Robust Estimation and Detection Schemes in Non-Standard Conditions for Radar, Array Processing and Imaging

Jean-Philippe Ovarlez[1,2]

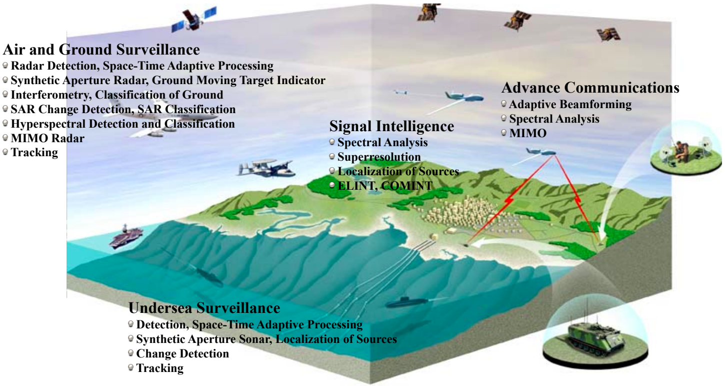
[1]SONDRA, CentraleSupélec, France
[2]French Aerospace Lab, ONERA DEMR/TSI, France

Joint works with F. Pascal, P. Forster, G. Ginolhac, M. Mahot, J. Frontera-Pons, A. Breloy, G. Vasile, and many others

12ème École d'Été de Peyresq
en Traitement du Signal et des Images
25 juin au 01 juillet 2017

General Introduction

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



General Introduction

Survey on

- General statistical non-Gaussian modeling (spherically, elliptically random processes),
- Robust covariance matrix estimation schemes (MLE, M-estimators),
- Robust detection schemes (Adaptive Normalized matched Filter).

3 Main Parts

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

Contents

- **Part A:**
  Background on Radar, Array Processing, SAR and Hyperspectral Imaging
- **Part B:**
  Robust Detection and Estimation Schemes
- **Part C:**
  Applications and Results in Radar, STAP and Array Processing, SAR Imaging, Hyperspectral Imaging

Part A

Background on Radar, Array Processing, SAR and Hyperspectral Imaging

Part A: Contents

Outline

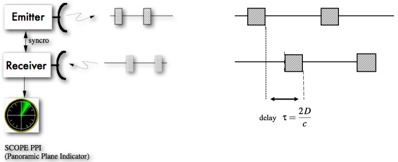
Introduction

RADAR = RAdio Detection And Ranging

- emits and receives electromagnetic waves,
- detects targets,
- estimates target parameters (range, radial velocity, angles of presentation, acceleration, amplitude (related to Radar Cross Section), etc.)
- images, recognizes, classifies,

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.

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

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

Velocity Measurement

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

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

then

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

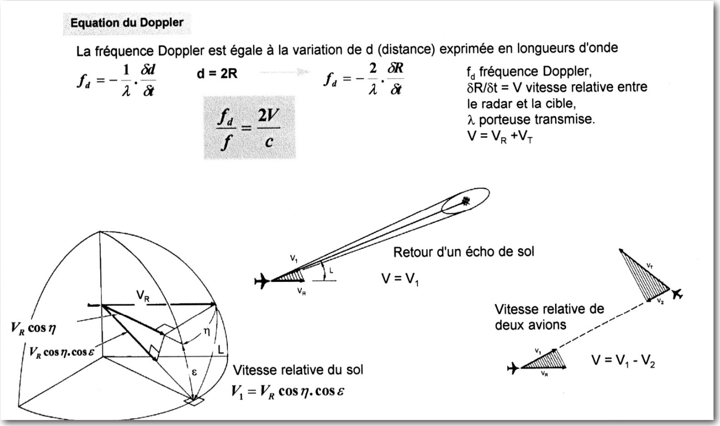
$$
\boxed{s_r(t) = \alpha'\, s_e\left(t - \frac{2\, D_0}{c}\right) \exp\left(-2 i\, \pi\, f_d\, t\right) + b(t)\,.}
$$

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

Play/Pause

Doppler Effect



Ambiguity function and distance criterion

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:

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

According to the physical transformation of X, we obtain the so-called Ambiguity functions [Woodward, 1953, Kelly and Wishner, 1965]:

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

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:

