Optimization, Ph.D. Thesis 2014-2017, Paris Saclay University,
- **A. Combernoux**, Low-rank detection and estimation using random matrix theory approaches for antenna array processing, Ph.D. Thesis 2013-2016, Paris Saclay University,
- **J. Frontera-Pons**, Robust Detection and Classification for Hyperspectral Imaging, Ph.D. Thesis 2011-2014, Paris Saclay University,
- **M. Mahot**, Robust Covariance Matrix Estimation in Signal Processing, Ph.D. Thesis 2009-2012, ENS Paris Saclay,
- **P. Formont**, Statistical and Geometric Tools for the Classification of Highly Textured Polarimetric SAR Images, Ph.D. Thesis 2009-2012, Paris Saclay University,
- **C. Y. Chong**, Signal Processing for MIMO Radars: Detection Under Gaussian and non-Gaussian Environments and Application to STAP, Ph.D. Thesis 2008-2011, Paris Saclay University,
- **G. Pailloux**, Noise Structured Covariance Estimation in Adaptive Detection, Ph.D. Thesis 2007-2009, Paris Saclay University,
- **M. Duquenoy**, Time-Frequency Analysis Applied to Polarimetric SAR Imaging, Ph.D. Thesis 2004-2007, Paris Saclay University,
- **F. Pascal**, Detection and Estimation in non-Gaussian Environment, Ph.D. Thesis 2003-2006, Nanterre University,
- **M. Tria**, Analysis of SAR Images using Continuous Multidimensional Wavelet Transform, Ph.D. Thesis 2001-2003, Paris Saclay University,
- **E. Jay**, Detection in non-Gaussian Environment, Ph.D. Thesis 1998-2001, Cergy-Pontoise University.

Bibliography (1)

- Y. I. Abramovich and N. K. Spencer. Diagonally loaded normalised sample matrix inversion (Insmi) for outlier-resistant adaptive filtering. *IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*, pp.1105-1108, 2007.
- Y. I. Abramovich and O. Besson. Regularized covariance matrix estimation in complex elliptically symmetric distributions using the expected likelihood approach - part 1: The over-sampled case. *Signal Processing, IEEE Transactions on*, **61**(23), pp.5807-5818, 2013.
- Y. I. Abramovich and O. Besson. Regularized covariance matrix estimation in complex elliptically symmetric distributions using the expected likelihood approach - part 2: The under-sampled case. *Signal Processing, IEEE Transactions on*, **61**(23), pp.5819-5829, 2013.
- T. J. Barnard and D. D. Weiner. Non-Gaussian clutter modeling with generalized spherically invariant random vectors. *Signal Processing, IEEE Transactions on*, **44**(10), pp.2384-2390, 1996.
- J. A. Barrachina, C. Ren, C. Morisseau, G. Vieillard, J.-P. Ovarlez, Comparison Between Equivalent Architectures of Complex-valued and Real-valued Neural Networks - Application on Polarimetric SAR Image Segmentation. *Journal of Signal Processing System*, **95**, pp.57-66, 2023.
- J. A. Barrachina, C. Ren, C. Morisseau, G. Vieillard, J.-P. Ovarlez. Theory and Implementation of Complex-Valued Neural Networks. arXiv:2302.08286v1, 2023.
- J. A. Barrachina, C. Ren, C. Morisseau, G. Vieillard, J.-P. Ovarlez, Impact of PolSAR pre-processing and balancing methods on complex-valued neural networks segmentation tasks. *IEEE Open Journal of Signal Processing*, **4**, pp.157-166, 2023.
- S. Bausson, F. Pascal, P. Forster, J.-P. Ovarlez and P. Larzabal, First and second order moments of the normalized sample covariance matrix of spherically invariant random vectors. *IEEE Signal Processing Letters*, **14**(6), pp.425-428, 2007,

Bibliography (2)

