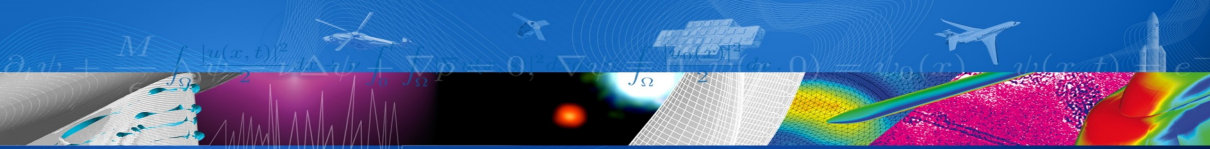


Tutorial on Robust Estimation and Detection Schemes in non-Standard Conditions for Radar, Array Processing and Imaging

Jean-Philippe Ovarlez

ONERA, The French Aerospace Lab,
ElectroMagnetism and Radar Department,
Advanced Signal Processing Unit
& CentraleSupélec SONDRA Laboratory

Séminaire Recherche M2R ATSI - 9 janvier 2024

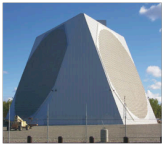
ONERA

THE FRENCH AEROSPACE LAB

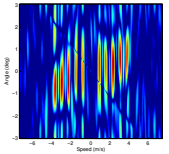
Plan

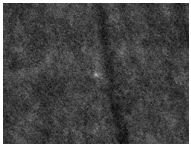
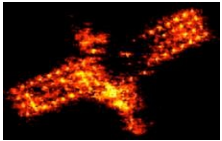
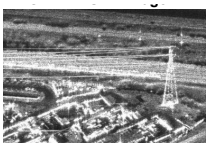
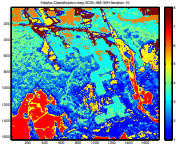
Radar and Imaging Sensors

RADAR = RAdio Detection And Ranging


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



Detection maps **ISAR Image** **SAR Image** **SAR Classification**

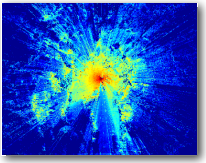
- but also: estimates parameters (range, radial velocity, angles of presentation, acceleration, amplitude (related to Radar Cross Section), etc.),
- images, classifies, recognizes.

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

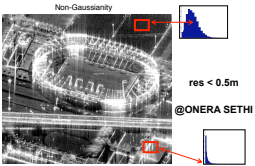
Radar and Imaging Sensors - New challenges

Positioning: facing the new non-standard conditions

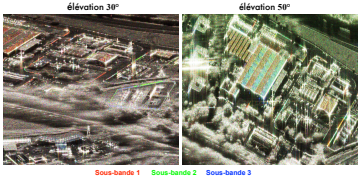
- **Complex Environments**: ground, dynamic environments (sea, ionosphere), heterogeneous, non-Gaussian, reverberating.
- **Complex targets**: small RCS, extended targets, fluctuating, dispersive, anisotropic.
- **Sensor Diversity**: temporal, spatial, polarimetric, interferometric, spectral.
- **Improvement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



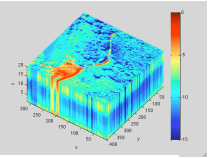
Heterogeneous Environments



Non-Gaussian Environments

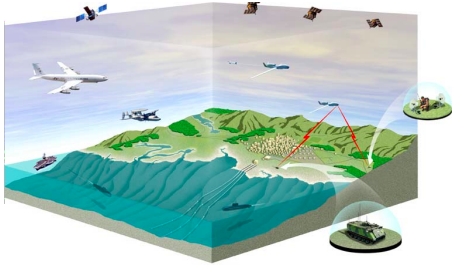


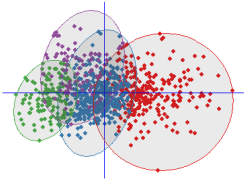
Non-Stationary Targets and Environments



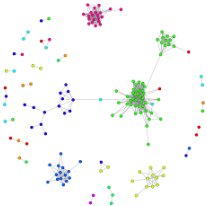
Curse of Dimensionality

Applicative Context





Finance
- Time Series
- Portofolio Optimization
- Risk Management
- Classification
- Prediction



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
- Graphes Analysis
- Learning Techniques

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 target embedded in these complex environments,
- Exploitation of emerging statistical Signal Processing techniques: Time-Frequency Analysis, Random Matrix Theory, Clustering, Compressive Sensing, etc.

Special issue: *M. S. Greco et al.*, Introduction to the Issue on Advanced Signal Processing Techniques for Radar Applications, IEEE Journal of Selected Topics in Signal Processing, 2015.
Book: *M. S. Greco and A. De Maio*, *Modern Radar Detection Theory*, Scitech Publishing, IET, 2015.

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 target embedded in these complex environments,
- Exploitation of emerging statistical Signal Processing techniques: Time-Frequency Analysis, Random Matrix Theory, Clustering, Compressive Sensing, etc.

Special issue: *M. S. Greco et al.*, Introduction to the Issue on Advanced Signal Processing Techniques for Radar Applications, IEEE Journal of Selected Topics in Signal Processing, 2015.
Book: *M. S. Greco and A. De Maio*, *Modern Radar Detection Theory*, Scitech Publishing, IET, 2015.

Tutorial Description

Survey on

- General statistical non-Gaussian modeling (Compound Gaussian, Spherically, Elliptically random processes),
- Robust covariance matrix estimation schemes (MLE, M-estimators),
- Robust detection schemes (Adaptive Normalized Matched Filter).

2 Main Parts and 1 Annex

- **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.
- **Annex:** Robust Model Order Selection Using Random Matrix Theory.

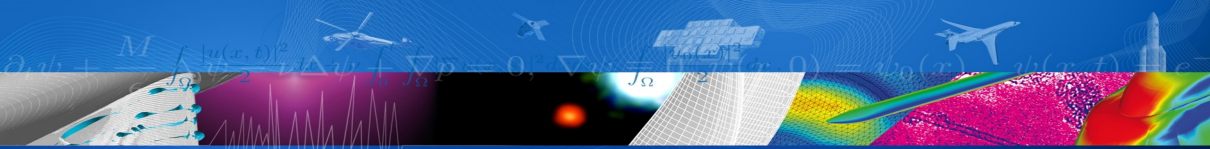
Acknowledgements to my colleagues

- Frédéric Pascal, L2S CentraleSupelec, Gif sur Yvette, France,
- Philippe Forster, ENS Cachan, France,
- Guillaume Ginolhac, Annecy University, France,
- Romain Couillet, L2S CentraleSupelec, Gif sur Yvette, France,
- and all my former PhD Students: E. Jay, M. Mahot, P. Formont, J. Frontera-Pons, A. Mian, E. Terreaux, etc.

References

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

`http://www.jeanphilippeovarlez.com`



Tutorial on Robust Estimation and Detection Schemes in non-Standard Conditions for Radar, Array Processing and Imaging

Jean-Philippe Ovarlez

ONERA, The French Aerospace Lab,
ElectroMagnetism and Radar Department,
Advanced Signal Processing Unit
& CentraleSupélec SONDRA Laboratory

Séminaire Recherche M2R ATSI - 9 anvier 2024

ONERA

THE FRENCH AEROSPACE LAB

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
- **Annex:** Robust Model Order Selection Using Random Matrix Theory.

Part A

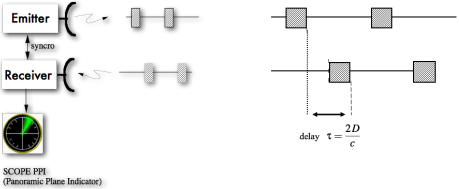
Background on Statistical Processing for Radar, Array Processing, SAR and Hyperspectral Imaging

Part A: Contents

Outline

Parameter Estimation - Range Measurement

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



- Radar emitted signal: $s_e(t) = u(t)\,\exp\left(2i\,\pi\,f_0\,t\right)$ where f_0 is the carrier frequency, and $u(.)$ 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 $b(.)$ is an additive noise.

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

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

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

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:

- $f_0 >> B$, where B is the bandwidth of baseband signal $u(.)$,
- $v << c$,
- $2\,B\,T << c/v$,

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

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

where $|\alpha'| = |\alpha|$ and where $f_d = \dfrac{2\,v}{c}\,f_0$ 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

Doppler Effect

Equation du Doppler

La fréquence Doppler est égale à la variation de d (distance) exprimée en longueurs d'onde

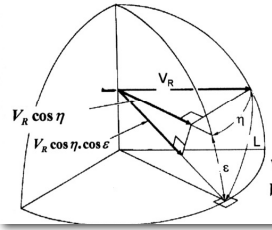
$$f_d = -\frac{1}{\lambda}\cdot\frac{\delta d}{\delta t}$$

d = 2R

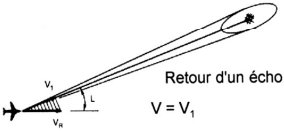
$$f_d = -\frac{2}{\lambda}\cdot\frac{\delta R}{\delta t}$$

$$\frac{f_d}{f} = \frac{2V}{c}$$

f_d fréquence Doppler,
$\delta R/\delta t = V$ vitesse relative entre le radar et la cible,
λ porteuse transmise.
$V = V_R + V_T$

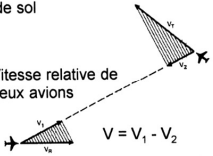


Vitesse relative du sol
$V_1 = V_R \cos\eta.\cos\varepsilon$



Retour d'un écho de sol

$V = V_1$



Vitesse relative de deux avions

$V = V_1 - V_2$

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):

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

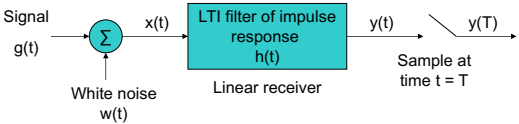
According to the physical transformation of X, we obtain the so-called Ambiguity functions *[Woodward 53, Kelly 65]*:

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

Link with the so-called Matched Filter and Pulse Compression

Goal: a) A matched filter is a filter used in communications to "match" a particular transit waveform. b) To optimize the design of the filter so as to minimize the effects of noise at the filter output and improve the detection of the pulse signal.

- A basic problem that often arises in the study of communication systems is that of detecting a pulse transmitted over a channel that is corrupted by channel noise (i.e. AWGN)
- Let us consider a received model, involving a linear time-invariant (LTI) filter of impulse response $h(t)$. The filter input $x(t)$ consists of a pulse signal $g(t)$ corrupted by additive channel noise $w(t)$ of zero mean and power spectral density $N_0/2$. The resulting output $y(t)$ is composed of $g_0(t)$ and $n(t)$, the signal and noise components of the input $x(t)$, respectively: $\begin{cases} x(t) = g(t) + w(t), \\ y(t) = g_0(t) + n(t) \end{cases}$ for $0 \leq t \leq T$



Link with the so-called Matched Filter and Pulse Compression

Signal to noise ratio is: $SNR = \frac{|g_0(T)|^2}{\sigma_n^2} = \frac{|g_0(T)|^2}{E\left[n^2(t)\right]}$, where $|g_0(T)|^2$ is the instantaneous power of the filtered signal, $g(t)$ at point $t = T$, and σ_n^2 is the variance of the zero mean white Gaussian filtered noise (the signal is sampled at t = T to have the max power of the filtered signal). We have: $|g_0(t)|^2 = \left| \int G(f)\, H(f) e^{j2\pi ft}\, df \right|^2$ and $\sigma_n^2 = E\left[n^2(t)\right] = R_n(0)$ where $R_n(\tau) = \int \frac{N_0}{2}\, |H(f)|^2\, e^{2i\pi f\tau}\, df$. We have:

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

The equality holds only for $H(f) = k\, G^*(f)\, e^{-2i\pi fT}$, $\forall k \in \mathbb{C}$, 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 $\left\{d_i = \dfrac{c\,\tau_i}{2}\right\}_{i\in[1,N]}$. The received signal $s_r(t)$ is:

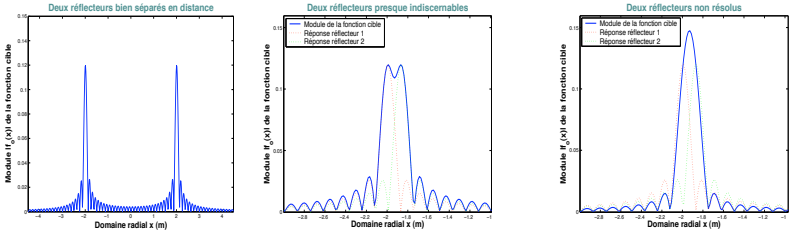
$$s_r(t) = \sum_{i=1}^{N} \alpha_i\, s_e(t-\tau_i) \overset{t\to f}{\Longrightarrow} S_r(f) = \sum_{i=1}^{N} \alpha_i\, S_e(f)\, e^{-2i\,\pi\, f\, \tau_i}\,.$$

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

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

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

Range resolution



(a) Distance réflecteurs : 4 m (b) Distance réflecteurs : 13 cm (c) Distance réflecteurs : 12.5 cm. (à la limite de résolution $\delta x = 12.5$ cm)

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

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

Velocity resolution

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

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

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

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

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

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

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:

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

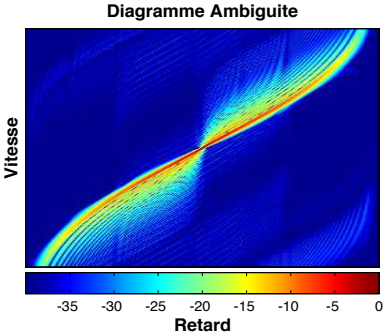
The radar processing (Matched Filter) leads to evaluate for all (τ,ν), the following expression:

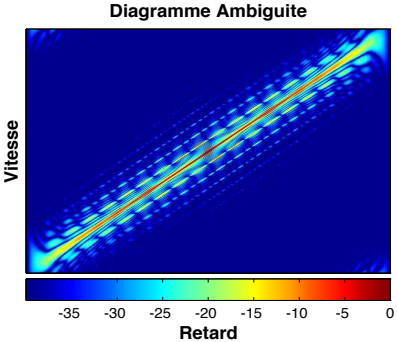
$$R(\tau,\nu) = \int_{-\infty}^{+\infty} s_r(t)\, s_e^*(t-\tau)\, e^{-2i\,\pi\,\nu\,t}\, dt\,.$$

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

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

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 $\sigma_t^2 = \int_{-\infty}^{+\infty} t^2 \, |s_e(t)|^2 \, dt \approx T^2$, $\sigma_f^2 = \int_{-\infty}^{+\infty} f^2 \, |S_e(f)|^2 \, df \approx B^2$ and the modulation index $m = \frac{-1}{2\pi} \, \mathrm{Im} \int_{-\infty}^{+\infty} t \, s_e(t) \, \frac{ds_e^*(t)}{dt} \, dt$. Under white Gaussian noise with variance σ^2, range and doppler accuracies are given by the following Cramer-Rao bounds *[Kay 93, Kay 98]*:

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

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

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

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

Outline

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?

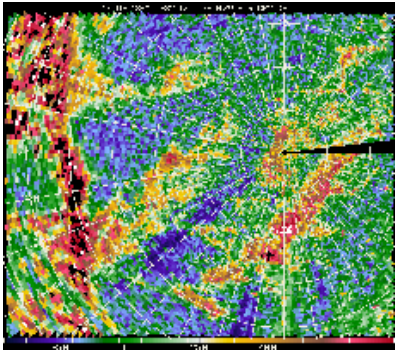
Clutter refers to radio frequency (RF) echoes returned from targets which are uninteresting to the radar operators and interfere with the observation of useful signals.
Such targets include natural objects such as ground, sea, precipitations (rain, snow or hail), sand storms, animals (especially birds), 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 ????

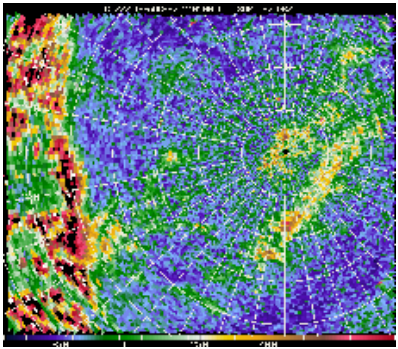
Noise and Clutter in Radar

Example of clutter map for different azimuth resolutions

resolution 3°



resolution 1°



Link Between Covariance Matrix and Power Spectral Density

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

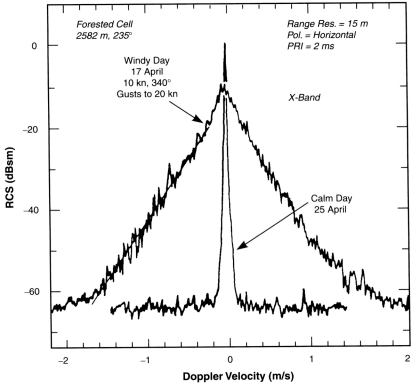


FIGURE 4.23 Power spectra of X-band radar returns from windblown trees.

[*Billingsley 1993*]

- Examples of some PSD models:

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

where $-1/(2\,T_r) \leq f \leq 1/(2\,T_r)$ (or $-\lambda/(4\,T_r) \leq v \leq \lambda/(4\,T_r)$).

- Autocorrelation function (Wiener-Khintchine Theorem):

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

- Covariance Matrix:

$$\mathbf{\Sigma} = E\left[\mathbf{z}\,\mathbf{z}^H\right] = \begin{pmatrix} \rho(0) & \ldots & \rho((m-1)\,T_r) \\ \vdots & \ddots & \vdots \\ \rho((m-1)\,T_r) & \ldots & \rho(0) \end{pmatrix}.$$

Outline

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 (SNR improvement),
- 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:

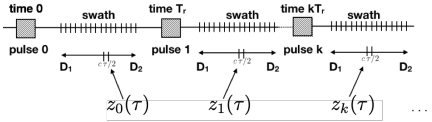
$$R(\tau,\nu) = \int_{-\infty}^{+\infty} s_r(t)\,s^*(t-\tau)\,e^{-2i\,\pi\,\nu\,t}\,dt = \sum_{n=0}^{m-1} e^{-2i\,\pi\,\nu\,n\,T_r} \int_0^{T_r} s_r(u+n\,T_r)\,s_e^*(u-\tau)\,\cancel{e^{-2i\,\pi\,\nu\,u}}\,du\,.$$

Range-Doppler Radar Processing

When supposing non migrating target (target stays in the same range bin during the duration T of the pulse, i.e. $BT \leq \frac{c}{2\,v}$) and neglecting the Doppler variation in the pulse, we can rewrite the processing as:

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

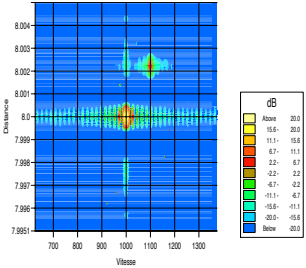
where $\mathbf{z} = (z_0(\tau), z_1(\tau), \ldots, z_{m-1}(\tau))^T$ and $\mathbf{p} = \left(1, e^{2i\,\pi\,\nu\,T_r}, \ldots, e^{2i\,\pi\,\nu\,(m-1)\,T_r}\right)^T$.



- 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 $n\,T_r$,

Range-Doppler Radar Processing

- For each range bin $c\,\tau/2$, compute the Discrete Fourier Transform over the m coefficients $\{z_n(\tau)\}_{n\in[0,m-1]}$ to characterize Doppler spectrum in the domain $\nu \in [0, 1/T_r]$.
- For non fluctuating target, the coefficients $\{z_n(\tau)\}_{n\in[0,m-1]}$ are generally constant over pulse train. This constant will be denoted by A in the following, A being the amplitude of the target over the burst.



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

Range Doppler Radar Processing

- Coherent Doppler processing brings an improvement of m on the Doppler resolution with regards to the one pulse processing ($\delta\nu = 1/(m\,T_r)$) 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.

Radar users have to choose the swath (range domain $c\,(k-1)\,Tr/2 \leq d_i\ < c\,k\,Tr/2$) relative the potential presence of targets and the Doppler support relative to the velocity of targets ($-c/(4\,Tr\,f_0) \leq v_i < c/(4\,Tr\,f_0)$).

Unfortunately, a large non-ambiguous swath and large non-ambiguous Doppler support cannot be chosen simultaneously.