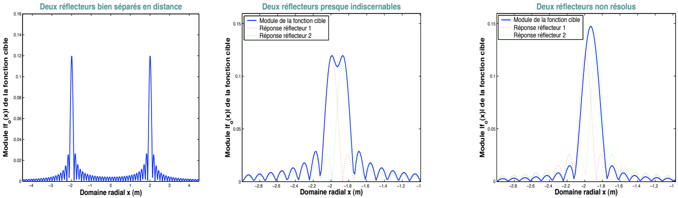
$$s_r(t) = \sum_{i=1}^{N} \alpha_i\, s_e(t-\tau_i) \overset{t\to f}{\Longrightarrow} S_r(f) = \sum_{i=1}^{N} \alpha_i\, S_e(f)\, e^{-2i\,\pi\,f\,\tau_i}\,.$$

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 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 $\left\{\nu_i = \dfrac{2\,v_i}{c}\,f_0\right\}_{i\in[1,N]}$.

The received signal $S_r(f)$ is:

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

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

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

The velocity resolution δV 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 $\left\{d_i = \dfrac{c\,\tau_i}{2}\right\}_{i\in[1,N]}$. The received signal $S_r(f)$ is:

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

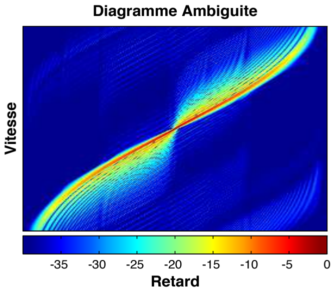
The radar processing leads to evaluate for all (τ,ν), the following expression:

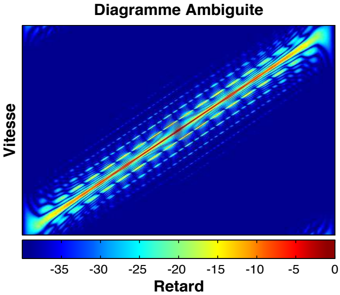
$$R(\tau,\nu) = \int_{-\infty}^{+\infty} s_r(t)\, s_e^*(t-\tau)\, e^{-2i\,\pi\,\nu\,t}\, dt\,.$$

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 of the signal $\sigma_t^2 = \int_{-\infty}^{+\infty} |s_e(t)|^2 \, dt \approx T^2$, $\sigma_f^2 = \int_{-\infty}^{+\infty} |S_e(f)|^2 \, df \approx B^2$ and the modulation index $m = \dfrac{-1}{2\pi} \mathrm{Im} \int_{-\infty}^{+\infty} t \, s_e(t) \, \dfrac{ds_e^*(t)}{dt} \, dt$.
  Under white Gaussian noise with variance σ^2, range and doppler accuracies are given by the following Cramer-Rao bounds [Kay, 1993]:

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

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

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

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

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 the following expression:

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

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	$\frac{c\,T_r}{2}$	$\frac{c}{2\,f_0\,T_r}$

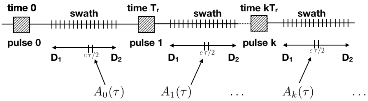
Characteristics of m pulses train with duration T, bandwidth B, PRF $1/T_r$ and carrier frequency f_0

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:

$$
\begin{aligned}
R(\tau,\nu) &= \sum_{n=0}^{m-1} e^{-2i\,\pi\,\nu\,n\,T_r} \int_0^{T_r} s_r(u + n\,T_r + \tau)\, s_e^*(u)\, du\,, \\
&= \sum_{n=0}^{m-1} A_n(\tau)\, e^{-2i\,\pi\,\nu\,n\,T_r} = \mathbf{A}^T\,\mathbf{p}\,,
\end{aligned}
$$

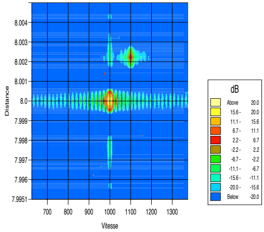
where $\mathbf{A} = (A_0(\tau), A_1(\tau), \ldots, A_{m-1}(\tau))^T$ and $\mathbf{p} = \left(1, e^{-2i\,\pi\,\nu\,T_r}, \ldots, e^{-2i\,\pi\,\nu\,(m-1)\,T_r}\right)^T$.



- 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 $A_n(\tau)$ corresponding to the time correlation between received signal and each emitted pulse $s_e(t)$,

Range Doppler Radar Processing

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



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

Noise in Radar

Thermal noise

Thermal noise for most radars corresponds to additive complex white Gaussian noise $\mathcal{CN}(\mathbf{0}, \mathbf{I})$.
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 ????