- J. Bertrand, P. Bertrand and J.-P. Ovarlez. Dimensionalized wavelet transform with application to radar imaging. *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'91)*, pp.2909-2912, Toronto, Canada, 1991.
- J. Bertrand, P. Bertrand and J.-P. Ovarlez. Frequency directivity scanning in laboratory radar imaging. *International Journal of Imaging Systems and Technology*, **5**(1), pp.39-51, 1994.
- J. Bertrand, P. Bertrand. The concept of hyperimage in wide-band radar imaging. *Geoscience and Remote Sensing, IEEE Transactions on*, **34**(5), pp.1144-1150, 1996.
- G. Bienvenu and L. Kopp. Principe de la goniométrie passive adaptative. In *Proc. du colloque GRETSI sur le traitement du signal et des images*, 1979.
- M. Bilodeau and D. Brenner. *Theory of multivariate statistics*. Springer Science & Business Media, 2008.
- A. Breloy, G. Ginolhac, F. Pascal and P. Forster. Clutter subspace estimation in low-rank heterogeneous noise context. *Signal Processing, IEEE Transactions on*, **63**(9), pp. 2173-2182, 2015.
- A. Breloy, Y. Sun, P. Babu, G. Ginolhac and D. P. Palomar. Robust rank-constrained Kronecker covariance matrix estimation. In *Proc. of Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, pp.810-814, Nov. 2016.
- J. P. Burg, D. G. Luenberger and D. L. Wenger. Estimation of structured covariance matrices. *Proc. IEEE*, **70**(9), pp.963-974, 1982.
- L. Cai and H. Wang. A Persymmetric Multiband GLR Algorithm. *Aerospace and Electronic Systems, IEEE Transactions on*, pp. 806-816, 1992.

Bibliography (3)

- Y. Chen, A. Wiesel, and A. O. Hero. Robust shrinkage estimation of high-dimensional covariance matrices. *Signal Processing, IEEE Transactions on*, **59**(9), pp.4097-4107, 2011.
- A. Collas, A. Breloy, C. Ren, G. Ginolhac, J.-P. Ovarlez. Riemannian Optimization for Non-Centered Mixture of Scaled Gaussian Distributions. *Signal Processing, IEEE Transactions on*, **71**, pp.2475-2490, 2023.
- A. Collas, F. Bouchard, A. Breloy, G. Ginolhac, C. Ren, J.-P. Ovarlez, Probabilistic PCA from Heteroscedastic Signals: Geometric Framework and Application to Clustering. *Signal Processing, IEEE Transactions on*, **69**, pp.6546-6560, Nov. 2021.
- E. Conte and M. Longo. Characterisation of radar clutter as a spherically invariant random process. *IEE Proceedings F (Communications, Radar and Signal Processing)*, **134**(2), pp.191-197, April 1987.
- E. Conte, M. Lops, and G. Ricci. Adaptive detection schemes in compound-Gaussian clutter. *Aerospace and Electronic Systems, IEEE Transactions on*, **34**(4), pp.1058-1069, 1998.
- E. Conte and A. De Maio. Exploiting persymmetry for CFAR detection in compound-Gaussian clutter. *Aerospace and Electronic Systems, IEEE Transactions on*, **39**(2), pp.719-724, April 2003.
- E. Conte and A. De Maio. Maximum likelihood estimation of a structured covariance matrix with a condition number constraint. *Signal Processing, IEEE Transactions on*, **60**(6), pp.3004-3021, 2012.
- E. Conte, M. Lops and G. Ricci. Asymptotically optimum radar detection in compound-Gaussian clutter, *Aerospace and Electronic Systems, IEEE Transactions on*, **31**(2), pp.617-625, 1995.
- R. Couillet and M. Debbah. Random matrix methods for wireless communications. Cambridge University Press, 2011.

Bibliography (4)

- R. Couillet and M. McKay. Large dimensional analysis and optimization of robust shrinkage covariance matrix estimators. *Journal of Multivariate Analysis*, **131**, pp.99-120, 2014.
- R. Couillet, M. S. Greco, J.-P. Ovarlez and F. Pascal. RMT for whitening space correlation and applications to radar detection. *IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing(CAMSAP)*, pages 149-152, 2015.
- R. Couillet. Robust spiked random matrices and a robust G-MUSIC estimator. *Journal of Multivariate Analysis*, **140**, pp.139-161, 2015.
- R. Couillet, F. Pascal, and J. W. Silverstein. The random matrix regime of Maronna's M -estimator with elliptically distributed samples. *Journal of Multivariate Analysis*, **139**, pp.56-78, 2015.
- A. De Maio. Maximum likelihood estimation of structured persymmetric covariance matrices. *Elsevier Signal Processing*, **83**(3), pp.633-640, 2003.
- C. Doz, R. Couillet, C. Ren, J.-P. Ovarlez, Large dimensional analysis of LS-SVM transfer learning: Application to POLSAR classification. *Acoustics, Speech and Signal Processing (ICASSP), 2023 IEEE International Conference on*, Rhodes Island, Greece, June 2023.
- P. Develter, J. Bosse, O. Rabaste, P. Forster, J.-P. Ovarlez. On the False Alarm Probability of the Normalized Matched Filter for off-grid targets: A geometrical approach and its validity conditions. *Signal Processing, IEEE Transactions on*, to appear, December 2023,
- G. Drasković and F. Pascal. New insights into the statistical properties of M -estimators. *Signal Processing, IEEE Transactions on*, **66**(16), pp.4253-4263, 2008.
- S. Fortunati, F. Gini, and M. S. Greco. Matched, mismatched, and robust scatter matrix estimation and hypothesis testing in complex t-distributed data. *EURASIP Journal on Advances in Signal Processing*, 2016-123, Nov. 2016.