	Range	Velocity
Resolution	$\frac{c}{2\,B}$ (depends on the signal)	$\frac{c}{2\,f_0\,m\,T_r}$ (does not depend on signal)
Ambiguity	$\frac{c\,T_r}{2}$	$\frac{c}{2\,f_0\,T_r}$

Characteristics of m pulses train with duration T, bandwidth B, PRF $1/T_r$ and carrier frequency f_0

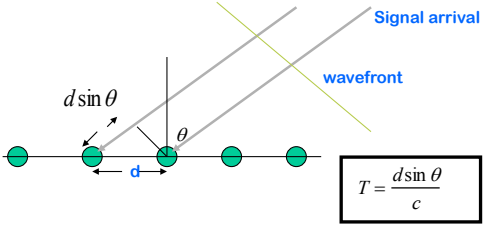
Outline

Array Processing

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

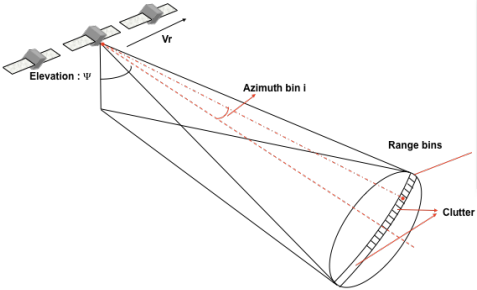
If the radar receives signal on antenna array, each antenna is collecting $s_r(t)$ delayed by the time shift $T = n\,d\,\sin\theta/c$ depending on its spatial position $n\,d$ ($n \in [0, N_s]$) on the array. Supposing that the array is non-dispersive ($N_s\,d\sin\theta << c/B$) , the concatenated $N_s \times m$-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:

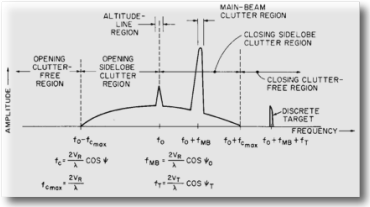
$$\mathbf{y} = A\,\mathbf{p} \otimes \left(1, e^{2i\pi\,f_0\,d\,\sin\theta/c}, \ldots, e^{2i\pi\,f_0\,(N_s-1)\,d\,\sin\theta/c}\right)^T + \mathbf{b}\,.$$



Outline

Space Time Adaptive Processing

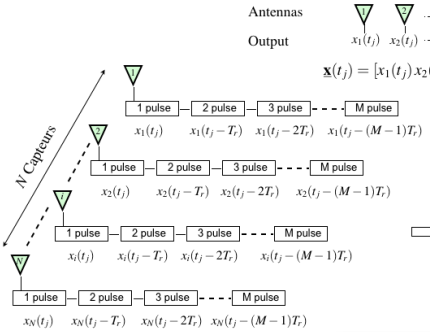




Goal: with a moving radar antenna array, detect targets embedded in ground clutter echos *[Ward 1998]*

- the radar collects strong ground clutter echos in the same angular domain as targets to detect.
- the spectral Doppler spread can be important

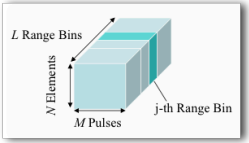
Data Collected in the Array



$$\underline{\mathbf{x}}(t_j) = [x_1(t_j)\, x_2(t_j) \ldots x_i(t_j) \ldots x_N(t_j)]^T$$

data collected in MNx1 data vector with period Tr in the range bin j

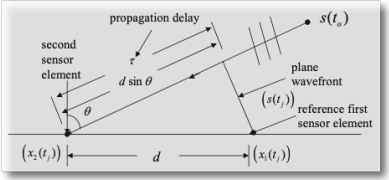
$$\underline{\mathbf{X}}_j = \begin{pmatrix} \underline{\mathbf{x}}(t_j) \\ \underline{\mathbf{x}}(t_j - T_r) \\ \vdots \\ \underline{\mathbf{x}}(t_j - kT_r) \\ \vdots \\ \underline{\mathbf{x}}(t_j - (M-1)T_r) \end{pmatrix}$$



Data collection on L range bins

Joint Use of the Temporal and Spatial Dimension

Use of the difference march for a source localized in the direction θ



$$\underline{\mathbf{x}}(t_j) = \begin{pmatrix} x_1(t_j) \\ x_2(t_j) \\ \vdots \\ x_k(t_j) \\ \vdots \\ x_N(t_j) \end{pmatrix} = s(t_j) \begin{pmatrix} 1 \\ e^{-2i\pi \frac{d}{\lambda} \sin\theta} \\ \vdots \\ e^{-2i\pi (k-1) \frac{d}{\lambda} \sin\theta} \\ \vdots \\ e^{-2i\pi (N-1) \frac{d}{\lambda} \sin\theta} \end{pmatrix} = s(t_j)\, \underline{\mathbf{a}}_N(\theta)$$

Use of the relative Doppler $f_d = \dfrac{2v}{\lambda} \sin\theta$ between the sensors and the target for all the period of repetition Tr

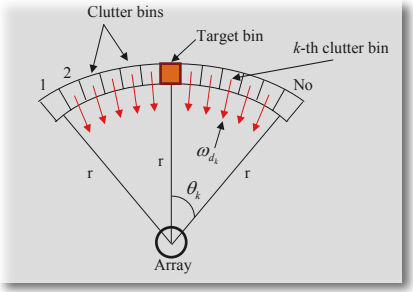
$$\underline{\mathbf{x}}(t_j - kT_r) = \underline{\mathbf{x}}(t_j)\, e^{-2i\pi(k-1) f_d T_r} = s(t_j)\, \underline{\mathbf{a}}_N(\theta)\, e^{-2i\pi(k-1) f_d T_r}$$

$$\underline{\mathbf{X}}_j = \begin{pmatrix} \underline{\mathbf{x}}(t_j) \\ \underline{\mathbf{x}}(t_j - T_r) \\ \vdots \\ \underline{\mathbf{x}}(t_j - kT_r) \\ \vdots \\ \underline{\mathbf{x}}(t_j - (M-1)T_r) \end{pmatrix} = s(t_j) \begin{pmatrix} \underline{\mathbf{a}}_N(\theta) \\ \underline{\mathbf{a}}_N(\theta)\, e^{-2i\pi f_d T_r} \\ \vdots \\ \underline{\mathbf{a}}_N(\theta)\, e^{-2i\pi(k-1) f_d T_r} \\ \vdots \\ \underline{\mathbf{a}}_N(\theta)\, e^{-2i\pi(M-1) f_d T_r} \end{pmatrix} = s(t_j) \begin{pmatrix} 1 \\ e^{-2i\pi f_d T_r} \\ \vdots \\ e^{-2i\pi(k-1) f_d T_r} \\ \vdots \\ e^{-2i\pi(M-1) f_d T_r} \end{pmatrix} \otimes \underline{\mathbf{a}}_N(\theta) = s(t_j)\, \underline{\mathbf{b}}_M(f_d) \otimes \underline{\mathbf{a}}_N(\theta)$$

$$\boxed{\underline{\mathbf{X}}_j = s(t_j)\, \underline{\mathbf{b}}_M(f_d) \otimes \underline{\mathbf{a}}_N(\theta) = s(t_j)\, \underline{\mathbf{a}}(\theta, f_d)}$$

STAP Detection in the Clutter

Detection Scheme in the j-th range bin: $\underline{\mathbf{X}}_j = \begin{cases} \alpha_0\, \underline{\mathbf{a}}\,(\theta_0, f_{d,0}) + \underline{\mathbf{X}}_c\,, & \text{under } H_1 \text{ hypothesis} \\ \underline{\mathbf{X}}_c\,, & \text{under } H_0 \text{ hypothesis} \end{cases}$



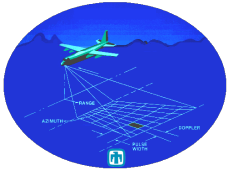
- $\underline{\mathbf{X}}_c = \sum_{k=1}^{N_0} c_k\,(f_{d,k})\; \underline{\mathbf{a}}\,(\theta_k, f_{d,k}) + \underline{\mathbf{n}}\,,$
- $f\,(d,k) = \dfrac{2\,v\,T_r}{d}\,\dfrac{2\,d\,\sin\theta_k}{\lambda} = \beta\,\dfrac{2\,d\,\sin\theta_k}{\lambda}\,,$
- $\underline{\mathbf{n}}$ is the thermal noise,
- $\{c_k\}_{k=1,N_0}$ the clutter amplitudes in direction θ_k,
- α_0, θ_0, $f_{d,0}$ are the unknown target parameters.

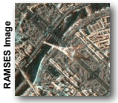
The clutter $\underline{\mathbf{X}}_c$ is generally assumed to be Gaussian distributed with unknown covariance matrix $\mathbf{\Sigma}$.

Outline

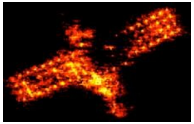
Background on SAR and Radar Imaging

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:

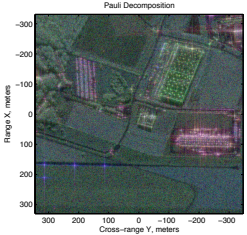
- Current use of very high spectral bandwidth and very high angular bandwidth leading to very high spatial resolution,
- Application to monitoring (detection, change detection), classification, 3D reconstruction, EM analysis, etc.

These applications require some physical diversity to reach good performances.

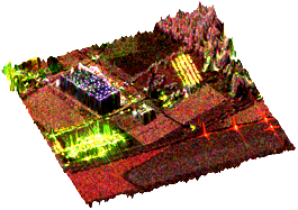
Multi-Channel SAR Images

Multi-channel SAR images automatically propose this diversity through:

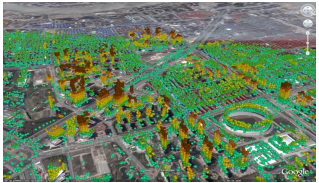
- polarimetric channels (POLSAR), interferometric channels (INSAR), polarimetric and interferometric channels (POLINSAR),
- multi-temporal, multi-passes SAR Image, etc.



EM behavior of the terrain in POLSAR images



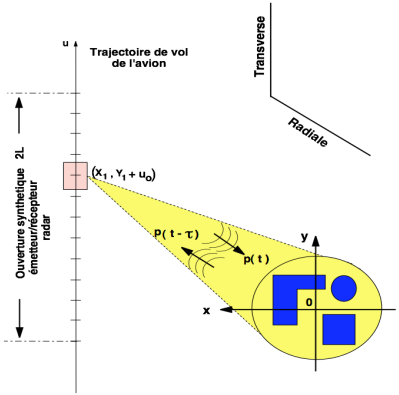
Estimation of the height in POLINSAR images



Analysis of the structures displacement in Shangai 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**.

SAR Processing



Goal of SAR Imaging: Invert the relation:

$$s_r(t,u) = \int\int_{\mathbb{R}^2} I(x,y)\, s_e\left(t - \frac{2}{c}\sqrt{(X_1 - x)^2 + (Y_1 + u - y)^2}\right) dx\, dy$$

Range Migration (RMA) SAR Processing Steps

$$s_r(t,u) = \int\int_{\mathbb{R}^2} I(x,y)\, s_e\left(t - \frac{2}{c}\sqrt{(X_1-x)^2+(Y_1+u-y)^2}\right)\, dx\, dy\,,$$

$$\Downarrow \quad t \overset{\mathcal{F}}{\Rightarrow} k = \frac{2\,f}{c}\,,$$

$$S_r(k,u) = S_e(k)\int\int_{\mathbb{R}^2} I(x,y)\, \exp\left(-2i\,\pi\,k\,\sqrt{(X_1-x)^2+(Y_1+u-y)^2}\right)\, dx\, dy\,,$$

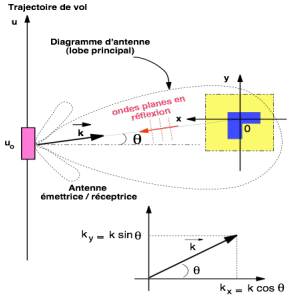
$$\Downarrow \quad u \overset{\mathcal{F}^{-1}}{\Rightarrow} k_u\,,$$

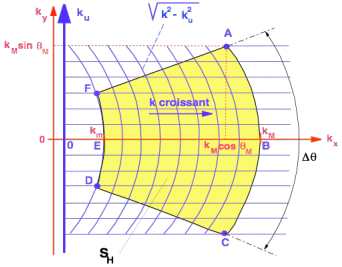
$$S_r(k,k_u) = S_e(k)\int\int_{\mathbb{R}^2} I(x,y)\, \exp\left(-2i\,\pi\,\left((X_1-x)\sqrt{k^2-k_u^2}+(Y_1-y)\,k_u\right)\right)\, dx\, dy\,,$$

$$\Downarrow \quad \begin{cases} k_x = \sqrt{k^2-k_u^2} \\ k_y = k_u \end{cases}$$

$$\boxed{S_r(k_x,k_y) = S_e(k)\,\exp\left(-2i\,\pi\,k_x\,X_1 + k_y\,Y_1\right)\int\int_{\mathbb{R}^2} I(x,y)\,\exp\left(2i\,\pi\,\left(k_x\,x+k_y\,y\right)\right)\, dx\, dy}$$

Range Migration Algorithm Principle



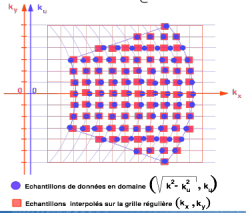


Support spectral du signal reçu $S_r(k_x, k_y)$

Interpolation de Stolt

$\sqrt{k^2 - k_u^2} \longrightarrow k_x$

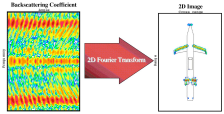
$k_u \longrightarrow k_y$



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 simply related to the SAR image $I(\mathbf{x})$ through:

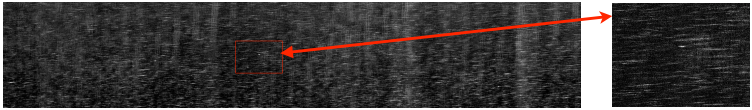
$$H(\mathbf{k}) = \int_{\mathcal{D}_{\mathbf{x}}} I(\mathbf{x}) \exp\left(-2\,i\,\pi\,\mathbf{k}^T\,\mathbf{x}\right) d\mathbf{x}\,.$$

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

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

With this model, all information relative to frequency f and angle θ are lost. Hence, spectral and angular diversities are lost *[Bertrand 94, Ovarlez 17]*.

Detection in SAR Images



Conventional SAR detection framework on a mono-channel SAR image mainly consists in locally thesholding the complex amplitude of pixel x_i:

- Global thresholding (Gaussian hypothesis): $\lambda = -\sigma^2 \log P_{fa}, \qquad \Lambda(x_i) = |x_i|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \lambda$,,
- Adaptive thresholding (Gaussian hypothesis) on N pixels:

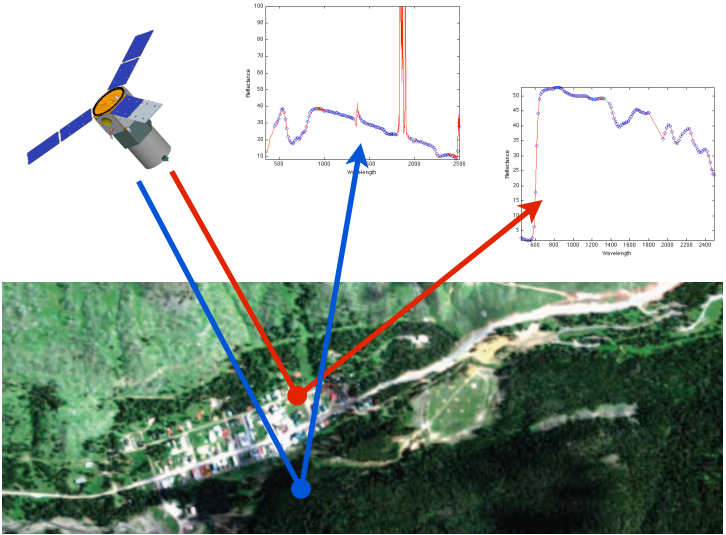
$$\Lambda(x_i) = \frac{|x_i|^2}{\frac{1}{N}\sum_{k\neq i}^{N} |x_k|^2} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda, \qquad \lambda = N\left(P_{fa}^{-1/N} - 1\right),$$

- Statistic-based thresholding (other distributions): $\lambda = f(P_{fa}), \qquad \Lambda(x_i) = g(x_i) \underset{H_0}{\overset{H_1}{\gtrless}} \lambda$.

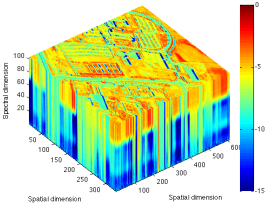
Adaptive multi-channels SAR detection framework can be extended with diversity contained in the steering vector $\mathbf{p}$ (polarimetry, interferometry, sub-looks and sub-bands decomposition *[Mian 19]*).

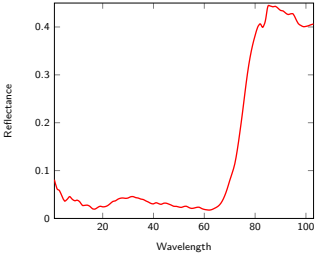
Outline

Hyperspectral Imaging (HSI)



Hyperspectral Imaging (HSI)





- **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 **p** - Regulation of False Alarm *[Frontera-Pons 17]*.

Outline

General Formulation of All the Detection Problems

Set of two binary hypotheses

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

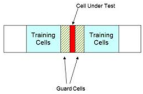
- $\mathbf{z}$ is a m-vector of data collected in a given measurement support. It can be range support, spatial support (Imaging), etc.
- The complex amplitude A of the target to detect is considered here deterministic (no fluctuation)
- The m-vector b represents the additive noise (thermal noise, photon 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, Interferomety, Wavelength, Spatial, joint Angular and Spectral information (STAP)

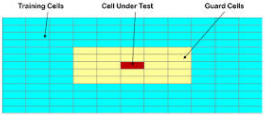
The problem here consists in choosing between H_1 hypothesis and H_0 hypothesis.

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:

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





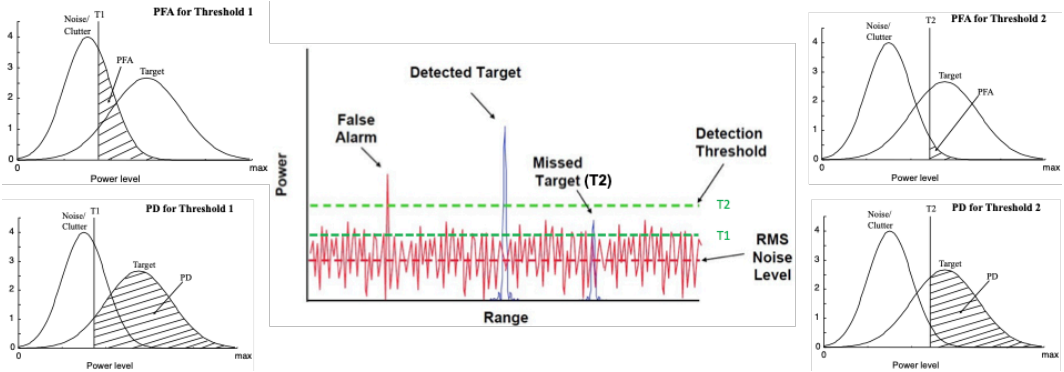
where the $\mathbf{z}_i$'s are n "signal-free" independent secondary data used to estimate the noise parameters. $\Rightarrow$ Neyman-Pearson criterion *[Kay 93, Kay 98]*

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

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

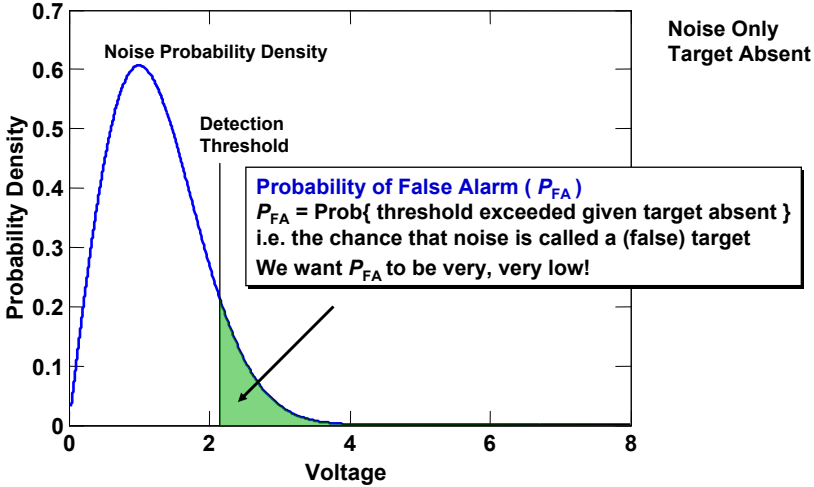
- Probability of False Alarm (type-I error): $P_{fa} = \mathbb{P}(\Lambda(\mathbf{z}) > \lambda/H_0)$
- Probability of Detection: $P_d = \mathbb{P}(\Lambda(\mathbf{z}) > \lambda/H_1)$ for different Signal-to-Noise Ratios (SNR),

Pd/Pfa



$$P_{fa} = \mathbb{P}(\Lambda(\mathbf{z}) > \lambda / H_0)\,, \qquad P_d = \mathbb{P}(\Lambda(\mathbf{z}) > \lambda / H_1)\,.$$

Pd/Pfa



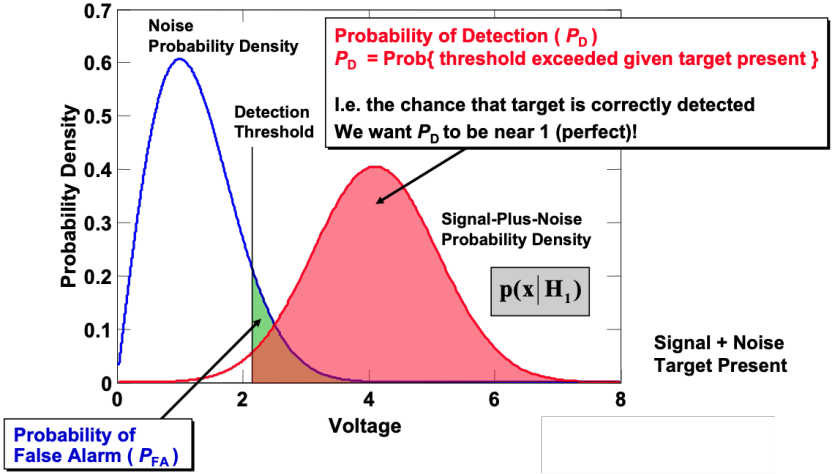
Probability of False Alarm (P_{FA})

P_{FA} = Prob{ threshold exceeded given target absent }

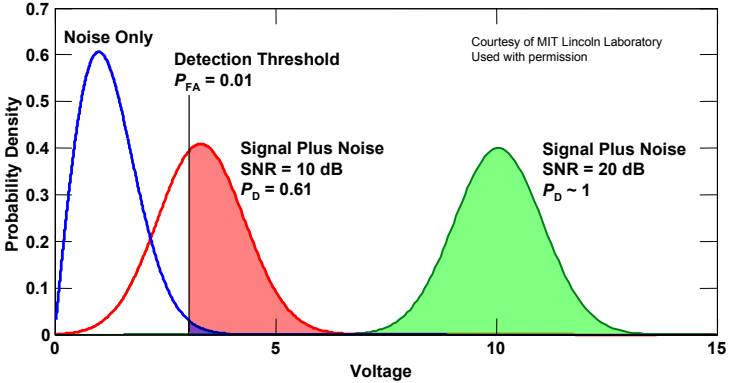
i.e. the chance that noise is called a (false) target

We want P_{FA} to be very, very low!