Formulation of the Range-Doppler Radar Detection Problem

Set of two binary hypotheses

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

where

- $\mathbf{y}$ is a m-vector of data collected in the same given range bin $c\,\tau/2$ and characterizing the reflected signal for each emitted pulse of the burst.
- The complex amplitude A is considered here deterministic.
- The m-vector b represents the additive noise (thermal noise, clutter, jam, etc.) characterized by a known (or unknown) PDF.
- The m-vector $\mathbf{p}$ represents the so-called *steering vector*: here $\mathbf{p} = (1, \exp(2i\,\pi\,\nu\,T_r), \exp(2i\,\pi\,\nu\,2\,T_r), \ldots, \exp(2i\,\pi\,\nu\,(m-1)\,T_r)]^T$, where the Doppler frequency ν is unknown and has to be estimated.

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

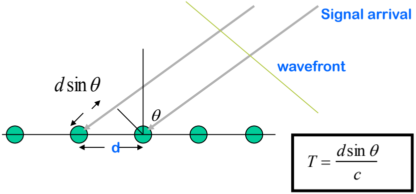
Outline

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:

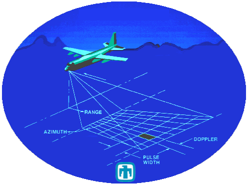
$$\mathbf{y} = A\,\mathbf{p} \otimes \left(1, e^{2i\pi\, f_0\, d\, \sin\theta/c}, \dots, e^{2i\pi\, f_0\, (N_s-1)\, d\, \sin\theta/c}\right)^T + b(t)\,.$$



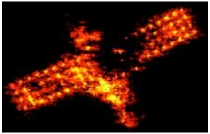
Outline

Background on SAR and Radar Imaging

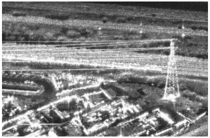
ONERA RAMSES Image



RAMSES Image



ONERA ISAR Image



ONERA RAMSES Image

Radar Imaging [Mensa, 1981, Soumekh, 1994, Soumekh, 1999] allows to build more and more precise images:

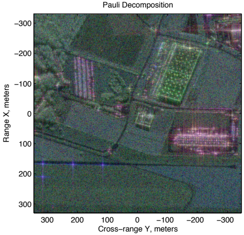
- Current use of very high spectral bandwidth and very high angular bandwidth leading to very high spatial resolution,
- Application to monitoring (detection, change detection), classification, 3D reconstruction, EM analysis, etc.

These applications require some physical diversity to reach good performances.

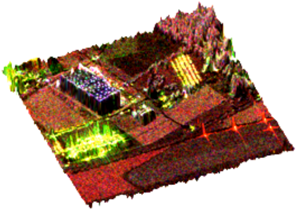
Multi-Channel SAR Images

Multi-channel SAR images automatically propose this diversity through:

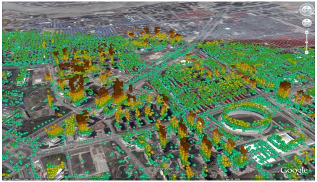
- polarimetric channels (POLSAR), interferometric channels (INSAR), polarimetric and interferometric channels (POLINSAR),
- multi-temporal, multi-passes SAR Image, etc.



EM behavior of the terrain in POLSAR images



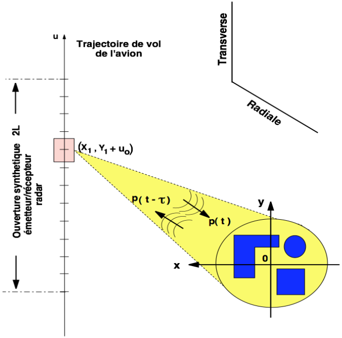
Estimation of the height in POLINSAR images



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**.

SAR Processing



Goal of SAR Imaging: Invert the relation:

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

Range Migration (RMA) SAR Processing Steps

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

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

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

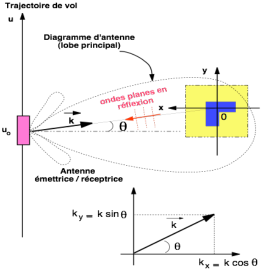
$$\Downarrow \quad u \overset{\mathcal{F}^{-1}}{\Rightarrow} k_u\,,$$