Bibliography (5)

- G. Frahm, *Generalized elliptical distributions: theory and applications*, Ph.D. thesis, Universitätsbibliothek, 2004.
- 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. *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*, Munich, Germany, July 2012.
- J. Frontera-Pons, M. A. Veganzones, S. Velasco-Forero, F. Pascal, J.-P. Ovarlez and J. Chanussot. Robust anomaly detection in hyperspectral imaging. *IEEE Geoscience and Remote Sensing Symposium*, pp.4604-4607, 2014.
- J. Frontera-Pons, M. A. Veganzones, F. Pascal and J.-P. Ovarlez. Hyperspectral anomaly detectors using robust estimators. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **9**(2), pp.720-731, 2016.
- J. Frontera-Pons, M. A. Veganzones, S. Velasco-Forero, F. Pascal, J.-P. Ovarlez and J. Chanussot. Robust anomaly detection in hyperspectral imaging. *IEEE Geoscience and Remote Sensing Symposium*, pp.4604-4607, 2014.
- J. Frontera-Pons, M. A. Veganzones, F. Pascal and J.-P. Ovarlez. Hyperspectral anomaly detectors using robust estimators. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **9**(2), pp.720-731, 2016.
- J. Frontera-Pons, F. Pascal and J.-P. Ovarlez. Adaptive nonzero-mean Gaussian detection. *Geoscience and Remote Sensing, IEEE Transactions on*, **55**(2), pp.1117-1124, 2017.
- F. Gini and M. S. Greco. Covariance matrix estimation for CFAR detection in correlated heavy-tailed clutter. *Signal Processing*, **82**(12), pp.1847-1859, 2002.
- M. S. Greco, F. Gini, and M. Rangaswamy, Statistical analysis of measured polarimetric clutter data at different range resolutions, *IEE Proceedings - Radar, Sonar and Navigation*, **153**(6), pp.473-481, 2006.
- G. Ginolhac, P. Forster, J.-P. Ovarlez and F. Pascal. STAP à Rand Réduit, Robuste et Persymétrique. *Traitement du Signal*, **28**(1-2), pp.143-170. Numéro Spécial : Traitements Spatio-Temporels Adaptatifs (STAP). 2011.

Bibliography (6)

- G. Ginolhac, P. Forster, F. Pascal and J.-P. Ovarlez. Derivation of the bias of the normalized sample covariance matrix in a heterogeneous noise with application to low-rank STAP filter. *Signal Processing, IEEE Transactions on*, **60**(1), pp.514-518, 2012.
- G. Ginolhac, P. Forster, F. Pascal and J.-P. Ovarlez. Performance of two low-rank STAP filters in a heterogeneous noise. *Signal Processing, IEEE Transactions on*, **61**(1), pp. 57-61, 2013.
- G. Ginolhac and P. Forster. Approximate distribution of the low-rank adaptive normalized matched filter test statistic under the null hypothesis. *Aerospace and Electronic Systems, IEEE Transactions on*, **52**(4), pp.2016-2023, 2016.
- A. Haimovich. The eigencanceler: adaptive radar by eigenanalysis methods. *Aerospace and Electronic Systems, IEEE Transactions on*, **32**(2), pp.532-542, 1996.
- P. J. Huber. Robust estimation of a location parameter, *The Annals of Mathematical Statistics*, **35**(1), 1964.
- D. Kelker. Distribution theory of spherical distributions and a location-scale parameter generalization. *Sankhya: The Indian Journal of Statistics, Series A*, **4**(32), pp.419-430, 1970.
- J. T. Kent and D. E. Tyler. Redescending M-estimates of multivariate location and scatter *Annals of Statistics*, **19**(4), pp.2102-2119, 1991.
- I. P. Kirsteins and D. W. Tufts. Adaptive detection using low-rank approximation to a data matrix. *Aerospace and Electronic Systems, IEEE Transactions on*, **30**(1), pp.55-67, 1994.
- O. Lerda, A. Mian, G. Ginolhac, J.-P. Ovarlez, D. Charlot. Robust Detection for Mills Cross Sonar. *Oceanic Engineering, IEEE Transactions on*, to appear, Decembre 2023.