Pd/Pfa

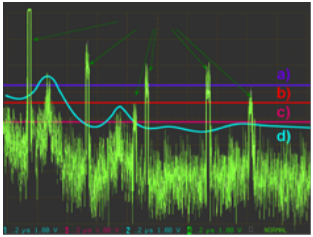

Probability of Detection (P_D)
P_D = Prob{ threshold exceeded given target present }
I.e. the chance that target is correctly detected
We want P_D to be near 1 (perfect)!
Noise
Probability Density
Detection
Threshold
Signal-Plus-Noise
Probability Density
$p(x|H_1)$
Signal + Noise
Target Present
Probability of
False Alarm (P_{FA})
Probability Density
Voltage
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
2
4
6
8

Pd/Pfa



- P_D increases with target SNR for a fixed threshold (P_{FA})
- Raising threshold reduces false alarm rate and increases SNR required for a specified Probability of Detection

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.

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:

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

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 $\boldsymbol{\mu}$:

$$\widehat{\boldsymbol{\mu}} = \underset{\boldsymbol{\mu}}{\operatorname{argmax}}\, p_{\mathbf{z}/H_0}(\mathbf{z}, \boldsymbol{\mu})\,.$$

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

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

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

Sample Covariance Matrix: MLE of the Gaussian problem

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

Outline

Modeling Homogeneous Gaussian Noise/Clutter

Problem to solve in Gaussian environment

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

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

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

$\Longrightarrow$ **All is known for Gaussian assumption!**

Sample Covariance Matrix (SCM)

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

Properties of the SCM in homogeneous Gaussian noise/clutter

Properties of the SCM

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

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

$$\text{where} \quad \begin{array}{l} \mathbf{C} = \left(\mathbf{\Sigma}^* \otimes \mathbf{\Sigma}\right) \\ \mathbf{P} = \left(\mathbf{\Sigma}^* \otimes \mathbf{\Sigma}\right) \mathbf{K}_{m^2,m^2}\,. \end{array}$$

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

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 $\mathbf{z}_k$:

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

- This estimate is unbiased, efficient, Wishart distributed,
- n can represent any samples support 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

Example 1 - Detection Schemes in Gaussian Noise

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:

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

The Log-Likelihood function test $\log \dfrac{p_{\mathbf{z}/H_1}(\mathbf{z})}{p_{\mathbf{z}/H_0}(\mathbf{z})}$ can be simplify as: $\Lambda(\mathbf{z}) = \mathrm{Re}\left(\mathbf{p}^H\,\mathbf{\Sigma}^{-1}\,\mathbf{z}\right) \underset{H_0}{\overset{H_1}{\gtrless}} \lambda$.

The statistic of the test becomes:

$$\boxed{\Lambda(\mathbf{z}) \sim \mathcal{N}\left(0, \mathbf{p}^H\,\mathbf{\Sigma}^{-1}\,\mathbf{p}\right) \text{ under } H_0} \text{ and}$$

$$\boxed{\Lambda(\mathbf{z}) \sim \mathcal{N}\left(\mathrm{Re}\left(A^H\,\mathbf{p}^H\,\mathbf{\Sigma}^{-1}\,\mathbf{p}\right), \mathbf{p}^H\,\mathbf{\Sigma}^{-1}\,\mathbf{p}\right) \text{ under } H_1}$$

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:

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

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

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

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

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

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

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

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

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

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

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

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:

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

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

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

.

Example 3 - Normalized Matched Filter (2)

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

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

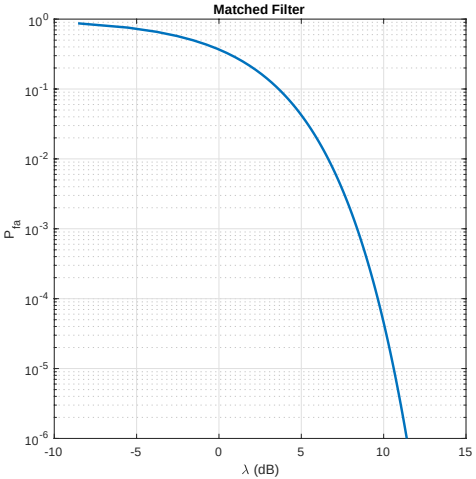
We can note that the NMF is invariant with respect to a change scale for $\mathbf{p}$, $\mathbf{z}$ or $\boldsymbol{\Sigma}$. Let $SNR = |A|^2 \mathbf{p}^H \boldsymbol{\Sigma}^{-1} \mathbf{p}$ be the Signal to Noise Ratio of the target to be detected. Under H_0 hypothesis, $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}_m,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

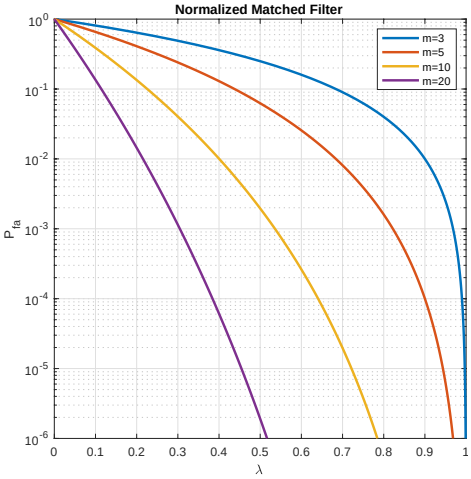
$$P_{fa} = \mathbb{P}\left(\Lambda_{NMF}(\mathbf{z}) > \lambda_{NMF}/H_0\right) = \left(1 - \lambda_{NMF}\right)^{m-1} \,,$$

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

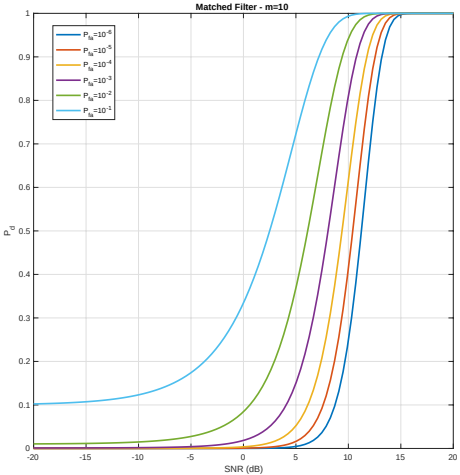
We can note that the threshold λ_{NMF} does not depend on unknown variance σ^2. The test is CFAR under H_0 hypothesis.

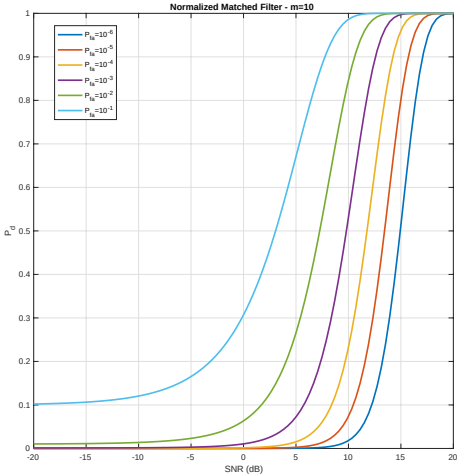
MF and NMF False Alarm regulation





MF and NMF Probability of Detection





Example 4 - Kelly and Adaptive Matched Filter (1)

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

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:

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

With formulas $\dfrac{\delta\,\log|\mathbf{\Sigma}^{-1}|}{\delta\,\mathbf{\Sigma}^{-1}} = \mathbf{\Sigma}^T$ and $\dfrac{\delta\,\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\,\mathbf{B}\right)}{\delta\,\mathbf{\Sigma}^{-1}} = \mathbf{B}^T$, we obtain:

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

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:

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

By denoting $\mathbf{S}=\sum_{k=1}^{n}\mathbf{z}_k\,\mathbf{z}_k^H$, we obtain $\underset{\mathbf{\Sigma}}{\mathrm{argmax}}\; p_{\mathbf{z},\{\mathbf{z}_k\}_k,\mathbf{\Sigma},A/H_1}(\mathbf{z}) = \dfrac{(\mathbf{z}-A\,\mathbf{p})\,(\mathbf{z}-A\,\mathbf{p})^H+\mathbf{S}}{n+1}$ and replacing these two expressions in the Generalized Log Likekihood Gatio leads to:

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

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

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

and $\min\limits_{A}|\mathbf{S}|\,\left(\|\mathbf{z}_s-A\,\mathbf{p}_s\|^2+1\right) = |\mathbf{S}|\,\left(\left\|\mathbf{P}_{\mathbf{p}_s}^{\perp}\,\mathbf{z}_s\right\|^2+1\right)$ 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:

$$\Lambda(\mathbf{z}) = \frac{\left|\mathbf{z}\,\mathbf{z}^H + \mathbf{S}\right|}{\min\limits_{A}\left|(\mathbf{z} - A\,\mathbf{p})\,(\mathbf{z} - A\,\mathbf{p})^H + \mathbf{S}\right|} = \frac{1 + \mathbf{z}_s^H\,\mathbf{z}_s}{1 + \mathbf{z}_s^H\,\mathbf{P}_{\mathbf{p}_s}^{\perp}\,\mathbf{z}_s} = \frac{1 + \mathbf{z}^H\,\mathbf{S}^{-1}\,\mathbf{z}_s}{1 + \mathbf{z}^H\,\mathbf{S}^{-1}\,\mathbf{z} - \dfrac{\left|\mathbf{p}^H\,\mathbf{S}^{-1}\,\mathbf{z}\right|^2}{\mathbf{p}^H\,\mathbf{S}^{-1}\,\mathbf{p}}} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda\,,$$

which is known as the so-called *Kelly*'s test *[Kelly 86]*:

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

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

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

The covariance matrix estimate $\hat{\mathbf{S}}_n = \dfrac{1}{n}\,\mathbf{S}$ is the *empirical* covariance matrix of the secondary data $\{\mathbf{z}_k\}_{k\in[1,n]}$ and is called *Sample Covariance Matrix* estimate. It corresponds to the Maximum Likelihood covariance matrix estimate under homogeneous Gaussian hypothesis.

Example 5 - Adaptive Normalized Matched Filter (1)

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

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

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]*:

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

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

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):

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

ANMF and Cosine Estimate

This detector is often called Cosine Estimator as it has the dimension of a cosinus squared between the steering vector $\mathbf{p}$ and the observation $\mathbf{z}$. Unlike the AMF Filter which measures the power of a scalar product, the ANMF measures an angle. It is therefore more sensible to a possible mismatch between $\mathbf{p}$ and $\mathbf{z}$.

Outline

Synthesis of CFAR Detection Schemes Under Gaussian Noise

- Adaptive Matched Filter *[Robey 92]*: $\Lambda_{AMF}(\mathbf{z}) = \dfrac{\left|\mathbf{p}^H \widehat{\mathbf{S}}_n^{-1} \mathbf{z}\right|^2}{\mathbf{p}^H \widehat{\mathbf{S}}_n^{-1} \mathbf{p}} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{AMF}$:

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

- Adaptive Kelly Filter *[Kelly 86]*: $\Lambda_{Kelly}(\mathbf{z}) = \dfrac{\left|\mathbf{p}^H \widehat{\mathbf{S}}_n^{-1} \mathbf{z}\right|^2}{\left(\mathbf{p}^H \widehat{\mathbf{S}}_n^{-1} \mathbf{p}\right)\left(n + \mathbf{z}^H \widehat{\mathbf{S}}_n^{-1} \mathbf{z}\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{Kelly}$:

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

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

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

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

Outline

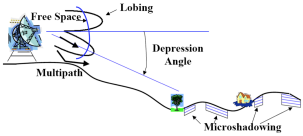
Examples of Gaussian Hypothesis Failure

High Resolution Radars

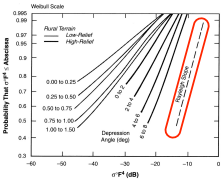
- Small number of scatterers in the cell under test - Varying number of scatterers from cell to cell - *Central Limit Theorem* non valid ⇒ non-Gaussianity *[Jakeman 80]*
- No validity of conventional tools based on Gaussian statistics *[Farina 87, Gini 00, Jay 02]*.

Low-Grazing angles Illumination Radar

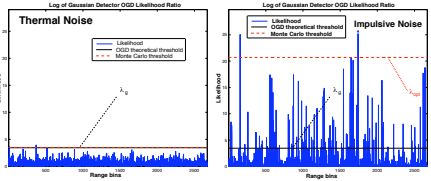
- Microshadowing ⇒ impulsive clutter *[Billingsley 93]*
- Transitions of clutter areas, heterogeneity of spatial area under test ⇒ difficulty to set up the detection test λ_{opt} and the Probability of False Alarm depending on the area.



Low-Grazing angle surveillance



Non-Gaussian behavior



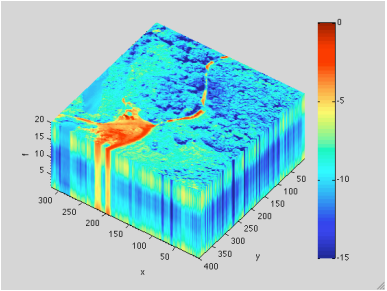
False Alarm regulation problem

Please refer to *[F. Gini, A. Farina and M. S. Greco 2001]*

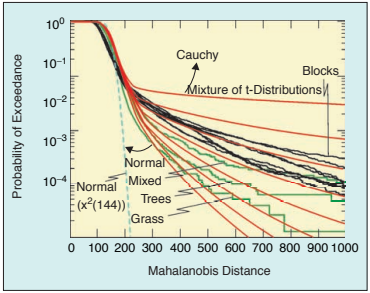
Examples of Gaussian Hypothesis Failure

- 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



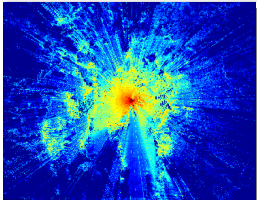
[*Manolakis 2002*]

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

Need of Better Approaches

Need to build alternatives to conventional approaches :



$\Rightarrow$ 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

- 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

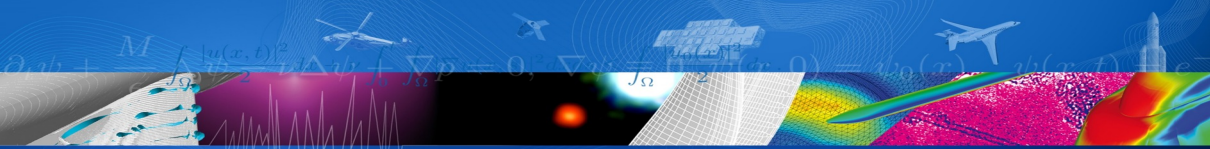
- A. Mian, *Robust Change Detection in Time Series SAR Images*, PhD Thesis, Université de Paris Saclay / ONERA, 2019.
- A. Mian, J.-P. Ovarlez, A. M. Atto and G. Ginolhac, Design of New Wavelet Packets Adapted to High-Resolution SAR Images With an Application to Target Detection, *Geoscience and Remote 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



Tutorial on Robust Estimation and Detection Schemes in non-Standard Conditions for Radar, Array Processing and Imaging

Jean-Philippe Ovarlez

ONERA, The French Aerospace Lab,
ElectroMagnetism and Radar Department,
Advanced Signal Processing Unit
& CentraleSupélec SONDRA Laboratory

Séminaire Recherche M2R ATSI - 9 janvier 2024

ONERA

THE FRENCH AEROSPACE LAB

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.
- **Annex:** Robust Model Order Selection Using Random Matrix Theory.

Part B

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

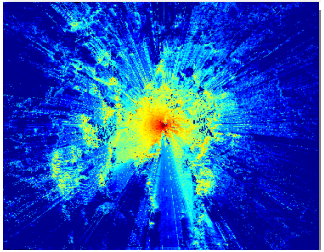
Applications and Results in Radar, STAP and Array Processing, SAR Imaging, Hyperspectral Imaging

Part B: Contents

Outline

Going to Robust Adaptive Detection

Generally, some parameters (*e.g.* second order statistic $\mathbf{\Sigma}$) are unknown and cannot be estimated through Gaussian methodology



$\Rightarrow$ 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

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 ($CE\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_z\right)$) if its PDF is *[Mahot 12, Ollila 12]*:

$$g_z(\mathbf{z}) = \pi^{-m} \left|\boldsymbol{\Sigma}\right|^{-1} h_z\left(\left(\mathbf{z}-\boldsymbol{\mu}\right)^H \boldsymbol{\Sigma}^{-1} \left(\mathbf{z}-\boldsymbol{\mu}\right)\right),$$

where $h_z : [0,\infty) \rightarrow [0,\infty)$ 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, $E\left[\mathbf{z}\,\mathbf{z}^H\right] = \alpha\,\boldsymbol{\Sigma}$ where α is known.

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

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:

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

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

- **Large class of distributions**: Gaussian ($p_\tau(\tau) = \delta(\tau - 1)$), K-distribution ($p_\tau$ 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*) $\mathbf{\Sigma}$,
- **Stochastic representation theorem**: $\boxed{\mathbf{z} =_d \boldsymbol{\mu} + \sqrt{\tau}\,\mathbf{A}\,\mathbf{x}}$, where $\tau \geq 0$ is the texture, independent of $\mathbf{x}$ and $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}_m, \mathbf{I})$.

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.

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

- Conditionnally 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.).

Outline

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:

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

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

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

This estimate does not depend on the texture τ_k but it is biased and share the same eigenvectors but have 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 $\mathbf{z}_i \sim CE_m(\mathbf{0}_m, \mathbf{\Sigma}, g_z)$ where $g_z(.)$ is known and where $\mathbf{\Sigma}$ is an unknown scatter matrix. The MLE $\widehat{\mathbf{\Sigma}}$ is set by solving

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

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

M-Estimator as MLE of the CES problem

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

Estimating the Covariance/Scatter Matrix

M-estimators:

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

PDF $g_z(.)$ specified: *MaximumLikelihood*-estimator of $\mathbf{\Sigma}$: $\widehat{\mathbf{\Sigma}} = \frac{1}{n} \sum_{i=1}^{n} \frac{-g_z'\left(\mathbf{z}_i^H \widehat{\mathbf{\Sigma}}^{-1} \mathbf{z}_i\right)}{g_z\left(\mathbf{z}_i^H \widehat{\mathbf{\Sigma}}^{-1} \mathbf{z}_i\right)} \mathbf{z}_i \mathbf{z}_i^H$,

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

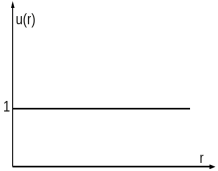
[Maronna 76, Kent 91, Maronna 06, Pascal 08, Mahot 13]

- 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

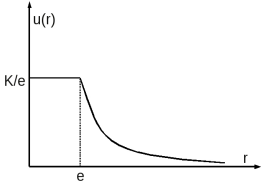
SCM:

$u(r) = 1$



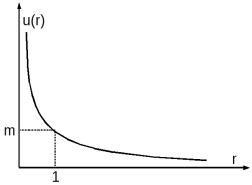
Huber's M-estimator:

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



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

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

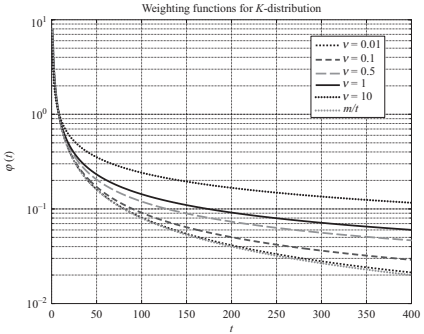
$$\widehat{\mathbf{\Sigma}}_{FPE} = \frac{m}{n} \sum_{k=1}^{n} \frac{\mathbf{z}_k \, \mathbf{z}_k^H}{\mathbf{z}_k^H \, \widehat{\mathbf{\Sigma}}_{FPE}^{-1} \, \mathbf{z}_k}.$$

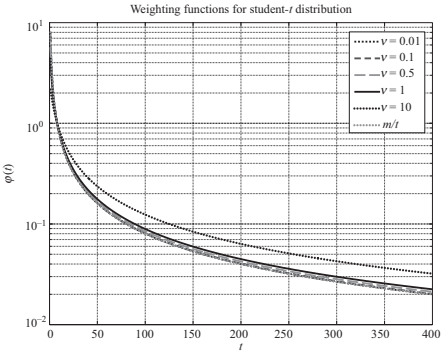
- The Tyler M-estimator does not depend on the texture (SIRV or CES distributions),
- Convergence of the algorithm: $\widehat{\mathbf{\Sigma}}_{n+1} = f\left(\widehat{\mathbf{\Sigma}}_n\right)$ with $f(\widehat{\mathbf{\Sigma}}) = \frac{m}{n} \sum_{k=1}^{n} \frac{\mathbf{z}_k \, \mathbf{z}_k^H}{\mathbf{z}_k^H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

  Existence, Uniqueness,
- $\widehat{\mathbf{\Sigma}}_{FPE}$ 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

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

Some Weighting Functions of M-estimators





$$u(t) = \varphi(t) = \frac{\sqrt{\nu}}{t}\,\frac{K_{\nu-m-1}\left(4\,\nu\,t\right)}{K_{\nu-m}\left(4\,\nu\,t\right)}, \qquad u(t) = \varphi(t) = \frac{\nu + 2\,m}{\nu + 2\,t}.$$

We have $\lim\limits_{\nu\to 0} \widehat{\mathbf{\Sigma}} = \widehat{\mathbf{\Sigma}}_{FPE}$ and $\lim\limits_{\nu\to\infty} \widehat{\mathbf{\Sigma}} = \widehat{\mathbf{S}}_n$.