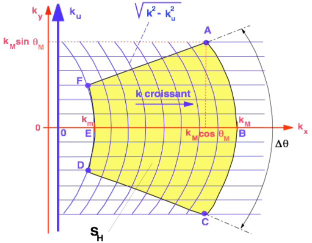
$$S_r(k,k_u) = S_e(k) \iint_{\mathbb{R}^2} I(x,y)\, \exp\left(-2i\,\pi\left((X_1-x)\sqrt{k^2-k_u^2} + (Y_1-y)\,k_u\right)\right) dx\, dy\,,$$

$$\Downarrow \quad \left\{ \begin{array}{l} k_x = \sqrt{k^2-k_u^2} \\ k_y = k_u \end{array} \right.$$

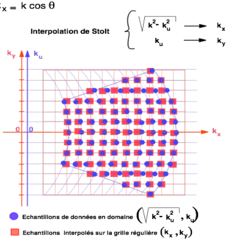
$$\boxed{S_r(k_x,k_y) = S_e(k)\, \exp\left(-2i\,\pi\, k_x\, X_1 + k_y\, Y_1\right) \iint_{\mathbb{R}^2} I(x,y)\, \exp\left(2i\,\pi\,(k_x\, x + k_y\, y)\right) dx\, dy}$$

Range Migration Algorithm Principle





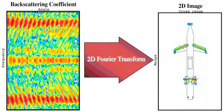
Support spectral du signal reçu $S_r(k_x, k_y)$



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:

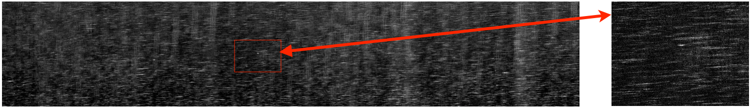
$$H(\mathbf{k}) = \int_{\mathcal{D}_{\mathbf{x}}} I(\mathbf{x}) \exp\left(-2\,i\,\pi\,\mathbf{k}^T\,\mathbf{x}\right) d\mathbf{x}$$

- 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 et al., 1994].

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

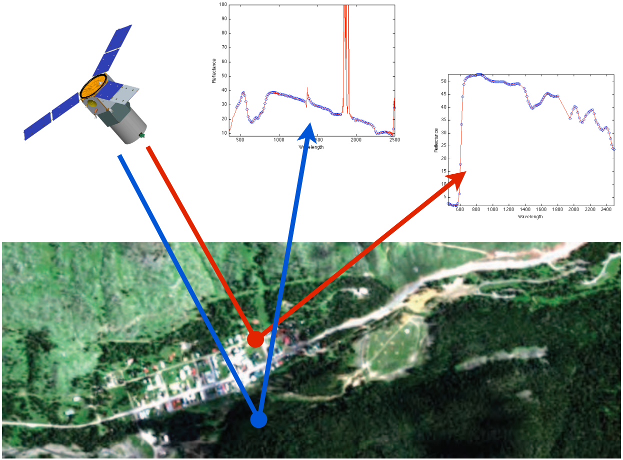
$$\lambda = N \left(P_{fa}^{-1/N} - 1 \right), \qquad \Lambda(x_i) = \frac{|x_i|^2}{\frac{1}{N} \sum_{k \neq i}^{N} |x_k|^2} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda\,,$$

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

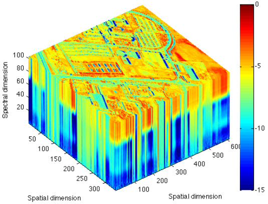
ÊAdaptive multi-channels SAR detection framework can be extended with diversity contained in the steering vector $\mathbf{p}$ (polarimetry, interferometry, sub-looks and sub-bands decomposition ([Ovarlez et al., 2017], see Ammar Mian's PhD talk).

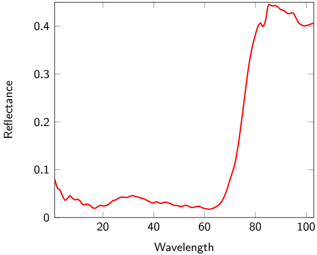
Outline

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-Pons et al., 2016].
- **"Pure" Detection**
  To detect targets characterized by a given spectral signature $\mathbf{p}$ - Regulation of False Alarm [Frontera-Pons et al., 2017].

Outline

Problem Statement

- When the noise parameters are known:

In a m-vector $\mathbf{z}$ of observation, detecting a complex deterministic signal $\mathbf{s} = A\,\mathbf{p}$ embedded in an additive noise $\mathbf{b}$ can be written as the following set of binary hypotheses test:

$$\begin{cases} \text{Hypothesis } H_0: & \mathbf{z} = \mathbf{b}\,, \\ \text{Hypothesis } H_1: & \mathbf{z} = \mathbf{s} + \mathbf{b}\,. \end{cases}$$

- When the noise parameters are unknown: (covariance, mean, etc.):

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

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

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

$\Rightarrow$ Neyman-Pearson criterion [Kay, 1993, Kay, 1998]

Detection Theory

When all parameters (noise, target) are known

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

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

  where λ is set for a given *PFA* (set by the user):

- Probability of False Alarm (type-I error):

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

- Probability of Detection (to evaluate the performance):

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

  for different Signal-to-Noise Ratios (SNR).