Bibliography (7)

- M. Mahot, F. Pascal, P. Forster and J.-P. Ovarlez. Asymptotic properties of robust complex covariance matrix estimates, *Signal Processing, IEEE Transactions on*, **61**(13), pp.3348-3356, 2013.
- R. A. Maronna. Robust M-estimators of multivariate location and scatter. *The Annals of Statistics*, **4**(1), pp.51-67, 1976.
- R. A. Maronna, D. R. Martin and J. V. Yohai. *Robust Statistics: Theory and Methods*, Wiley Series in Probability and Statistics, John Wiley & Sons, 2006.
- B. Mériaux, C. Ren, A. Breloy, M. N. El Korso, and P. Forster. Efficient estimation of Kronecker product of linear structured scatter matrices under t-distribution. *In Proc. of European Signal Processing Conference (EUSIPCO)*, 2020.
- B. Mériaux, C. Ren, M. N. El Korso, A. Breloy, and P. Forster. Asymptotic performance of complex M-estimators for multivariate location and scatter estimation. *IEEE Signal Processing Letters*, **26**(2), pp.367-371, February 2019.
- B. Mériaux, C. Ren, M. N. El Korso, A. Breloy, and P. Forster. Robust estimation of structured scatter matrices in (mis)matched models. *Signal Processing*, **165**, pp.163-174, December 2019.
- B. Mériaux, A. Breloy, C. Ren, M. N. El Korso, and P. Forster. Modified sparse subspace clustering for radar detection in non-stationary clutter. *In Proc. of IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2019.
- B. Mériaux, C. Ren, A. Breloy, M. N. El Korso, P. Forster, and J.-P. Ovarlez. On the recursions of Robust COMET algorithm for convexly structured shape matrix. *In Proc. of European Signal Processing Conference (EUSIPCO)*, 2019.
- A. Mian, J.-P. Ovarlez, A. M. Atto and G. Ginolhac. Design of New Wavelet Packets Adapted to High-Resolution SAR Images With an Application to Target Detection. *Geoscience and Remote Sensing, IEEE Transactions on*, **57**(6), pp.3919-3932, June 2019.

Bibliography (8)

- M. Muzeau, C. Ren, S. Angelliaume, M. Datcu, J.-P. Ovarlez, Self-Supervised Learning Based Anomaly Detection in Synthetic Aperture Radar Imaging. *IEEE Open Journal of Signal Processing*, **3**, pp.440-449, 2022.
- R. Nitzberg. Application of maximum likelihood estimation of persymmetric covariance matrices to adaptive processing. *Aerospace and Electronic Systems, IEEE Transactions on*, **16**(1), pp.124-127, 1980.
- E. Ollila, D. E. Tyler, V. Koivunen, and H. V. Poor. Complex elliptically symmetric distributions: Survey, new results, and applications. *Signal Processing, IEEE Transactions on*, **60**(11), pp.5597-5625, November 2012.
- E. Ollila and D. E. Tyler. Regularized M -estimators of scatter matrix. *Signal Processing, IEEE Transactions on*, **62**(22), pp.6059-6070, 2014.
- J.-P. Ovarlez. Détection en Environnement Non-Gaussien. ONERA Report RT 6/5275 SN 1996.
- J.-P. Ovarlez. Evaluation de la Détectabilité des Cibles Volant à Basse Altitude par des Radars Sols. ONERA Report RT 25/5272 SY, 1995.
- J.-P. Ovarlez, F. Pascal and A. Breloy. Asymptotic detection performance analysis of the robust adaptive normalized matched filter, *IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pp.137-140, 2015.
- J.-P. Ovarlez, F. Pascal, P. Forster, G. Ginolhac and M. Mahot. Traitement STAP et Modélisation SIRV : Robustesse et Persymétrie. *Traitement du Signal*, **28**(1-2), pp.113-142. Numéro Spécial : Traitements Spatio-Temporels Adaptatifs (STAP). 2011.