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 $\widehat{\mathbf{\Sigma}}$

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

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

$$\mathbf{C} = \sigma_1\ \left(\mathbf{\Sigma}^* \otimes \mathbf{\Sigma}\right) + \sigma_2\ \mathrm{vec}\left(\mathbf{\Sigma}\right)\mathrm{vec}(\mathbf{\Sigma})^H,$$
$$\mathbf{P} = \sigma_1\ \left(\mathbf{\Sigma}^* \otimes \mathbf{\Sigma}\right)\ \mathbf{K}_{m^2,m^2} + \sigma_2\ \mathrm{vec}\left(\mathbf{\Sigma}\right)\mathrm{vec}(\mathbf{\Sigma})^T,$$

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

An important property of complex M-estimators

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

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

with

$$\begin{cases} \mathbf{C} = \nu_1\, \mathbf{\Sigma}^* \otimes \mathbf{\Sigma} + \nu_2\, \mathrm{vec}(\mathbf{\Sigma})\, \mathrm{vec}(\mathbf{\Sigma})^H, \\ \mathbf{P} = \nu_1\, \left(\mathbf{\Sigma}^* \otimes \mathbf{\Sigma}\right)\, \mathbf{K}_{m^2,m^2} + \nu_2\, \mathrm{vec}\left(\mathbf{\Sigma}\right) \mathrm{vec}\left(\mathbf{\Sigma}\right)^T, \end{cases}$$

where ν_1 and ν_2 are any real numbers.

e.g.

	SCM	M-estimators	FPE
ν_1	1	σ_1	$(m+1)/m$
ν_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 more secondary data ($\sigma_1 = (m+1)/m$ times more for Tyler)**: It implies that SCM can be simply replaced by any M-estimate in previous detectors without changing performance in Gaussian case (finite distance).

An important property of Tyler estimator

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

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

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

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

$$\begin{aligned}
\mathbf{C}_{FP} &= \frac{1}{m}\left(\mathbf{\Sigma}^T \otimes \mathbf{\Sigma}\right) + \frac{m-1}{m^2} \operatorname{vec}\left(\mathbf{\Sigma}\right) \operatorname{vec}\left(\mathbf{\Sigma}\right)^H, \\
\mathbf{P}_{FP} &= \frac{1}{m}\left(\mathbf{\Sigma}^T \otimes \mathbf{\Sigma}\right) \mathbf{K}_{m^2,m^2} + \frac{m-1}{m^2} \operatorname{vec}\left(\mathbf{\Sigma}\right) \operatorname{vec}\left(\mathbf{\Sigma}\right)^H
\end{aligned}$$

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

An important property of complex M-estimators

- Let $H(.)$ 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 $H(\widehat{\mathbf{\Sigma}})$)

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

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

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

where $H'\left(\mathbf{\Sigma}\right) = \left(\dfrac{\partial H(\mathbf{\Sigma})}{\partial \mathrm{vec}(\mathbf{\Sigma})}\right)$.

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

Outline

CES distribution $\Rightarrow$ two-step GLRT ANMF

Adaptive Normalized Matched Filter detector

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

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

- The ANMF is **scale-invariant (homogeneous of degree 0)**, i.e. $\forall \alpha, \beta \in \mathbb{R}\,,\ \Lambda_{ANMF}(\alpha\,\mathbf{z}, \beta\,\widehat{\mathbf{\Sigma}}) = \Lambda_{ANMF}(\mathbf{z}, \widehat{\mathbf{\Sigma}})$.
- Its **asymptotic distribution** (conditionally to $\mathbf{z}$!) is known *[Pascal 15, Ovarlez 15]*.

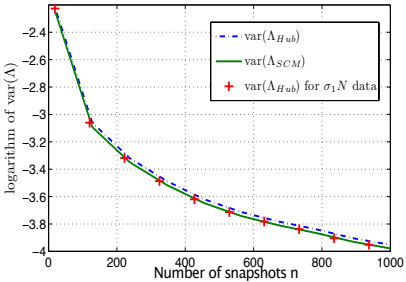
$$\sqrt{n}\left(H\left(\widehat{\mathbf{\Sigma}}\right) - H(\mathbf{\Sigma})\right) \xrightarrow{d} \mathcal{CN}\left(0, 2\,\nu_1\,H(\mathbf{\Sigma})\,\left(H\left(\mathbf{\Sigma}\right) - 1\right)^2\right) .$$

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

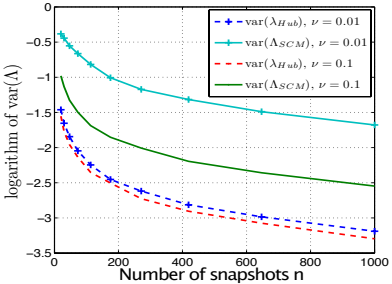
- **It is CFAR w.r.t the covariance/scatter matrix,**
- **It is CFAR w.r.t the texture.**

Illustrations of the Result on the ANMF

- $\Lambda = \mathrm{var}\left(H\left(\widehat{\mathbf{\Sigma}}\right) - H(\mathbf{\Sigma})\right)$. Here $\widehat{\mathbf{\Sigma}}$ = 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 Gaussian case but are clearly better in non-Gaussian case.

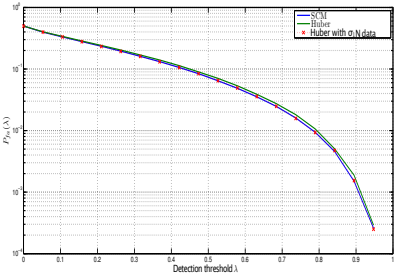
Illustrations of the Result on P_{fa}

- Figure 1: Gaussian context :

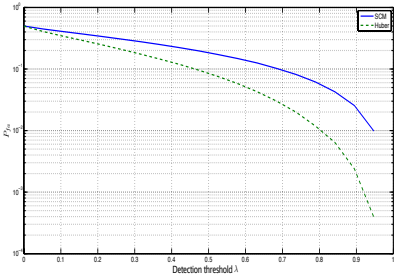
$$P_{fa} = (1 - \lambda_{ANMF})^{n-m+1} \; {}_2F_1\left(n - m + 2, n - m + 1; n + 1; \lambda_{ANMF}\right) .$$

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

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



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)

- 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.

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

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

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

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

The MUSIC statistic is

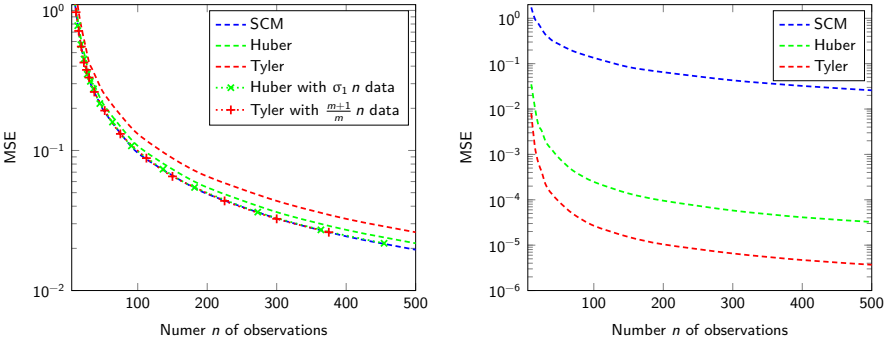
$$\begin{cases} H(\boldsymbol{\Sigma}) = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\, \gamma(\boldsymbol{\theta}) & \text{where } \gamma(\boldsymbol{\theta}) = \mathbf{s}(\boldsymbol{\theta})^H \mathbf{E}_W \mathbf{E}_W^H \mathbf{s}(\boldsymbol{\theta}), \\ H(\widehat{\boldsymbol{\Sigma}})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

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

This function respects assumptions of theorem 3!

The Mean Square Error (MSE) between the estimated angle $\widehat{\boldsymbol{\theta}}$ and the real angles $\boldsymbol{\theta}$ can then 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.



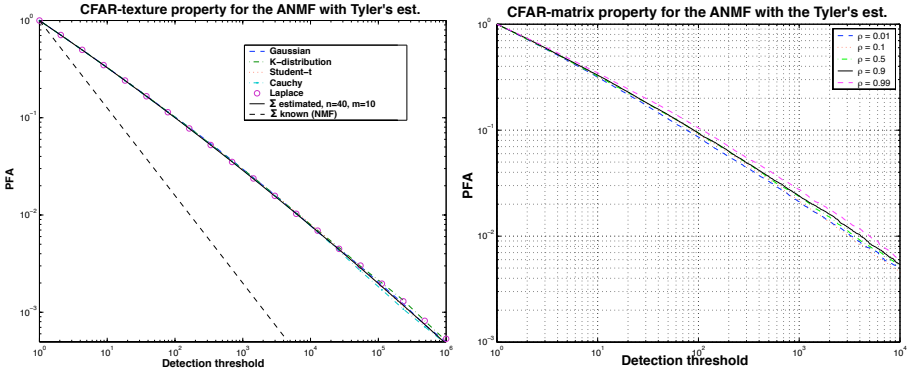
(a) White additive Gaussian noise (b) K-distributed additive noise ($\nu = 0.1$)

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

Similar conclusions as for detection can be drawn...

Illustration of the ANMF CFAR Properties For CES Noise

False Alarm regulation for ANMF built with Tyler's estimate



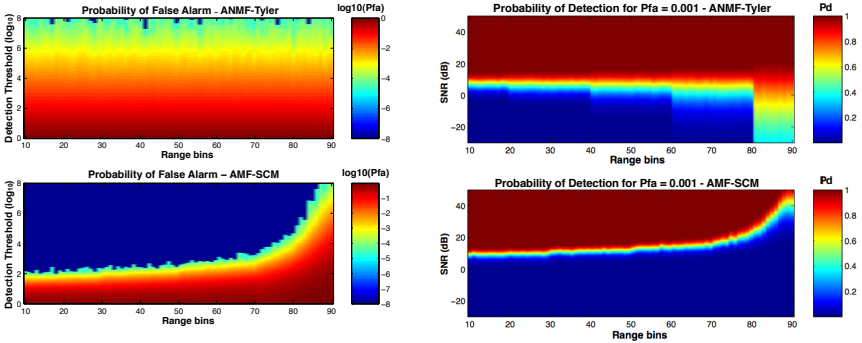
(a) CFAR-texture (b) CFAR-matrix

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

Properties of ANMF-Tyler Detector on Clutter Transitions

- K-distributed clutter transitions: from Gaussian to impulsive noise,
- Estimation of the covariance matrix with secondary data in 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

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:

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

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

- Tyler (or FP) Covariance Matrix case:

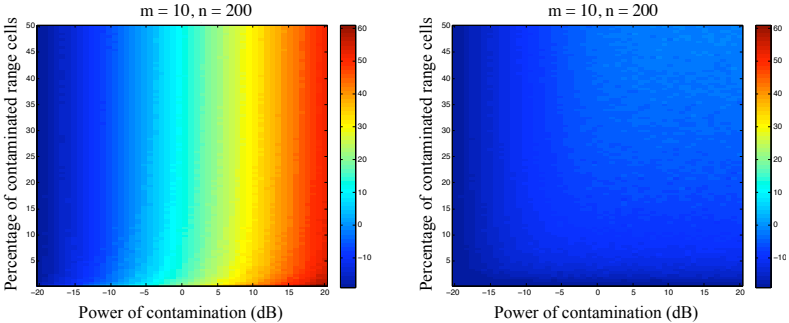
$$\widehat{\mathbf{\Sigma}}_{FPEPol} = \frac{m}{n}\sum_{k=1}^{n} \frac{\mathbf{y}_k\,\mathbf{y}_k^H}{\mathbf{y}_k^H\,\widehat{\mathbf{\Sigma}}_{FPEPol}^{-1}\,\mathbf{y}_k}\,, \qquad E\left[\widehat{\mathbf{\Sigma}}_{FPEPol}\right] = \mathbf{\Sigma} + \frac{m+1}{n}\left[E\left[\frac{\mathbf{p}_0\,\mathbf{p}_0^H}{\mathbf{p}_0^H\,\mathbf{\Sigma}^{-1}\,\mathbf{p}_0}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

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

Robustness of M-estimators

Gaussian vectors $\mathbf{y}_k$ polluted by outliers

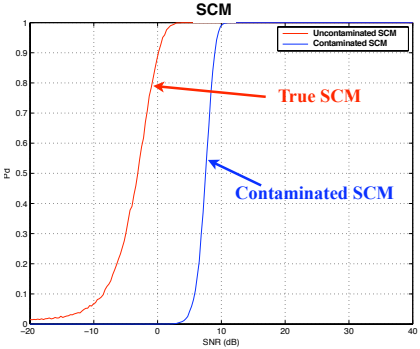
$$\hat{\mathbf{S}}_n = \frac{1}{n}\sum_{k=1}^{n} \mathbf{y}_k\,\mathbf{y}_k^H\,, \qquad \widehat{\mathbf{\Sigma}}_{FPE} = \frac{m}{n}\sum_{k=1}^{n} \frac{\mathbf{y}_k\,\mathbf{y}_k^H}{\mathbf{y}_k^H\,\widehat{\mathbf{\Sigma}}_{FPE}^{-1}\,\mathbf{y}_k}\,.$$



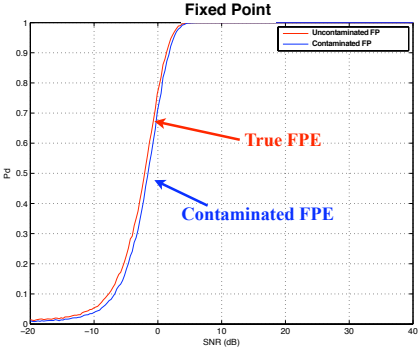
Plot of the error between the covariance matrix estimated with and without ouliers.

Robustness of ANMF: Impact on detection performance

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



AMF + SCM



ANMF + FPE

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

Outline

Motivations

The estimation of $\mathbf{\Sigma}$ does not take into account any prior knowledge on the covariance matrix:

How to improve detection performance by exploiting prior information on $\mathbf{\Sigma}$?

$\Longrightarrow$ Use of some prior knowledge on the structure of the covariance matrix:

- 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

Outline

Using Persymmetry Property

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

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

then:

- $\mathbf{s} = \mathbf{T}\,\mathbf{p}$ is a real vector (if $\mathbf{p}$ is centrosymmetric, i.e. $\mathbf{p} = \mathbf{J}_m \, \mathbf{p}^*$),
- $\mathbf{R} = \mathbf{T}\,\mathbf{\Sigma}\,\mathbf{T}^H$ 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
$\begin{cases} H_0 : \mathbf{y} = \mathbf{c}, & \mathbf{c}_1, \ldots, \mathbf{c}_n \\ H_1 : \mathbf{y} = A\,\mathbf{p} + \mathbf{c}, & \mathbf{c}_1, \ldots, \mathbf{c}_n \end{cases}$	$\longrightarrow$	$\begin{cases} H_0 : \mathbf{z} = \mathbf{n}, & \mathbf{n}_1, \ldots, \mathbf{n}_n \\ H_1 : \mathbf{z} = A\,\mathbf{s} + \mathbf{n}, & \mathbf{n}_1, \ldots, \mathbf{n}_n \end{cases}$

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 though $\mathbf{T}$ onto a real covariance matrix $\mathbf{R}$.

The Persymmetric FP Covariance Matrix Estimate

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

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

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

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

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

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

The Persymmetric Adaptive Normalized Matched Filter

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

$$\Lambda\left(\widehat{\mathbf{R}}_{PFP}\right) = \frac{\left|\mathbf{s}^\top \widehat{\mathbf{R}}_{PFP}^{-1} \mathbf{z}\right|^2}{\left(\mathbf{s}^\top \widehat{\mathbf{R}}_{PFP}^{-1} \mathbf{s}\right)\left(\mathbf{z}^H \widehat{\mathbf{R}}_{PFP}^{-1} \mathbf{z}\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda. \tag{6}$$

Properties:

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

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

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

Outline

Conventional Low Rank Detectors

Principle of Low Rank Matched Filter approaches found for example in *[Kirsteins 94]* (Principal Component Inverse) and *[Haimovich96]* (Eigencanceler) and *[Rangaswamy 04]*.

Let suppose that the rank r of clutter covariance matrix $\mathbf{\Sigma}$ is known:

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

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

$$\widehat{\mathbf{S}}_n = \frac{1}{n}\sum_{k=1}^{n} \mathbf{y}_k\,\mathbf{y}_k^H = (\mathbf{U_r}\,\mathbf{U_0}) \begin{pmatrix} \mathbf{\Sigma}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_0 \end{pmatrix} (\mathbf{U_r}\,\mathbf{U_0})^H,$$

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

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

Extended Low Rank Detectors

In case of heterogeneous and non-Gaussian clutter, we know that $\widehat{\mathbf{S}}_{SCM}$ 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:

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

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

More recent works can be found in *[Breloy 15, Sun 16, Breloy 16, Ginolhac 16]*.

Outline

Shrinkage of Tyler's estimators

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

Chen estimator *[Chen 11]*

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

subject to the constraint $\mathrm{Tr}(\mathbf{\Sigma}_C) = m$ 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]*

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

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

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

Outline

Some RMT-based results for detection schemes

The RMT (ex: *[Couillet 11]*) allows 1) to understand the statistical behaviour 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. At a finite distance (practical m, N values), the corrected results are often valid.

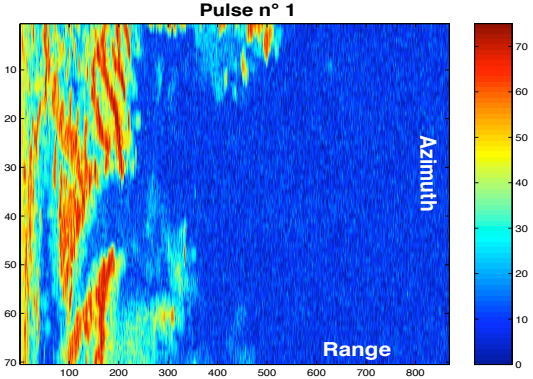
- Sources localisation applications *[Couillet 15]*: the based-RMT Music algorithm (G-Music) is known to have higher performance than those of conventional algorithms when using all the eigenvalues of the covariance matrix.
- MIMO-STAP *[Combernoux 16]*
- Adaptive Radar Detection when secondary data are correlated *[Couillet 15]*.
- Hyperspectral Anomaly Detection - Unmixing: the goal of E. Terreaux PhD thesis *[Terreaux 17]* 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.

Outline

False Alarm Regulation on THALES Ground Clutter

Data Description

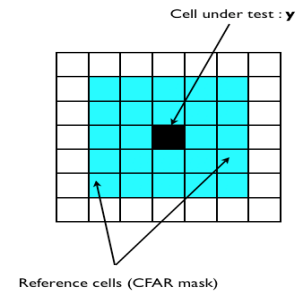
- "Range-azimuth" map from ground clutter data collected by a radar from THALES Air Defense, placed 13 meters above ground and illuminating area at low grazing angle.
- Ground clutter complex echoes collected in 868 range bins for 70 different azimuth angles and for $m = 8$ pulses.



False Alarm Regulation on THALES Ground Clutter

Data processing

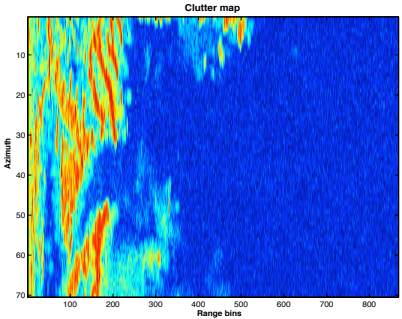
- Rectangular CFAR mask 5×5 for $0 \leq k \leq m$ different steering vectors $\mathbf{p}_k$.