General Detection Theory

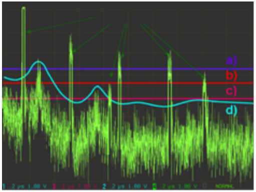
When some parameters (noise, target) are unknown:

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

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

where θ and μ represent respectively the unknown target parameter vector and the unknown noise parameter vector.

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.

Ê

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 $\mathbf{b} \sim \mathcal{CN}(\mathbf{0}_m, \boldsymbol{\Sigma})$ with known covariance matrix $\boldsymbol{\Sigma}$. The probability density functions of the received m-vector $\mathbf{z}$ under each hypothesis are given by:

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

The Log-Likelihood function test leads to:

$$\Lambda(\mathbf{z}) = \log \frac{p_{\mathbf{z}/H_1}(\mathbf{z})}{p_{\mathbf{z}/H_0}(\mathbf{z})} = 2\,\mathrm{Re}\left(A^H\,\mathbf{p}^H\,\boldsymbol{\Sigma}^{-1}\,\mathbf{z}\right) + |A|^2\;\mathbf{p}^H\,\boldsymbol{\Sigma}^{-1}\,\mathbf{p} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda\,.$$

$$\boxed{\Lambda(\mathbf{z}) \sim \mathcal{N}\left(|A|^2\;\mathbf{p}^H\,\boldsymbol{\Sigma}^{-1}\,\mathbf{p}, 2\,m\,|A|^2\right)\,.}$$

Example 2 - Matched Filter

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 $\mathbf{b} \sim \mathcal{CN}(\mathbf{0}_m, \boldsymbol{\Sigma})$ with known covariance matrix $\boldsymbol{\Sigma}$. The probability density functions of the received m-vector $\mathbf{z}$ under each hypothesis are given by:

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

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

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

Example 2 - Matched Filter - Derivation of Performances

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 $\Lambda(\mathbf{z}) \sim \dfrac{1}{2}\, \chi^2(2)$. We have:

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

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

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

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

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

Example 3 - Kelly and Adaptive Matched Filter (1)

Problem under study:

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

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

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

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

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

Example 3 - Kelly and Adaptive Matched Filter (2)

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

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

By denoting $\mathbf{S} = \sum_{k=1}^{n} \mathbf{z}_k\,\mathbf{z}_k^H$, we obtain $\underset{\boldsymbol{\Sigma}}{\mathrm{argmax}}\, p_{\mathbf{z},\{\mathbf{z}_k\}_k,\boldsymbol{\Sigma},A/H_1}(\mathbf{z}) = \dfrac{(\mathbf{z}-A\,\mathbf{p})\,(\mathbf{z}-A\,\mathbf{p})^H + \mathbf{S}}{n+1}$

and replacing these two expressions in the Generalized Log Likekihood ratio 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 $\min\limits_{A}\,|\mathbf{S}|\;\left(\|\mathbf{z}_s-A\,\mathbf{p}_s)\|^2+1\right) = |\mathbf{S}|\;\left(\|\mathbf{P}^{\perp}_{\mathbf{p}_s}\,\mathbf{z}_s\|^2+1\right)$ where $\mathbf{P}^{\perp}_{\mathbf{p}_s} = \mathbf{I}_m - \mathbf{p}_s\,\mathbf{p}_s^H/\mathbf{p}_s^H\,\mathbf{p}_s$.

Example 3 - 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, 1986]:

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

This detector has good properties but often (usually) replaced by a simpler one, the *Adaptive Matched Filter* [Robey et al., 1992]:

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

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 4 - Detection in quasi-homogeneous Gaussian Noise - Adaptive Normalized Matched Filter

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

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

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

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

The corresponding detector [Scharf and Friedlander, 1994, Kraut and Scharf, 1999] 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):

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

Modeling Homogeneous Gaussian noise/clutter

The Sample Covariance Matrix (SCM)

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

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

The Shrinkage or Diagonal Loading SCM [Ledoit and Wolf, 2004]: useful when $m \geq n$