Bibliography (9)

- G. Pailloux, P. Forster, J.-P. Ovarlez and F. Pascal. Persymmetric adaptive radar detectors, *Aerospace and Electronic Systems, IEEE Transactions on*, **47**(4), pp.2376-2390, 2011.
- F. Pascal and J.-P. Ovarlez. Asymptotic properties of the robust ANMF, *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp.2594-2598, 2015.
- F. Pascal, Y. Chitour, J.-P. Ovarlez, P. Forster and P. Larzabal. Covariance structure maximum-likelihood estimates in compound Gaussian noise: Existence and algorithm analysis, *Signal Processing, IEEE Transactions on*, **56**(1), pp.34-48, 2008.
- F. Pascal, Y. Chitour, and Y. Quek. Generalized robust shrinkage estimator and its application to STAP detection problem. *Signal Processing, IEEE Transactions on*, **62**(21), pp.5640-5651, 2014.
- F. Pascal, P. Forster, J.-P. Ovarlez, and P. Larzabal. Performance analysis of covariance matrix estimates in impulsive noise. *Signal Processing, IEEE Transactions on*, **56**(6), pp.2206-2217, June 2008.
- M. Rangaswamy, F. C. Lin and K. R. Gerlach. Robust adaptive signal processing methods for heterogeneous radar clutter scenarios. *Proceedings of the 2003 IEEE Radar Conference*, Huntsville, AL, USA, pp.265-272, 2003.
- M. Rangaswamy, F. C. Lin, K. R. Gerlach. Robust adaptive signal processing methods for heterogeneous radar clutter scenarios, *Signal Processing*, **84**(9), pp.1653-1665, 2004.
- I. Reed and X. Yu. Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, **38**(10), pp.1760-1770, 1990.
- K. J. Sangston, F. Gini, and M. S. Greco. Adaptive detection of radar targets in compound-Gaussian clutter. *Proc. of IEEE Radar Conference*, pp.587-592, May, 2015.

Bibliography (10)

- K. J. Sangston, F. Gini, and M. S. Greco. Coherent radar target detection in heavy-tailed compound-Gaussian clutter. *Aerospace and Electronic Systems, IEEE Transactions on*, **48**(1), pp.64-77, January 2012.
- R. O. Schmidt. Multiple emitter location and signal parameter estimation. *Acoustics Speech and Signal Processing, IEEE Transactions on*, **34**(3), pp.276-280, 1986.
- I. Soloveychik and A. Wiesel. Tyler's covariance matrix estimator in elliptical models with convex structure. *Signal Processing, IEEE Transactions on*, **62**(20), pp.5251-5259, October 2014.
- Y. Sun, P. Babu and D. P. Palomar. Robust estimation of structured covariance matrix for heavy-tailed elliptical distributions. *Signal Processing, IEEE Transactions on*, **14**(64), pp.3576-3590, 2016.
- Y. Sun, A. Breloy, P. Babu, D. P. Palomar, F. Pascal and G. Ginolhac. Low-complexity algorithms for low-rank clutter parameters estimation in radar systems. *Signal Processing, IEEE Transactions on*, **64**(8), pp.1986-1998, 2015.
- Y. Sun, P. Babu and D. P. Palomar. Regularized Tyler's scatter estimator: Existence, uniqueness, and algorithms. *Signal Processing, IEEE Transactions on*, **62**(19), pp.5143-5156, 2014.
- D. E. Tyler. A distribution-free M-estimator of multivariate scatter, *The Annals of Statistics*, **15**(1), 1987.
- D. E. Tyler. Some results on the existence, uniqueness, and computation of the M-estimates of multivariate location and scatter. *SIAM Journal on Scientific and Statistical Computing*, **9**(2), pp.354-362, 1988.
- D. E. Tyler. Some issues in the robust estimation of multivariate location and scatter. In *Directions in Robust Statistics and Diagnostics*, **34**, pp.327-336. Springer New York, 1991.
- A. Wiesel. Geodesic convexity and covariance estimation. *Signal Processing, IEEE Transactions on*, **60**(12), pp.6182-6189, Dec. 2012.
- K. Yao. A representation theorem and its application to spherically invariant random processes, *Information Theory, IEEE Transactions on*, **19**(2), 1973.