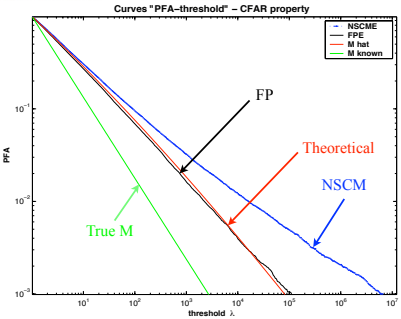
$$\mathbf{p}_k = \begin{pmatrix} 1 \\ \exp\left(\frac{2i\pi(k-1)}{m}\right) \\ \exp\left(\frac{2i\pi(k-1)2}{m}\right) \\ \vdots \\ \exp\left(\frac{2i\pi(k-1)(m-1)}{m}\right) \end{pmatrix}$$

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

False Alarm Regulation on THALES Ground Clutter



Azimut/range bins map

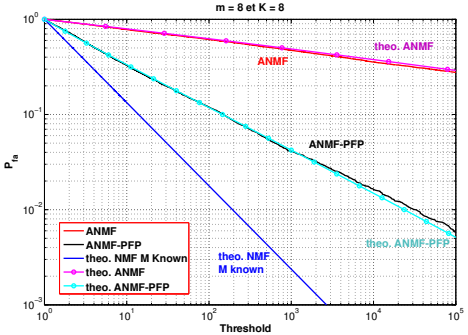


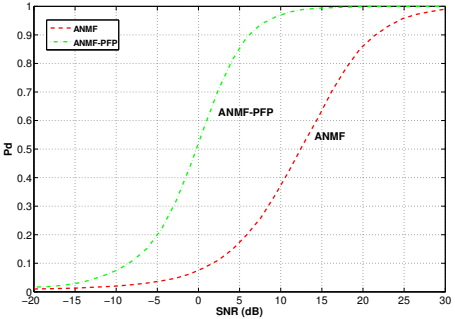
Relationship "P_{fa}-threshold"

Figure: ANMF with Tyler's M-estimate - False alarm regulation for $\mathbf{p}_0 = (1 \ldots 1)^T$.

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

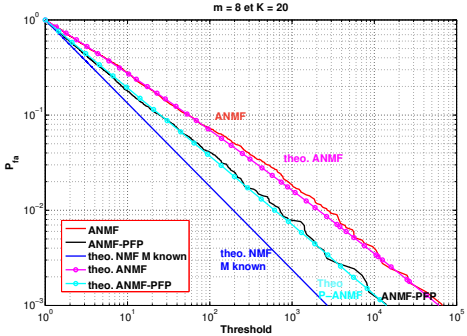
False Alarm Regulation on THALES Ground Clutter

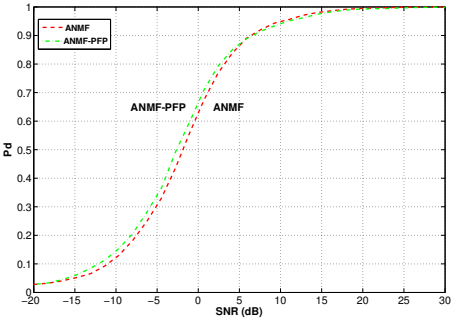




Persymmetric Tyler-ANMF and Tyler ANMF on THALES dataset - $m = 8$, $n = 8$

False Alarm Regulation on THALES Ground Clutter

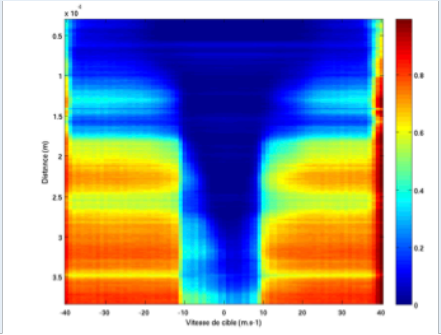




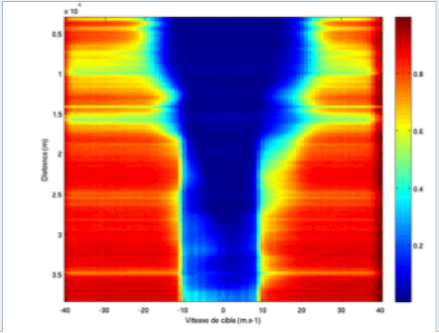
Persymmetric Tyler-ANMF and Tyler ANMF on THALES dataset - $m = 8$, $n = 20$

Detection Performance on THALES Sea Clutter

Non-Stationary and Heterogeneous THALES Sea clutter



X Detector on Dieppe sea clutter



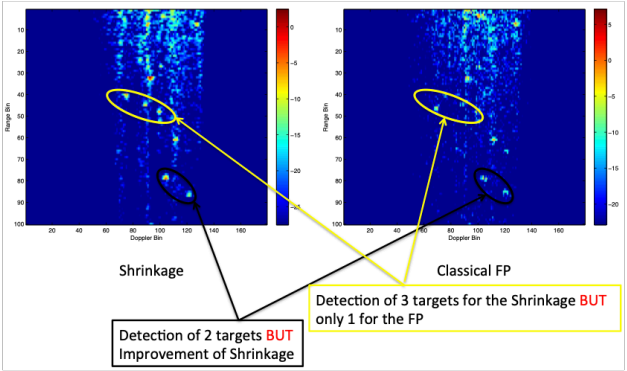
ANMF-FP Detector on Dieppe sea clutter

Application of Shrinkage to HFSWR: Detection

Singapore NTU HFSWR dataset

Shrinkage FP Estimator

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



[1] F. Pascal, Y. Chitour and Y. Quek, *Generalized Robust Shrinkage Estimator and Its Application to STAP Detection Problem*, Signal Processing, IEEE Transactions on, Vol.62, No.21, pp.5640 - 5651, 2014

Outline

Detection Performance on STAP Data

Problem

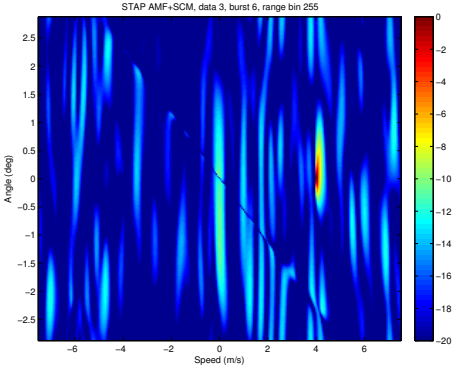
Using joint spatial and time measurements, estimate the position (angle) and the Doppler frequency (speed) of the target
$\Rightarrow$ use of the ANMF with a particular steering vector *[Ovarlez 2011]*

Data parameters: experimental clutter with synthetic target

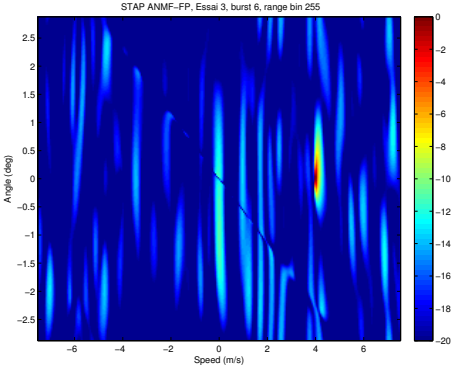
X-Band $\simeq 10^9$ Hz, wavelength $\lambda = 0.03$m, flight speed v =100m/s, distance to the scene 30km, 5 deg of incidence, PRF (Pulse Repetition Frequency) of 1 kHz, inter-sensor distance $d = 0.3$m, 12 trials with $n = 410$ range bins, $M = 64$ pulses and $N = 4$ sensors.

- 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



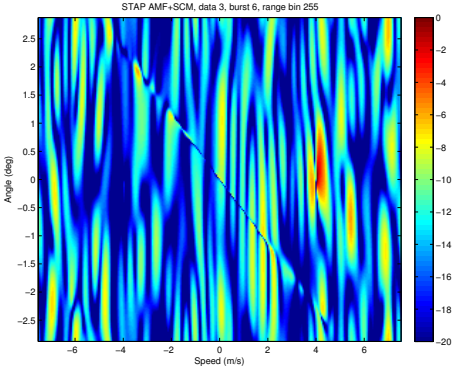
AMF detector with the SCM



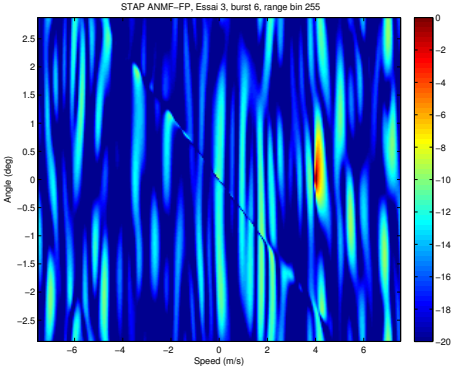
ANMF detector with Tyler's estimate.

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

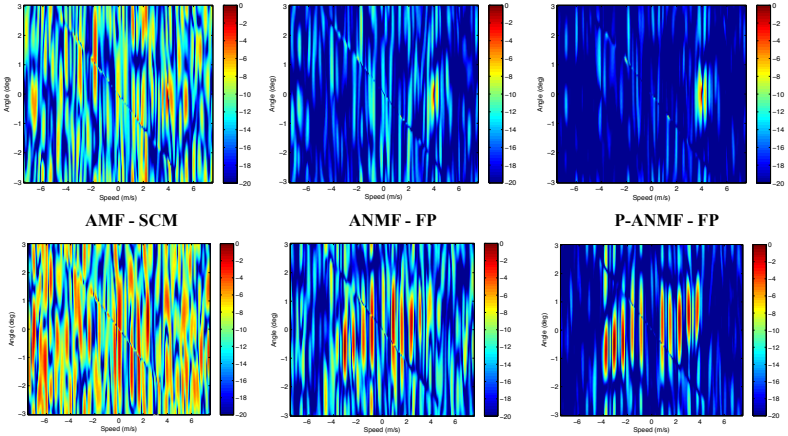


Figure: Doppler-angle map for the range bin 255 with $n = 404$ secondary data, $m = 256$ *[Pailloux 10]*.

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

- **Only one target detection**
- **Non contaminated secondary data**

$$N = 4,\ M = 64,\ n = 408$$

$$n < 2\,M\,N\,,\quad n > 2\,r$$

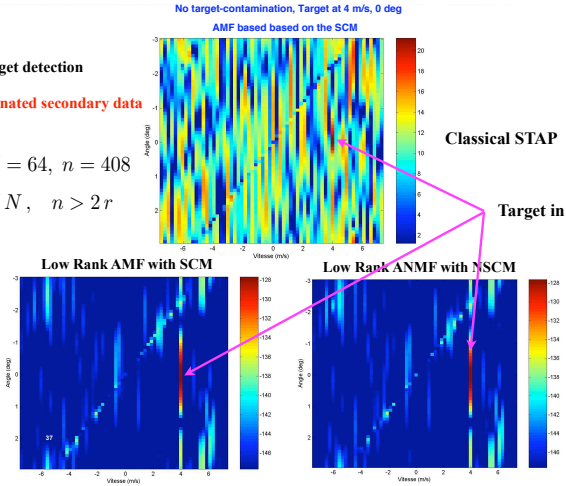


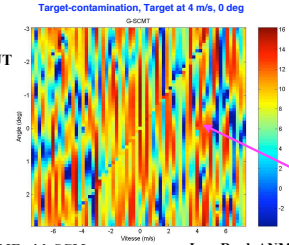
Figure: Doppler-angle map for the range bin 255 with $n = 100 < m$ secondary data and $m = 256$. (guard cells are removed)

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

- **Only one target (4m/s) in the CUT**
- **Contaminated secondary data (two targets at 4m/s and -4m/s)**

$$N = 4,\ M = 64,\ n = 410$$

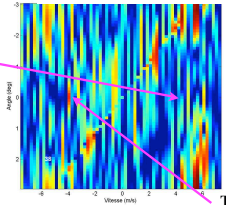
$$n < 2\,M\,N\,,\quad n > 2\,r$$



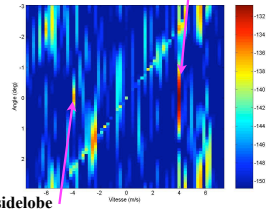
Classical STAP

Target in the CUT

Low Rank AMF with SCM



Low Rank ANMF with NSCM



Whitened target

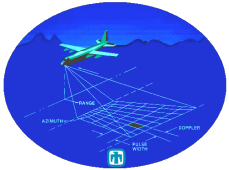
Target sidelobe

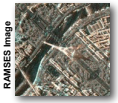
Figure: Doppler-angle map for the range bin 255 with $n = 100 < m$ secondary data and $m = 256$. (guard cells are removed)

Outline

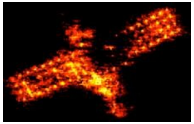
Background on SAR and Radar Imaging

ONERA RAMSES Image





RAMSES Image



ONERA ISAR Image

ONERA RAMSES Image

Radar Imaging allows to build more and more precise images:

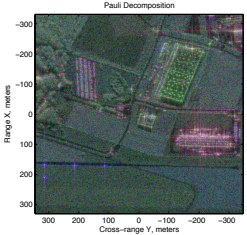
- Current use of very high spectral bandwidth and very high angular bandwidth leading to very high spatial resolution,
- Application to monitoring (detection, change detection), classification, 3D reconstruction, EM analysis, etc.

These applications require some physical diversity to reach good performances.

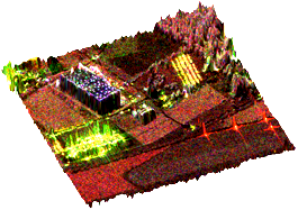
Multi-Channel SAR Images

Multi-channel SAR images automatically propose this diversity through:

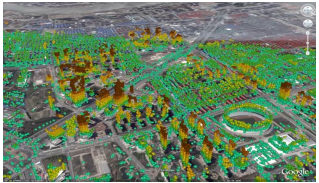
- polarimetric channels (POLSAR), interferometric channels (INSAR), polarimetric and interferometric channels (POLINSAR),
- multi-temporal, multi-passes SAR Image, etc.



EM behavior of the terrain in POLSAR images



Estimation of the height in POLINSAR images



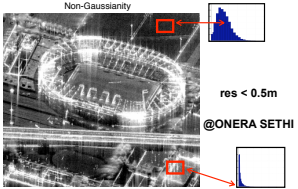
Analysis of the structures displacement in Shangai with multi-temporal SAR images (@Telespazio)

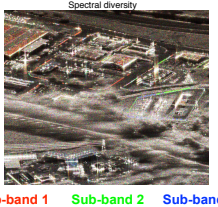
Almost all the conventional techniques of detection, parameters estimation, speckle filtering techniques, classification in multi-channel SAR images (e.g. polarimetric covariance matrix, interferometric coherency matrix) are based on the **multivariate statistic**.

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 directly 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**.





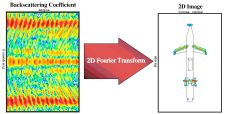
Challenging Problems

- 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 simply related to the SAR image $I(\mathbf{x})$ through:

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

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

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

With this model, all information relative to frequency f and angle θ are lost. 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.

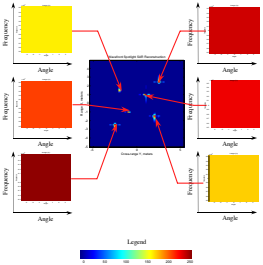
LTFD analysis and the physical group theory (Heisenberg or affine group) allow to construct **hyperimages** *[Bertrand 91, Bertrand 94, Bertrand 96]* through:

$$\tilde{I}(\mathbf{r}_0, \mathbf{k}_0) = < H(.), \Psi_{\mathbf{r}_0,\mathbf{k}_0}(.) > = \int_{\mathcal{D}_\mathbf{k}} H(\mathbf{k})\, \Psi^*_{\mathbf{r}_0,\mathbf{k}_0}(\mathbf{k})\, d\mathbf{k}\,,$$

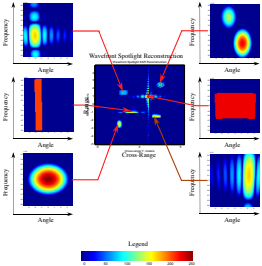
where $\Psi_{\mathbf{r}_0,\mathbf{k}_0}(\mathbf{k})$ is a family of wavelet bases (Gabor, wavelet) generated from a mother wavelet $\phi(f,\theta)$ through the chosen physical group of transformation (translations, scale in frequency, etc.) and where $\mathcal{D}_\mathbf{k}$ is the spectral/angular support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

Highlighting the Spectral and Angular Behaviors of Scatterers

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



Isotropic and white scatterers.

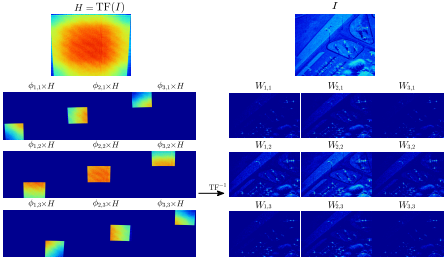


Anisotropic and colored scatterers.

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

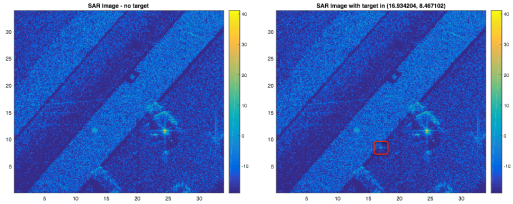
Each pixel i of the mono-channel SAR image can now be characterized by a N-vector $\mathbf{x}_i = \left[W^i_{1,1}, \ldots, W^i_{N_f,N_\theta}\right]^T$ of information ($N = N_f\, N_\theta$) related to **dispersion** in frequency domain and **anisotropy** in angular domain. Which multivariate statistic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

Detection Performance on SAR Image

Analysis of Performance

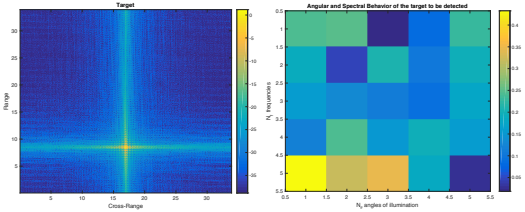
- Evaluation the CFAR property of the AMF and ANMF detectors,
- Comparison of the target detection performance between AMF and ANMF.

Dataset from SANDIA National Laboratories



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

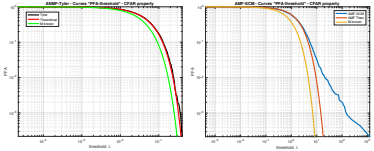
Artificial embedded target



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).

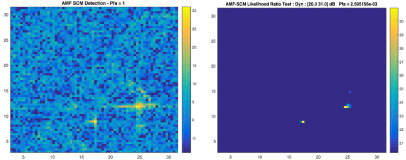
Detection Performance on SAR Image

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

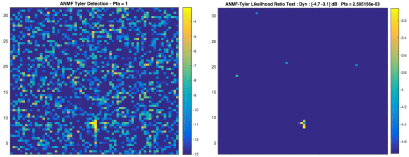


Left: FA Regulation with ANMF-Tyler. Right: FA Regulation with AMF-SCM. $N_\theta = 5$, $N_f = 5$, $K = 88$.

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



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



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

Change Detection on SAR Image Time Series

Contributions to SAR Image Time Series Analysis

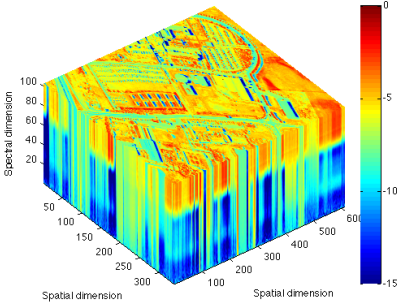
Guillaume Ginolhac and Arnaud Breloy

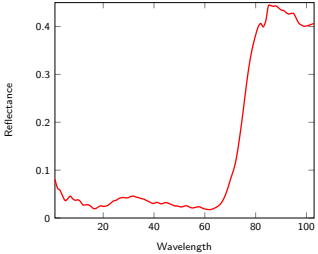
RadarConf Tutorial - 26 September 2020

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

Outline

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 **p** - Regulation of False Alarm *[Ovarlez 11, Frontera 17]*.

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

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

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

Joint estimation of location and scale *[Bilodeau 08]*

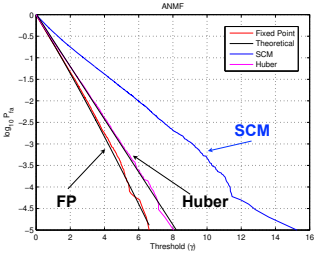
Hyperspectral Imaging

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

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

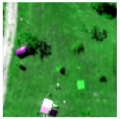
$$P_{fa} = (1-\lambda)^{\frac{n-1}{\sigma_1} - m + 1} \, {}_2F_1\left(\frac{n-1}{\sigma_1} - m + 2, \frac{n-1}{\sigma_1} - m + 1; \frac{n-1}{\sigma_1} - 1; \lambda\right), \text{ where } \sigma_1 = (m+1)/m\,.$$

- 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 reaches the same performance than those of the detector built with the SCM estimate.



Hyperspectral Imaging

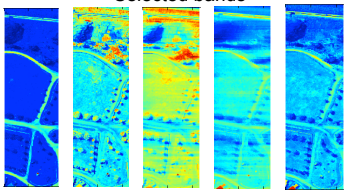
Original data set

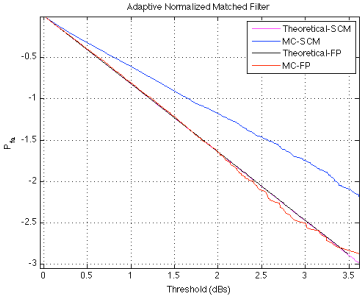


Extracted region :

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

Selected bands

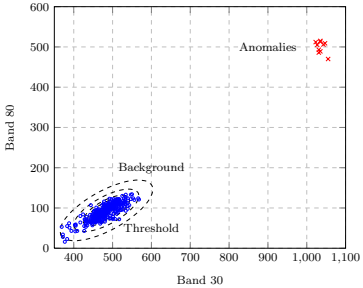




Hyperspectral Imaging (HSI)

GLRT RX Anomaly Detector: Mahalanobis Distance *[Reed 90]*

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



denoting $\widehat{\boldsymbol{\mu}} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} c_i$

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

(Hotelling T^2 distributed)

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

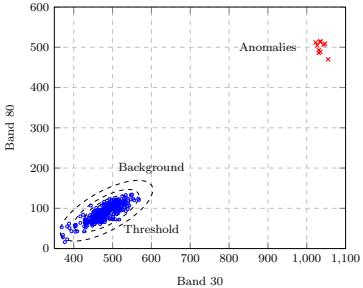
- Derived and valid only under Gaussian hypotheses,
- Its false alarm rate is independent of the covariance matrix: CFAR-matrix property in homogeneous Gaussian data.

Hyperspectral Imaging

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

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

A known and $\mathbf{p}$ unknown

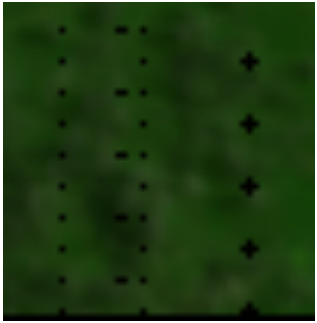


$$RXD_{M-est}(\mathbf{c}) = \left(\mathbf{c} - \widehat{\boldsymbol{\mu}}\right)^H \widehat{\boldsymbol{\Sigma}}^{-1} \left(\mathbf{c} - \widehat{\boldsymbol{\mu}}\right) \underset{H_0}{\overset{H_1}{\gtrless}} \lambda$$

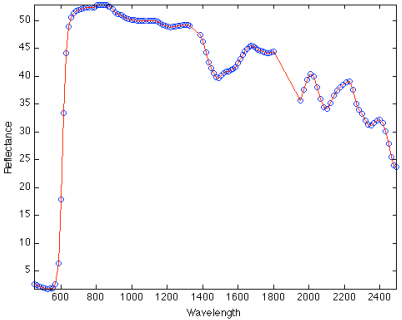
where $\widehat{\boldsymbol{\Sigma}}$ and $\widehat{\boldsymbol{\mu}}$ are M-estimates of the location and scale

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

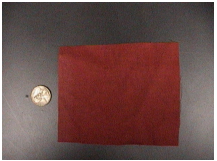
Anomaly Detection Results on Artificial Targets



Original image (Forest Region)

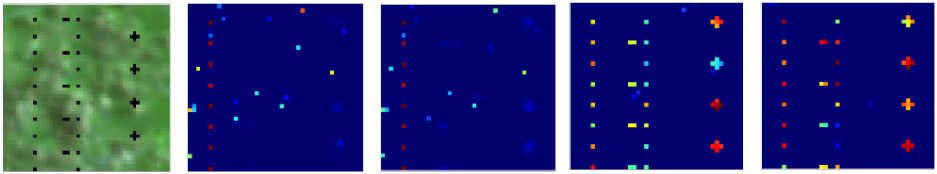


Target Spectrum



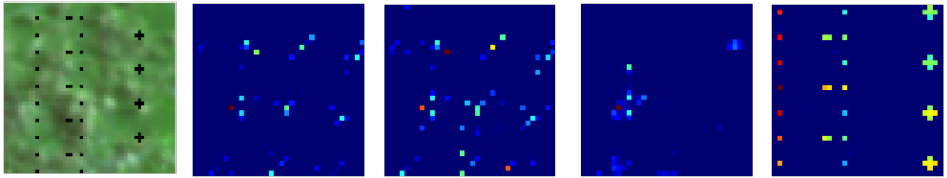
50 x 50 pixels, 126 spectral bands

Anomaly Detection Results on Artificial Targets



(a) Original (b) SCM (c) Shr-SCM (a) FP (b) Shr-FP

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



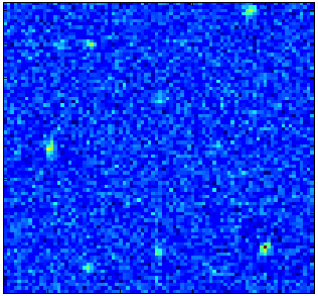
(a) Original (b) SCM (c) Shr-SCM (a) FP (b) Shr-FP

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

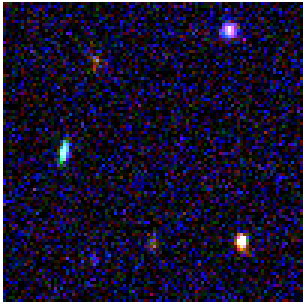
Problem of detecting galaxies in HS MUSE (Multi Unit Spectroscopic Explorer) data (465-930 nm)

Classical RXD



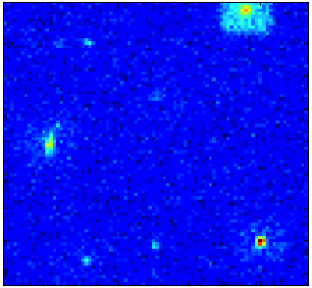
$RXD_{SCM}(\mathbf{c})$

Muse Image



300 x 300 pixels
3578 spectral bands

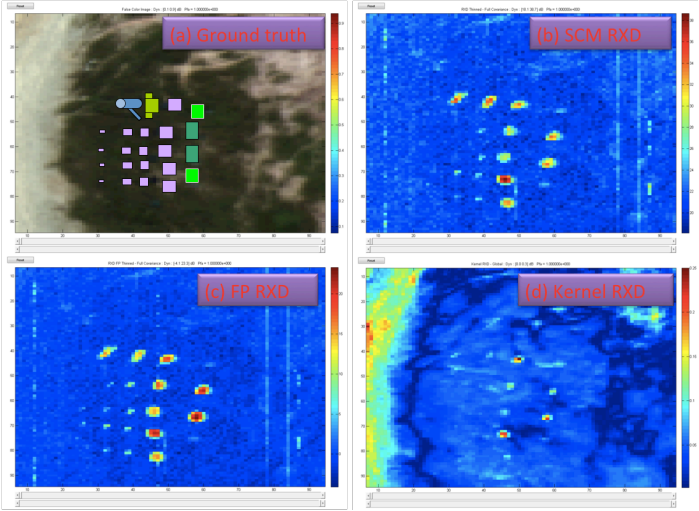
Extended RXD



$RXD_{Tyler}(\mathbf{c})$

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 to take into account the background complexity: the non-Gaussianity, the temporal background fluctuations and the spatial background power fluctuations,
- Using this model, the ANMF detector built with the Fixed Point (or other **M**-estimators) background covariance matrix estimator is shown to be CFAR-texture, CFAR-matrix and exhibits nice properties (robustness) and very good detection performance,
- Taking into account additional *a priori* properties on the covariance matrix structure (low rank, persymmetry, Toeplitz, ...) can lead to a appreciable gain for small numbers of secondary data,
- These methods have been applied for many problems involving covariance matrix estimation: STAP detection, SAR detection (FOPEN), Polarimetric/Interferometric SAR detection and classification, SAR and Hyperspectral Change Detection, SAR and Hyperspectral time-series analysis, Financial Portfolio Oprimization, 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, ...*], [*Q. Hoarau, G. Ginolhac*],
- Sparsity and high dimension: [A. Bitar, J.-P. Ovarlez].
- Performance of Estimation : [A. Renaux, B. Mériaux, S. Fortunati]

Bibliography

- 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.
- G. Frahm, *Generalized elliptical distributions: theory and applications*, PhD thesis, Universitätsbibliothek, 2004.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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), 2012.
- 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.
- M. Mahot. Estimation robuste de la matrice de covariance en traitement du signal. PhD thesis, École Normale Supérieure de Cachan, France, 2012.
- K. Yao, A representation theorem and its application to spherically invariant random processes, *Information Theory, IEEE Transactions on*, **19**(2), 1973.
- 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

- J. T. Kent and D. E. Tyler, Redescending M-estimates of multivariate location and scatter *Annals of Statistics*, **19**(4), pp.2102-2119, 1991.
- 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.
- F. Pascal, *Détection et Estimation en Environnement Non-Gaussien*, PhD thesis, University of Nanterre, France, 2006.
- 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.
- 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. Terreaux, *Robust model order selection using Random Matrix Theory*, PhD Thesis, University Paris Saclay, 2018.
- P. J. Huber, Robust estimation of a location parameter, *The Annals of Mathematical Statistics*, **35**(1), 1964.
- F. Gini and M. Greco. Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter. *Signal Processing*, **82**(12), pp.1847-1859, 2002.
- 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, 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.
- 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.

Bibliography

- 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.
- 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.
- 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.
- 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.
- 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.
- R. O. Schmidt. Multiple emitter location and signal parameter estimation. *Acoustics Speech and Signal Processing, IEEE Transactions on*, **34**(3), pp.276-280, 1986.
- J. P. Burg, D. G. Luenberger and D. L. Wenger. Estimation of structured covariance matrices. *Proc. IEEE*, **70**(9), pp.963-974, 1982.
- 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.
- 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.
- A. Haimovich. The eigencanceler: adaptive radar by eigenanalysis methods. *Aerospace and Electronic Systems, IEEE Transactions on*, **32**(2), pp.532-542, 1996.

Bibliography

- L. Cai and H. Wang. A Persymmetric Multiband GLR Algorithm. *Aerospace and Electronic Systems, IEEE Transactions on*, pp. 806-816, 1992.
- A. De Maio. Maximum likelihood estimation of structured persymmetric covariance matrices. *Elsevier Signal Processing*, **83**(3), pp.633-640, 2003.
- 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.
- G. Pailloux. *Estimation Structurée de la Matrice de Covariance et Application à la Détection Radar*, PhD thesis, University of Nanterre, France, 2010.
- 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.
- 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.
- 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.
- 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.

Bibliography

- 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. Combernoux. Low rank detection and estimation using random matrix theory approaches for antenna array processing. PhD Thesis, University Paris-Saclay. 2016.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.

Bibliography

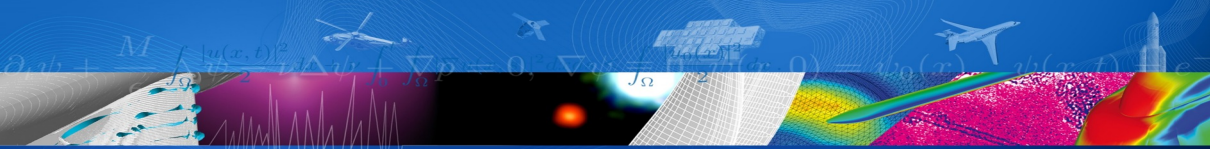
- Y. Sun, P. Babu and D. P. Palomar. Regularized Tyler's scatter estimator: Existence, uniqueness, and algorithms. *Signal Processing, IEEE Transactions on*, **62**(19), pp.5143-5156, 2014.
- Y. I. Abramovich and N. K. Spencer. Diagonally loaded normalised sample matrix inversion (Insmi) for outlier-resistant 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.
- 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.
- 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.
- E. Ollila and D. E. Tyler. Regularized *M*-estimators of scatter matrix. *Signal Processing, IEEE Transactions on*, **62**(22), pp.6059-6070, 2014.
- A. Wiesel. Geodesic convexity and covariance estimation. *Signal Processing, IEEE Transactions on*, **60**(12), pp.6182-6189, Dec. 2012.
- 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.
- 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

- 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.
- 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.
- M. Bilodeau and D. Brenner. *Theory of multivariate statistics*. Springer Science & Business Media, 2008.
- 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.
- 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.
- 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.
- 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, Dec. 2019.

Bibliography

- 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.
- 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.



Tutorial on Robust Estimation and Detection Schemes in non-Standard Conditions for Radar, Array Processing and Imaging

Jean-Philippe Ovarlez

ONERA, The French Aerospace Lab,
ElectroMagnetism and Radar Department,
Advanced Signal Processing Unit
& CentraleSupélec SONDRA Laboratory

Séminaire Recherche M2R ATSI - 11 janvier 2023

ONERA

THE FRENCH AEROSPACE LAB

Annex

Robust Model Order Selection Using Random Matrix Theory

Annex: Contents

Motivations

The Model Order Selection is a fundamental problem in Signal Processing:

- Radar, Sonar: Direction of Arrival, Source Localization, STAP, Date of Arrival, Spectral Analysis (ARMA), etc.
- Hyperspectral: Unmixing, etc.
- Finance: portfolio optimization, efficient portfolio composition, etc.

In spite of the multitude of techniques available for solving this problem, most of them use information theoretic approaches, such as:

- the Akaike Information Criterion (AIC),
- the Minimum Description Length (MDL).

Recently, the Random Matrix Theory (RMT) based-approach has also been proposed.

All these methods are classically based on white Gaussian noise assumption.

Contents

Outline

Problem formulation

Let $\{\mathbf{y}_i\}_{i\in[1,N]}$ be N observations of size m characterizing the $p < m$ mixed sources corrupted by additive noise:

$$\mathbf{y}_i = \sum_{j=1}^{p} s_{i,j}\,\mathbf{m}_j + \mathbf{n}_i\,, \qquad i \in [1, N]\,, \tag{1}$$

which can be rewritten more compactly as:

$$\mathbf{Y} = \mathbf{M}\,\mathbf{S} + \mathbf{N}\,, \tag{2}$$

where

- $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N] \in \mathbb{C}^{m\times N}$ are the observations,
- $\mathbf{M} \in \mathbb{C}^{m\times p}$ is the mixing matrix containing spectra of the p sources,
- $\mathbf{S} \in \mathbb{C}^{p\times N}$ is the channel gain matrix,
- $\mathbf{N} \in \mathbb{C}^{m\times N}$ is the additive noise matrix, independent of the source signal.

Generally, $p < m$ is unknown, M and S unknown and N is zero-mean white Gaussian noise.

Recall of Akaike Information Criterion and MDL

Goal: Given a set of N observations $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N]$ and a family of models, (e.g., a parametrized family of pdf $\{f(\mathbf{Y}|\mathbf{\Theta}_k)\}_k$), select the model k that best fits the data. *[Akaike 74]* proposal was to select the model which gives the minimum AIC defined as:

$$AIC(k) = -2 \log f\left(\mathbf{Y}|\hat{\mathbf{\Theta}}_k\right) + 2\,k\,,$$

where $\hat{\mathbf{\Theta}}$ is the MLE of vector $\mathbf{\Theta}_k$ and k is the number of free adjusted parameters in vector $\mathbf{\Theta}_k$.

[Schwartz 78] and *[Rissanen 78]* approaches yield the same type of criterion, given by:

$$MDL = -\log f\left(\mathbf{Y}|\hat{\mathbf{\Theta}}_k\right) + \frac{1}{2}\,k\,\log N\,.$$

Applications for Model Order Selection (1)

Let us consider the theoretical covariance matrix $\mathbf{R}$ of complex observations $\mathbf{y}_i$ with $\mathbf{n}_i$ white Gaussian noise:

$$\mathbf{R} = E\left[\mathbf{y}_i\,\mathbf{y}_i^H\right] = \mathbf{M}\,\mathbf{s}_i\,\mathbf{s}_i^H\,\mathbf{M} + \sigma^2\,\mathbf{I}_m = \mathbf{\Phi} + \sigma^2\,\mathbf{I}_m\,.$$

We assume here $\mathbf{M}$ full column rank ($\{\mathbf{m}_i\}_i$ linearly independent) and $\mathbf{s}_i\,\mathbf{s}_i^H$ non singular. We have:

- $\mathrm{rank}(\mathbf{\Phi}) = p$, the $m-p$ smallest eigenvalues of $\mathbf{\Phi}$ are equal to zero,
- $\mathrm{eig}(\mathbf{R}) = \left\{\lambda_1, \lambda_2, \ldots, \lambda_p, \sigma^2, \ldots, \sigma^2\right\}$.

We can define a family of covariance matrix $\mathbf{R}^{(k)} = \mathbf{\Phi}^{(k)} + \sigma^2\,\mathbf{I}_m$ as:

$$\text{Model (k):} \qquad \mathbf{R}^{(k)} = \sum_{i=1}^{k}\left(\lambda_i - \sigma^2\right)\,\mathbf{v}_i\,\mathbf{v}_i^H + \sigma^2\,\mathbf{I}_m\,,$$

where λ_1, ..., λ_k are the k highest eigenvalues of $\mathbf{R}^{(k)}$ and where $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are their corresponding eigenvectors. We can define also the vector $\mathbf{\Theta}_k$ of unknown parameters as:

$$\mathbf{\Theta}_k = \left(\lambda_1, \ldots, \lambda_k, \sigma^2, \mathbf{v}_1, \ldots, \mathbf{v}_k\right)^T\,.$$

Applications for Model Order Selection (2)

$$
\begin{aligned}
-\log f\left(\mathbf{Y}|\hat{\mathbf{\Theta}}_k\right) &= -\log \prod_{i=1}^{N} \frac{1}{\pi^m \left|\mathbf{R}^{(k)}\right|} \exp\left(-\mathbf{y}_i^H \left(\mathbf{R}^{(k)}\right)^{-1} \mathbf{y}_i\right), \\
&\approx N \log\left|\mathbf{R}^{(k)}\right| + \mathrm{Tr}\left[\left(\mathbf{R}^{(k)}\right)^{-1} \sum_{i=1}^{N} \mathbf{y}_i \mathbf{y}_i^H\right], \\
&\approx N \log\left|\mathbf{R}^{(k)}\right| + \mathrm{Tr}\left[\left(\mathbf{R}^{(k)}\right)^{-1} N \hat{\mathbf{R}}\right], \qquad (3)
\end{aligned}
$$

where $\hat{\mathbf{R}} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{y}_i \mathbf{y}_i^H$. Maximizing (3) with respect to each parameter of $\mathbf{\Theta}_k$ *[Anderson 63]* leads to:

- $\hat{\lambda}_i = l_i$ and $\hat{\mathbf{v}}_i = \mathbf{u}_i, i \in [1, k]$, $\hat{\sigma}^2 = \frac{1}{m-k}\sum_{i=k+1}^{m} l_i$.

where $l_1 > l_2 \ldots > l_m$ and $\mathbf{u}_1, \ldots, \mathbf{u}_m$ are the eigenvalues and eigenvectors of the Sample Covariance Matrix $\hat{\mathbf{R}}$.

Applications for Model Order Selection (3)

The number of free parameters is obtained by counting the number of degrees of freedom spanned by $\mathbf{\Theta}_k = \left(\lambda_1, \ldots, \lambda_k, \sigma^2, \mathbf{v}_1, \ldots, \mathbf{v}_k\right)^T$ *[Wax 85]*:

$$k+1 \text{ reals } +2\,k\,m \text{ reals} - 2\,k \text{ normalizations } -2\,k\,(k-1)/2 \text{ mutual orthogonalizations}$$

Substituting the Maximum Likelihood Estimates in the log-likelihood (3) leads to:

$$AIC(k) = -2\,N \log \frac{\prod_{i=k+1}^{m} \lambda_i}{\left(\frac{1}{m-k}\sum_{i=k+1}^{m} \lambda_i\right)^{m-k}} + 2\,k(2\,m-k)\,,$$

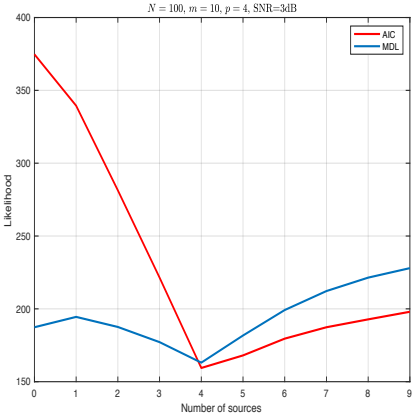
$$MDL(k) = -N \log \frac{\prod_{i=k+1}^{m} \lambda_i}{\left(\frac{1}{m-k}\sum_{i=k+1}^{m} \lambda_i\right)^{m-k}} + \frac{1}{2}\,k\,(2\,m-k)\,\log N\,.$$

Applications for Model Order Selection (4)

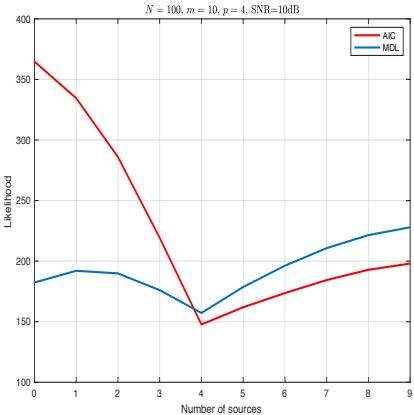
- AIC is shown to be not consistent and has a problem of over-estimation of the number of sources
- MDL is consistent and is generally prefered to AIC.
- Both techniques are based on white Gaussian noise. They do not work for correlated noise or non-Gaussian noise
- Both techniques may have some problems for high dimensional data

Examples (1)

SNR = 3dB



SNR = 10dB

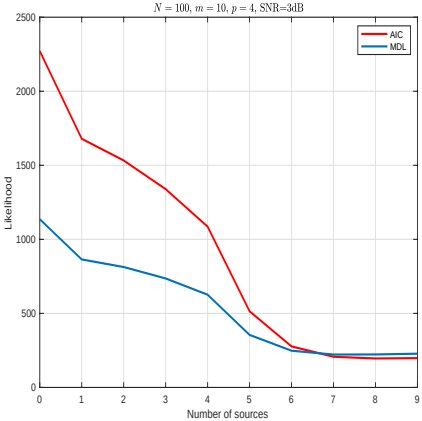


AIC and

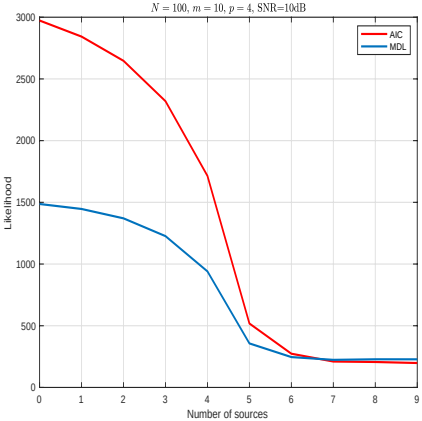
MDL model order selection (white Gaussian noise, $N = 100$, $m = 10$).

Examples (2)

SNR = 3dB



SNR = 10dB



AIC and

MDL model order selection (correlated Gaussian noise, $\rho = 0.9$, $N = 100$, $m = 10$).

Outline

RMT in SONDRA

The RMT *Couillet 11]* is not a magic tool or *scientific nutshell* but allows 1) to understand the statistical behaviour 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. At a finite distance (practical m, N values), the corrected results are often valid.

- Sources localisation applications [F. Pascal, R. Couillet, ...]: the based-RMT Music algorithm (G-Music) is known to have higher performance than those of conventional algorithms when using all the eigenvalues of the covariance matrix.
- MIMO-STAP: the goal of A. Combernoux PHD thesis *[Combernoux 16]* was to analyse/improve the detection and filtering performances of low-rank detectors.
- Adaptive Radar Detection: when secondary data are correlated *[Couillet 15]*
- Hyperspectral Anomaly Detection - Unmixing: the goal of E. Terreaux PhD thesis *[Terreaux 17]* 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.

Key ideas (1)

Let $\{\mathbf{y}_i\}_{i\in[1,N]}$ be distributed according to $\mathcal{CN}(\mathbf{0}_m, \mathbf{M})$. The Maximum Likelihood Estimate of $\mathbf{M}$ is given by $\widehat{\mathbf{M}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{y}_i\,\mathbf{y}_i^H = \frac{1}{N}\mathbf{Y}\,\mathbf{Y}^H$.

Asymptotic Regime

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

$$\widehat{\mathbf{M}} - \mathbf{M} \xrightarrow{a.s.} \mathbf{0}_m, \qquad \left\|\widehat{\mathbf{M}} - \mathbf{M}\right\| \xrightarrow{a.s.} 0\,.$$

Random Matrix Regime

- No longer valid if m, $N \to \infty$ with $m/N \to c \in [0, \infty[$: $\left\|\widehat{\mathbf{M}} - \mathbf{M}\right\| \nrightarrow 0$,
- For practical large m, N with $m \simeq N$, it can lead to dramatically wrong conclusions (even $m = N/100$).

Key ideas (2)

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

Random Matrix Regime

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

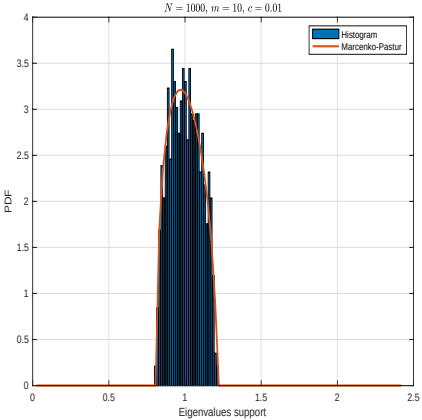
$$p(x) = \left(1 - \frac{1}{c}\right)_+ \delta(x) + \frac{1}{2\,\pi\,c\,x}\sqrt{(x-\lambda_-)_+\,(\lambda_+ - x)_+}$$

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

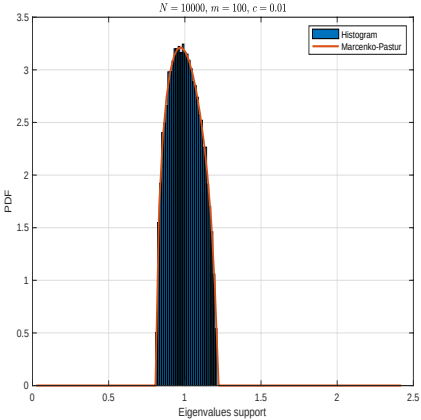
Not restricted to Gaussian statistics !

RMT Examples (1): classical asymptotic regime

$N = 1000,\ m = 10,\ c = 0.01$



$N = 10000,\ m = 100,\ c = 0.01$

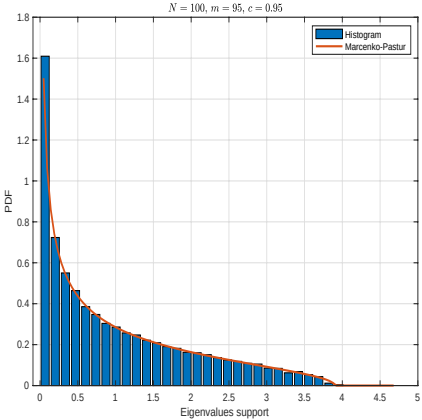


Eigenvalues

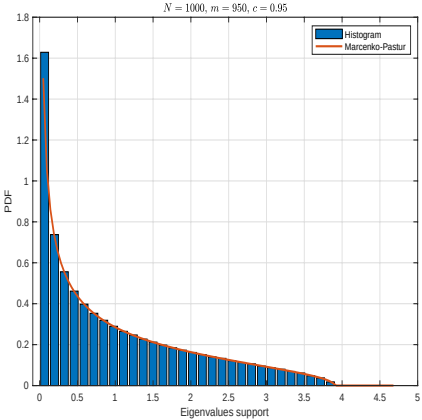
support for **white** Gaussian noise ($\sigma^2 = 1$, $\mathbf{C} = \sigma^2\,\mathbf{I}_m$).

RMT Examples (2): same RMT regime

$N = 100,\ m = 95,\ c = 0.95$



$N = 1000,\ m = 950,\ c = 0.95$

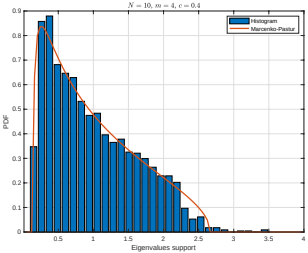


Eigenvalues

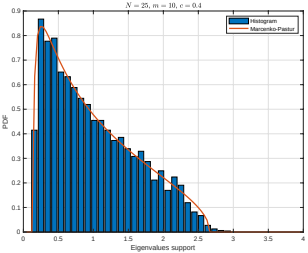
support for **white** Gaussian noise ($\sigma^2 = 1$, $\mathbf{C} = \sigma^2\,\mathbf{I}_m$).

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

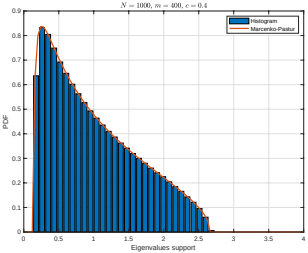
$N = 10,\ m = 4$



$N = 25,\ m = 10$



$N = 1000,\ m = 400$

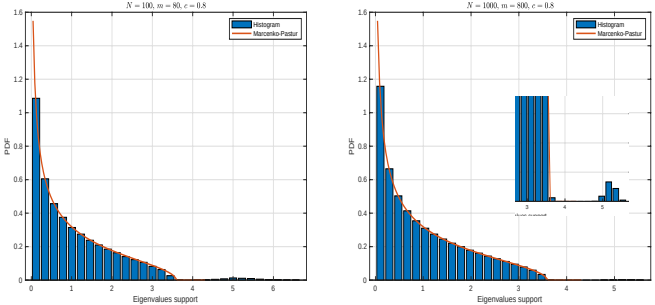


Eigenvalues support for **white** Gaussian noise ($\sigma^2 = 1$, $\mathbf{C} = \sigma^2\,\mathbf{I}_m$) and $c = 0.4$.

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 Marcenko-Pastur distribution.

$N = 100,\ m = 80,\ c = 0.8$ $\quad$ $N = 1000,\ m = 800,\ c = 0.8$



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

$$\lambda_{max}\left(\frac{1}{N}\mathbf{Y}\mathbf{Y}^H\right) \underset{N,m\to\infty}{\overset{a.s.}{\longrightarrow}} \sigma^2\left(1+\sqrt{c}\right)^2$$

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

$$\lambda_{max}\left(\frac{1}{N}\mathbf{Y}\mathbf{Y}^H\right) \underset{N,m\to\infty}{\overset{a.s.}{\longrightarrow}} \sigma^2\left(1+\theta\right)\left(1+\frac{c}{\theta}\right) \geq \sigma^2\left(1+\sqrt{c}\right)^2$$

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

Generalization: We consider N observations $\left\{ \mathbf{y}_k = \sum_{i=1}^{p} \sqrt{\theta_i}\, \mathbf{u}_i + \mathbf{n}_k \right\}_{k \in [1,N]}$ with $\|\mathbf{u}_i\| = 1$.

Let $\left\{ \hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \ldots \geq \hat{\lambda}_p \right\}$ be the p-largest eigenvalues of $\frac{1}{N} \mathbf{Y}\mathbf{Y}^H$

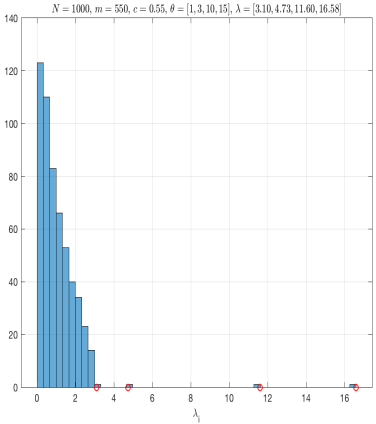
- if $\theta_i \geq \sigma^2 \sqrt{c}$, then

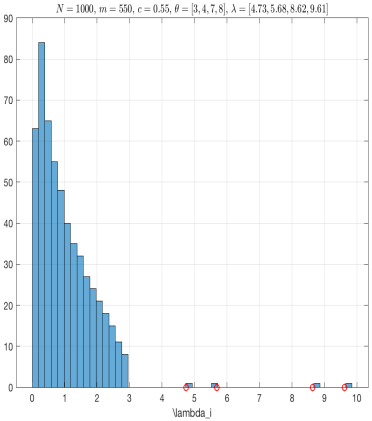
$$\hat{\lambda}_i \xrightarrow[N,m \to \infty]{a.s.} \sigma^2 \left(1 + \theta_i\right) \left(1 + \frac{c}{\theta_i}\right) \geq \sigma^2 \left(1 + \sqrt{c}\right)^2$$

$$\hat{\theta}_k = \frac{\hat{\lambda}_k - \sigma^2 (1 + c) + \sqrt{\left(\sigma^2 (1 + c) - \hat{\lambda}_k\right)^2 - 4\, c\, \sigma^4}}{2\, \sigma^2}$$

is a consistent estimator of θ_k, $k \in [1, p]$.

Source Detection with RMT





MUSIC Problem in large dimension

Let $\{\mathbf{y}_i\}_{i\in[1,N]}$ be N observations of size m characterizing the $p < m$ mixed sources corrupted by additive noise:

$$\mathbf{y}_i = \sum_{j=1}^{p} s_{i,j}\,\mathbf{m}_j + \sigma\,\mathbf{n}_i\,, \qquad i \in [1, N]\,,$$

which can be rewritten more compactly as:

$$\mathbf{Y} = \mathbf{M}\,\mathbf{S} + \sigma\,\mathbf{N}\,,$$

where

- $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N] \in \mathbb{C}^{m\times N}$ are the observations,
- $\mathbf{M} \in \mathbb{C}^{m\times p}$ is the mixing matrix containing spectra (steering vector) of the p sources,
- $\mathbf{S} \in \mathbb{C}^{p\times N}$ is the channel gain matrix,
- $\mathbf{N} \in \mathbb{C}^{m\times N}$ is the white Gaussian noise ($E\left[\mathbf{n}_i^H\,\mathbf{n}_i\right] = 1$), independent of the source signal.

Subspace estimation

Let $\mathbf{\Pi}_m$ the orthogonal projector on $\mathrm{span}(\mathbf{M})$, the column space of $\mathbf{M}$ and $\mathbf{I}_m - \mathbf{\Pi}_m$ the orthogonal projector on $\mathrm{span}(\mathbf{M})^{\perp}$.

Problem: Consistent estimation of $\mathbf{b}_m^H \left(\mathbf{I}_m - \mathbf{\Pi}_m\right) \mathbf{b}_m$ for any deterministic vector $\mathbf{b}_m$

Ex: MUSIC Algorithm

$$\mathbf{a}_m(\phi)^H \left(\mathbf{I}_m - \mathbf{\Pi}_N\right) \mathbf{a}_m(\phi) = 0 \Longleftrightarrow \phi \in \{\phi_1, \ldots, \phi_p\} \text{ directions of the } p \text{ sources.}$$

Since $\mathbf{C} = E\left[\mathbf{Y}\mathbf{Y}^H\right] = \mathbf{M}\,\mathrm{diag}\left(\theta_1, \ldots, \theta_p\right) \mathbf{M}^H + \sigma^2 \mathbf{I}_m$, $\mathbf{\Pi}_m$ is the orthogonal projector spanned by the eigenvectors $(\mathbf{u}_1, \ldots, \mathbf{u}_p)$ corresponding to the p-largest eigenvalues of $\mathbf{C}$.

Standard Estimator: $\hat{\mathbf{\Pi}}_m = \sum_{k=1}^{p} \hat{\mathbf{u}}_k \hat{\mathbf{u}}_k^H$, where $(\hat{\mathbf{u}}_1, \ldots, \hat{\mathbf{u}}_p)$ are the p-eigenvectors corresponding to the p-largest eigenvalues of $\frac{1}{N}\mathbf{Y}\mathbf{Y}^H$

Subspace estimation

In large dimension, $\hat{\mathbf{\Pi}}_m$ is no longer a good estimator of $\mathbf{\Pi}_m$! Let $\hat{\boldsymbol{\lambda}} = \left(\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_m\right)^T$ be the eigenvalues of $\frac{1}{N}\mathbf{Y}\mathbf{Y}^H$ and $\{\hat{\mathbf{u}}_1, \ldots, \hat{\mathbf{u}}_m\}$, their corresponding eigenvectors.

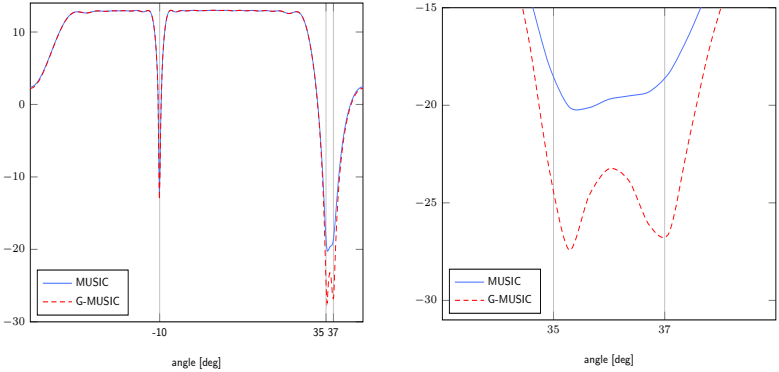
G-MUSIC *[Mestre 08]*:

$\hat{\mathbf{\Pi}}_m^G = \sum_{k=1}^{m} \phi(k)\, \hat{\mathbf{u}}_k\, \hat{\mathbf{u}}_k^H$ is a consistent estimator of $\mathbf{\Pi}_m$, where

$$\phi(k) = \begin{cases} 1 + \displaystyle\sum_{i=m-p+1}^{m} \left(\frac{\hat{\lambda}_i}{\hat{\lambda}_k - \hat{\lambda}_i} - \frac{\hat{\mu}_i}{\hat{\lambda}_k - \hat{\mu}_i} \right), & k \leq m-p \\ -\displaystyle\sum_{i=m-p+1}^{m} \left(\frac{\hat{\lambda}_i}{\hat{\lambda}_k - \hat{\lambda}_i} - \frac{\hat{\mu}_i}{\hat{\lambda}_k - \hat{\mu}_i} \right), & k > m-p \end{cases}$$

and where $\hat{\mu}_1 \leq \hat{\mu}_2 \leq \ldots \leq \hat{\mu}_m$ are the eigenvalues of $\mathrm{diag}(\hat{\boldsymbol{\lambda}}) - \sqrt{\hat{\boldsymbol{\lambda}}}\sqrt{\hat{\boldsymbol{\lambda}}}^T$.

Example



MUSIC against G-MUSIC for DoA detection of $p = 3$ signal sources, $m = 20$ sensors, $N = 150$ samples, SNR:10 dB. Angles of arrival: 10 deg, 35 deg, and 37 deg (Thanks to *[Couillet 11]*).

Outline

Model with correlated Gaussian noise

Let $\{\mathbf{y}_i\}_{i\in[1,N]}$ be N observations of size m characterizing the $p < m$ mixed sources corrupted by additive noise:

$$\mathbf{y}_i = \sum_{j=1}^{p} s_{i,j}\,\mathbf{m}_j + \mathbf{C}^{1/2}\,\mathbf{n}_i\,, \qquad i \in [1, N]\,,$$

which can be rewritten more compactly as:

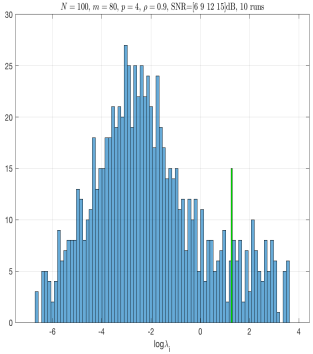
$$\mathbf{Y} = \mathbf{M}\,\mathbf{S} + \mathbf{C}^{1/2}\,\mathbf{N}\,,$$

where

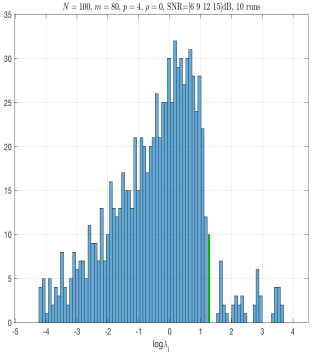
- $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N] \in \mathbb{C}^{m\times N}$ are the observations,
- $\mathbf{M} \in \mathbb{C}^{m\times p}$ is the mixing matrix containing spectra of the p sources,
- $\mathbf{S} \in \mathbb{C}^{p\times N}$ is the channel gain matrix,
- $\mathbf{N} \in \mathbb{C}^{m\times N}$ is the white Gaussian noise ($E\left[\mathbf{n}_i^H\,\mathbf{n}_i\right] = 1$), independent of the source signal,
- $\mathbf{C} \in \mathbb{C}^{m\times N}$ a **Toeplitz** Hermitian covariance matrix ($\mathrm{Tr}(\mathbf{C}) = m\,\sigma^2$).

Problems due to the correlation

$c = 0.8,\ p = 4,\ \rho = 0.9$



$c = 0.8,\ p = 4,\ \rho = 0$



SCM eigenvalues support for Gaussian noise and 4 random sources (SNR= [6 9 12 15]dB.
Left): **colored noise**. Right): **white noise**.

Consistent Estimation for $\mathbf{C}$: Gaussian Case

Proposition: *[Terreaux 17]*

As $m, N \to \infty$ such that $m/N \to c \in [0, \infty[$,

$$\left\| \mathcal{T}\left[\frac{1}{N}\mathbf{Y}\mathbf{Y}^H\right] - \mathbf{C} \right\| \xrightarrow{a.s.} 0\,,$$

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

A consistent estimator $\hat{\mathbf{C}}$ of the background noise covariance matrix $\mathbf{C}$ characterizing the background noise is therefore defined through observations $\mathbf{Y}$ as $\hat{\mathbf{C}} = \mathcal{T}\left[\frac{1}{N}\mathbf{Y}\mathbf{Y}^H\right]$.

We can now **whiten** the observation $\mathbf{Y}$ by $\hat{\mathbf{C}}^{-1/2}\mathbf{Y}$.

Behavior of whitened data: Gaussian Case

Let $\mathbf{Y}_w = \left(\mathcal{T}\left[\frac{1}{N}\mathbf{Y}\mathbf{Y}^H\right]\right)^{-1/2}\mathbf{Y}$ be the whitened data

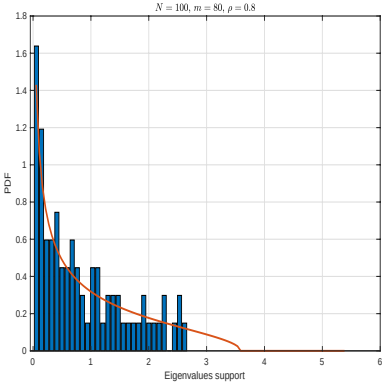
Proposition: *[Terreaux 17, Terreaux 18]*

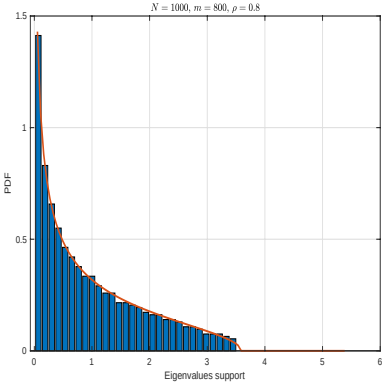
As $m, N \to \infty$ such that $m/N \to c \in [0, \infty[$, if $\mathbf{Y}_w$ does not contain sources, then:

$$\left\| \frac{1}{N}\mathbf{Y}_w\mathbf{Y}_w^H - \frac{1}{N}\mathbf{N}\mathbf{N}^H \right\| \xrightarrow{a.s.} 0,$$

- Without sources, the spectral distribution of the whitened data covariance matrix of $\mathbf{Y}_w$ follows a Marchenko-Pastur distribution (same spectral distribution of unobservable covariance matrix of $\mathbf{N}$) characterized by its support $\left[\left(1-\sqrt{c}\right)^2, \left(1+\sqrt{c}\right)^2\right]$,
- All eigenvalues greater than $\left(1+\sqrt{c}\right)^2$ can be considered as sources,
- Detection occurs if $\text{SNR} = \frac{s_j^2\,\mathbf{m}_j^H\mathbf{m}_j}{m\,\sigma^2} \geq \sqrt{c}$.

Some RMT results: Gaussian Case

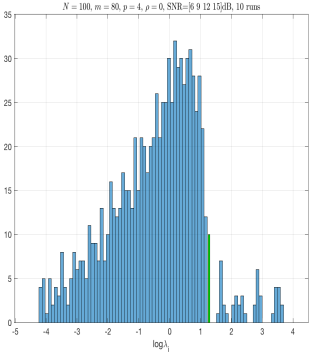




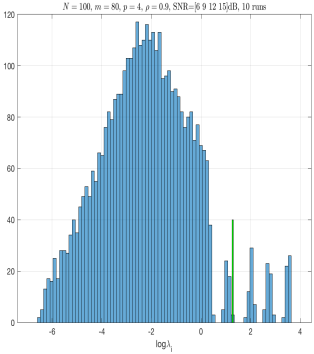
$$\left\| \frac{1}{N} \mathbf{Y}_w \mathbf{Y}_w^H - \frac{1}{N} \mathbf{N}\mathbf{N}^H \right\| \xrightarrow{a.s.} 0\,, \qquad E\left[\mathbf{N}\mathbf{N}^H\right] = \mathbf{I}_m$$

Example: Gaussian noise and 4 random sources

$c = 0.8$, $p = 4$, $\rho = 0$, 10 runs



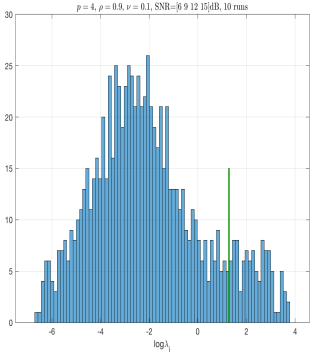
$c = 0.8$, $p = 4$, $\rho = 0.9$, 10 runs



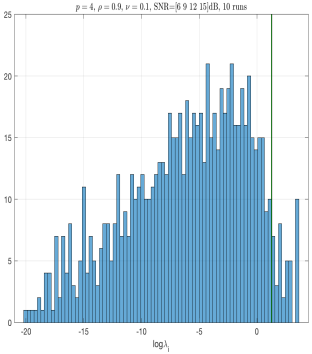
SCM eigenvalues support for Gaussian noise and sources (SNR= [6 9 12 15]dB). Left): **white noise**. Right): **whitened** colored noise .

Problems in non-Gaussian case

$c = 0.8,\ p = 4,\ \rho = 0.9,\ \nu = 0.1$



$c = 0.8,\ p = 4,\ \rho = 0.9,\ \nu = 0.1$



SCM eigenvalues support for K-distributed noise and sources (SNR= [6 9 12 15]dB). Left): **colored K-distributed noise**. Right): **whitened colored K-distribution noise**.

Model with correlated Non Gaussian noise

Let $\{\mathbf{y}_i\}_{i\in[1,N]}$ be N observations of size m characterizing the $p < m$ mixed sources corrupted by additive noise:

$$\mathbf{y}_i = \sum_{j=1}^{p} s_{i,j}\,\mathbf{m}_j + \sqrt{\tau_i}\,\mathbf{C}^{1/2}\,\mathbf{n}_i\,, \qquad i \in [1, N]\,,$$

which can be rewritten more compactly as:

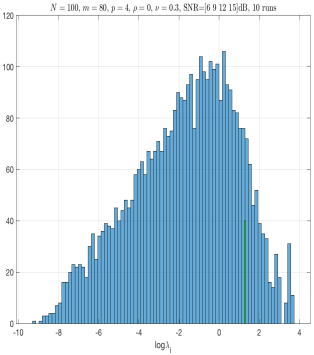
$$\mathbf{Y} = \mathbf{M}\,\mathbf{S} + \mathbf{C}^{1/2}\,\mathbf{N}\,\mathbf{T}^{1/2}\,,$$

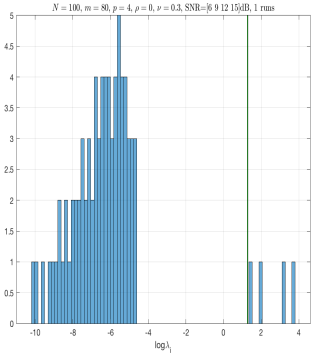
where

- $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N] \in \mathbb{C}^{m\times N}$ are the observations,
- $\mathbf{M} \in \mathbb{C}^{m\times p}$ is the mixing matrix containing spectra of the p sources,
- $\mathbf{S} \in \mathbb{C}^{p\times N}$ is the channel gain matrix, $\mathbf{T}$ is the texture diagonal matrix,
- $\mathbf{N} \in \mathbb{C}^{m\times N}$ is the white Gaussian noise ($E\left[\mathbf{n}_i^H\,\mathbf{n}_i\right] = 1$), independent of the source signal,
- $\mathbf{C} \in \mathbb{C}^{m\times N}$ a **Toeplitz** Hermitian covariance matrix ($\mathrm{Tr}(\mathbf{C}) = m\,\sigma^2$).

Key idea 1: to use Robust Covariance Matrix Estimation

$N = 100,\ m = 80,\ c = 0.8,\ p = 4$ $\quad$ $N = 100,\ m = 80,\ c = 0.8,\ p = 4$

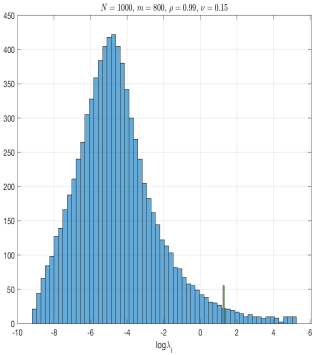




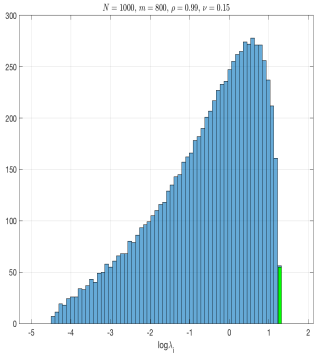
Eigenvalues support for **white K-distributed** noise ($\sigma^2 = 1$, $\nu = 0.3$) and 4 sources (SNR= [6 9 12 15]dB). Left): **SCM**. Right): **Tyler**.

Key idea 2: to whiten the correlated data

$c = 0.8,\ \nu = 0.15,\ \rho = 0.99$



$c = 0.8,\ \nu = 0.15,\ \rho = 0.99$



Tyler eigenvalues support for correlated K-distributed noise ($\sigma^2 = 1$). Left): **unwhitened data.** Right): **whitened data** (right).

Robust Consistent Estimation for $\mathbf{C}$: General case

Let $\hat{\mathbf{M}}_{FP} = \frac{m}{N} \sum_{i=1}^{N} \frac{\mathbf{y}_i \mathbf{y}_i^H}{\mathbf{y}_i^H \hat{\mathbf{M}}_{FP}^{-1} \mathbf{y}_i}$ be the Tyler M-estimator of $\mathbf{Y}$ scatter matrix.

Proposition: *[Terreaux 17]*

As $m, N \to \infty$ such that $m/N \to c \in [0, \infty[$,

$$\left\| \mathcal{T}\left[\hat{\mathbf{M}}_{FP}\right] - \mathbf{C} \right\| \xrightarrow{a.s.} 0\,,$$

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

A consistent estimator $\hat{\mathbf{C}}$ of the background noise covariance matrix $\mathbf{C}$ characterizing the background noise is therefore defined through observations $\mathbf{Y}$ as $\hat{\mathbf{C}} = \mathcal{T}\left[\hat{\mathbf{M}}_{FP}\right]$.

We can now **whiten** the observation $\mathbf{Y}$ by $\hat{\mathbf{C}}^{-1/2}\mathbf{Y}$.

Behavior of whitened data: General case

Let $\mathbf{Y}_w = \left(\mathcal{T}\left[\hat{\mathbf{M}}_{FP}\right]\right)^{-1/2} \mathbf{Y}$ be the whitened data and $\hat{\mathbf{W}}_{FP}$ be the Tyler M-estimator of $\mathbf{Y}_w$.

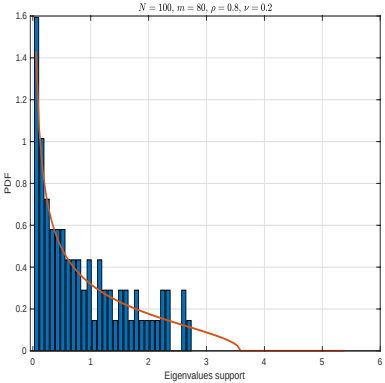
Proposition: *[Terreaux 17]*

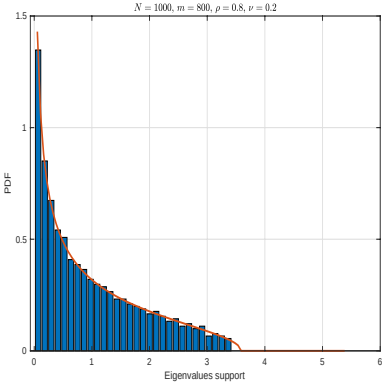
As $m, N \to \infty$ such that $m/N \to c \in [0, \infty[$, if $\mathbf{Y}_w$ does not contain sources, then:

$$\left\| \hat{\mathbf{W}}_{FP} - \frac{1}{N} \mathbf{N}\,\mathbf{N}^H \right\| \xrightarrow{a.s.} 0\,,$$

- Without sources, the spectral distribution of the whitened data scatter matrix of $\mathbf{Y}_w$ follows a Marchenko-Pastur distribution (same spectral distribution of unobservable covariance matrix of $\mathbf{N}$) characterized by its support $\left[\left(1-\sqrt{c}\right)^2, \left(1+\sqrt{c}\right)^2\right]$,
- All eigenvalues greater than $\left(1+\sqrt{c}\right)^2$ can be considered as sources,
- Detection occurs if SNR $= \dfrac{s_j^2\, \mathbf{m}_j^H \mathbf{m}_j}{m\,\sigma^2} \geq \sqrt{c}$.

Some RMT results: Non-Gaussian Case

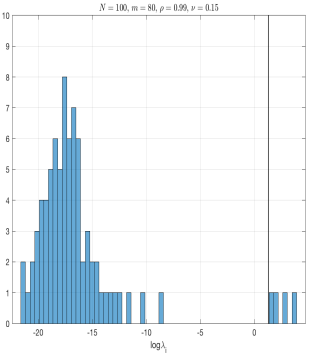


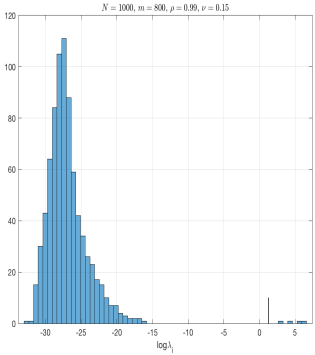


$$\left\| \hat{\mathbf{W}}_{FP} - \frac{1}{N} \mathbf{N}\,\mathbf{N}^H \right\| \xrightarrow{a.s.} 0\,, \qquad E\left[\mathbf{N}\,\mathbf{N}^H\right] = \mathbf{I}_m$$

Example

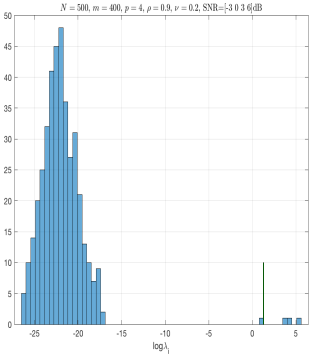
$c = 0.8,\ p = 4,\ \rho = 0.99,\ \nu = 0.15$ $\quad$ $c = 0.8,\ p = 4,\ \rho = 0.99,\ \nu = 0.15$

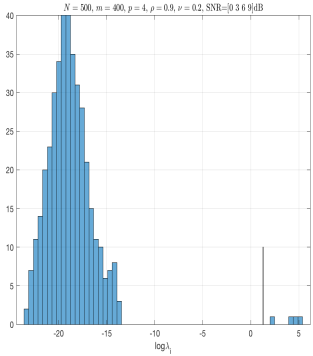




Tyler eigenvalues support for colored K-distributed noise and 4 sources (SNR= [3 6 9 10]dB).
Left): $N = 100$, $m = 80$. Right): $N = 1000$, $m = 800$.

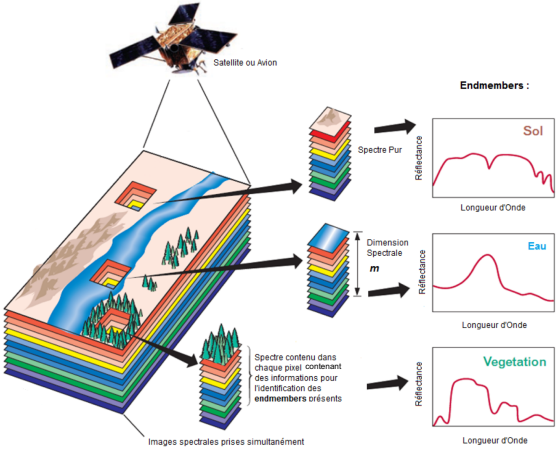
SNR Impact





Tyler eigenvalues support for whitened observations (4 random sources and colored K-distributed noise). Left): **SNR= [-3 0 3 6] dB.** Right): **SNR= [0 3 6 9]dB.**

Hyperspectral Imaging



General Problems

- Estimation of the endmembers number
- Detection/Estimation of sources / Anomaly Detection
- Unmixing

Considered Problems

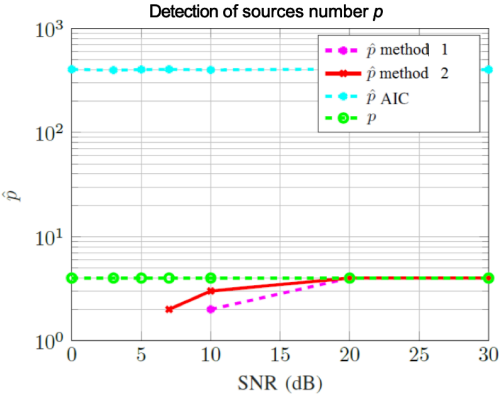
- Estimation of the number of endmembers
- Estimation of their spectrum

Hyperspectral Imaging

With the set of observations $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_N]$

- Estimation of the noise scatter matrix $\hat{\mathbf{C}} = \mathcal{T}\left[\hat{\mathbf{M}}\right]$ by Toeplitz rectification on:
  - Method 1: Maronna's M-estimators *[Maronna 76]* adapted to data statistic :
    $$\hat{\mathbf{M}} = \frac{1}{N}\sum_{i=0}^{N-1} u\left(\frac{1}{m}\,\check{\mathbf{y}}_i^H\,\hat{\mathbf{M}}^{-1}\,\mathbf{y}_i\right)\,\mathbf{y}_i\,\mathbf{y}_i^H.$$
  - Method 2: Tyler's M-estimator : $\hat{\mathbf{M}} = \dfrac{m}{N}\displaystyle\sum_{i=1}^{N}\frac{\mathbf{y}_i\,\mathbf{y}_i^H}{\mathbf{y}_i^H\,\hat{\mathbf{M}}^{-1}\,\mathbf{y}_i}.$
- Whitening observations: $\mathbf{Y}_w = \left(\mathcal{T}\left[\hat{\mathbf{M}}\right]\right)^{-1/2}\mathbf{Y}.$
- Thresholding the eigenvalue distribution of the whitened data scatter matrix $\hat{\mathbf{W}}$:
  - Method 1: Threshold depending on the function $u(.)$ and data for Maronna's M-estimator $\hat{\mathbf{W}}$,
  - Method 2: Threshold independent of data for Tyler's M-estimator $\hat{\mathbf{W}}$,

Hyperspectral Imaging



4 sources, $N = 2000$, $m = 900$, $\{\tau\}_{i \in [\![1, N]\!]}$ inverse-gamma

Hyperspectral Imaging

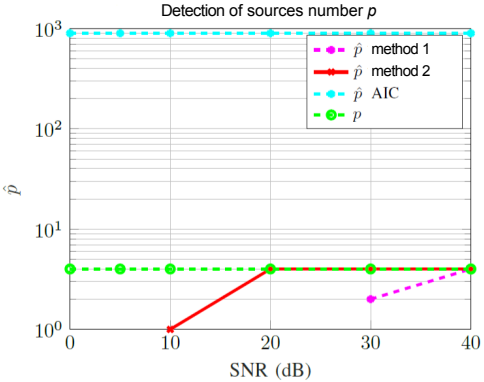
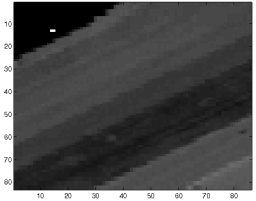


Fig 4 sources, $N = 2000$, $m = 900$, $\{\tau\}_{i \in [\![1, N]\!]}$ Student-t

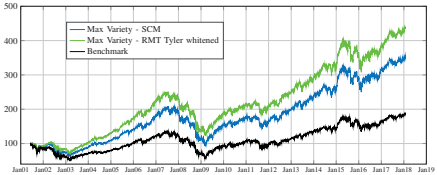
Hyperspectral Imaging

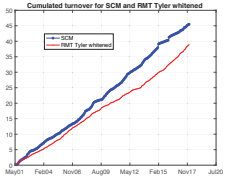


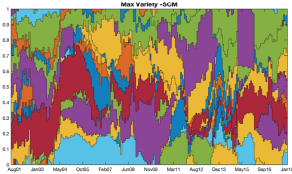
Estimation of the number of the most energetic endmembers

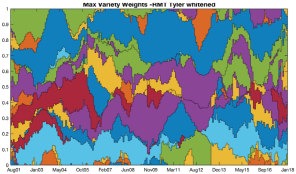
	Images	Indian Pines	SalinasA	PaviaU	Cars
	p	16	9	9	6
	$\hat{p}_{AIC}$	219	203	102	143
	$\hat{p}_{Hysime}$	**19**	14	60	19
Method 1	$\hat{p}_{FP}$	11	**9**	1	**3**
Method 2	$\hat{p}_{TYL}$	**13**	2	**10**	**13**

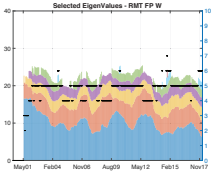
Portfolio Performance Optimization *[Jay 2018, Jay 2020]*











Max Variety Portfolios	Ann. Return	Ann. Volatility	Ratio (Ret / Vol)	Max DD
RMT Tyler Whithened	9,71%	12,9%	0,75	50,41%
SCM	8,51%	13,80%	0,62	55,02%
Benchmark	4,92%	15,19%	0,32	58,36%

Outline

Conclusions

This work has extended classical Model Order Selection techniques (AIC, MDL, etc.) for correlated and non-Gaussian additive noise.

- This extension was efficiently derived using latest results coming from RMT assuming **Toeplitz covariance structure assumption** for the noise covariance matrix,
- This quite *simple* technique can be easily applied on experimental data (radar, STAP, MIMO-STAP, SAR, HS, finance).

Bibliography

- H. Akaike, A new look at the statistical model identification. *Automatic Control, IEEE Transactions on* **19**(6), pp.716-723, 1974.
- T. W. Anderson, Asymptotic theory for principal component analysis. *Ann. J. Math. Stat.*, **34**, pp.122-148, 1963.
- A. Combernoux, Low rank detection and estimation using random matrix theory approaches for antenna array processing. PhD Thesis, University of Paris-Saclay, 2016.
- R. Couillet and M. Debbah, Random matrix methods for wireless communications. Cambridge University Press, 2011.
- 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)*, pp.149-152, 2015.
- E. Jay, E. Terreaux, J.-P. Ovarlez and F. Pascal. Improving portfolios global performance with robust covariance matrix estimation: Application to the maximum variety portfolio. *Proc. Eusipco*, 2018.
- E. Jay, T. Soler, E. Terreaux, J.-P. Ovarlez, F. Pascal, P. De Peretti, C. Chorro, Improving portfolios global performance using a cleaned and robust covariance matrix estimate", *Soft Computing, Springer Nature*, **24**(12), pp.8643-8654, March 2020.
- R. A. Maronna, Robust M-estimators of multivariate location and scatter. Annals of Statistics, **4**(1), pp.51-67, 1976.
- X. Mestre and M. A. Lagunas, Modified subspace algorithms for DoA estimation with large arrays. *Signal Processing, IEEE Transactions on*, **56**, pp.598-614, 2008.
- J. Rissanen, Modeling by shortest data description. *Automatica*, **14**(5), pp.465-471, 1978.
- G. Schwarz, Estimating the dimension of a model. *The Annals of Statistics*, **6**(2), pp.461-464, 1978.
- E. Terreaux, J.-P. Ovarlez and F. Pascal. Robust model order selection in large dimensional elliptically symmetric noise. arXiv preprint, 2017.
- E. Terreaux, J.-P. Ovarlez and F. Pascal. A Toeplitz-Tyler estimation of the model order in large dimensional regime. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp.4489-4493, 2018.
- M. Wax and T. Kailath . Detection of signals by information theoretic criteria. *Acoustics, Speech, and Signal Processing, IEEE Transactions on*, **33**(2), pp.387-392, 1985.