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

Properties of the SCM in homogeneous Gaussian noise/clutter environment

Properties of the SCM

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

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

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

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

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

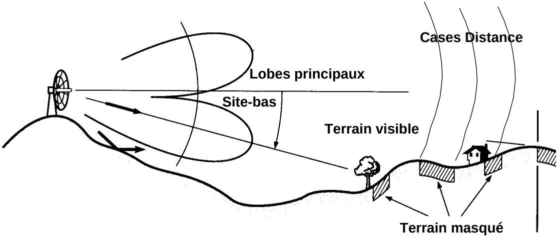
$$\widehat{\mathbf{S}}_n = \frac{1}{n}\sum_{k=1}^{n} \mathbf{z}_k\, \mathbf{z}_k^H\,.$$

- This estimate is unbiased, efficient, Wishart distributed,
- n can represent any samples support: in time, spatial, angular domain, $\mathbf{z}_k$ a vector of any information collected in any domain:
  - in **Radar Detection**, it can represent the time returns collected in a given range bin of interest, n is here the range bin support
  - in **Array Processing**, it can represent the spatial information collected by the antenna array at a given time, n is here the time support,
  - in **STAP**, it can represent the joint spatial and time information collected in a given range bin of interest, n is here the 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

- To have a SCM estimate invertible (whitening process), the number n of samples has to be bigger than the size m of the information collected $\mathbf{z}_k$,
- To improve the quality of the estimate, n has to be high but it means also that the space support has also to respect the initial Gaussian hypothesis (has to be statistically homogeneous) that is not always the case in the real world !
- Due to the increase of the radar resolution or due to the illumination angle, the number of the scatterers present in each cell (random walk) can become very small, the Central Limit Theorem being no longer valid [Jakeman, 1980]. Even if the number of scatterers is large enough to apply the CLT, this number can also randomly fluctuate from one resolution cell to another, leading to a backscattered signal locally Gaussian with random power (heterogeneous support)
- Robustness of the SCM: The n secondary data used to estimate the SCM may also contain another target returns, jammers, strong undesired scatterers which can lead to a poor or a biased estimate.

- Grazing angle Radar [Billingsley, 1993]

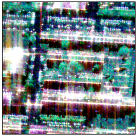


  - ⇒ Impulsive Clutter
  - ⇒ Spatial heterogeneity (e.g. in SAR or HS images)
- High Resolution Radar
  - ⇒ Small number of scatters in the Cell Under Test (CUT)
  - ⇒ Central Limit Theorem (CLT) is not valid anymore

- The SAR images are more and more complex, detailed, heterogeneous,
- The spatial statistic of SAR images is not at all Gaussian !

- Many Non Coherent Polarimetric Decomposition and classification techniques [Lee and Pottier, 2009] generally use an estimate $\widehat{\mathbf{S}}_n$ of the local spatial covariance matrix (coherency matrix), typically the Sample Covariance Matrix (SCM),
- All these techniques may give very different results when using another estimates that may fit better to the reality! Are they more physically valid? Which one to choose?

Non Coherent Polarimetric/Interferometric SAR Classification



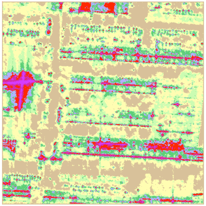
$I_1/I_3/I_2$

[*Vasile et al. 2008*]

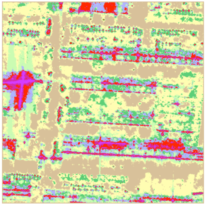
[*Formont et al. 2011*]

K-means H/α classification (8 classes)

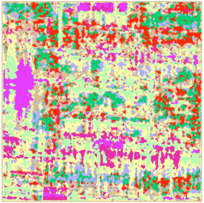
SPAN Gamma



H/α SCM-Wishart

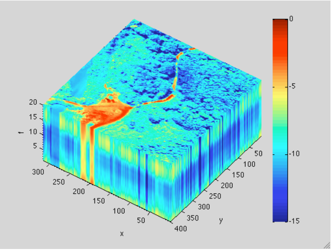


H/α FP-Wishart



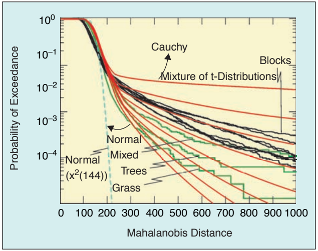
Classification on intensity only and H/α classification seem to be the same! The Gaussian SCM is contaminated by the power. Polarimetric information is lost [Vasile et al., 2010, Vasile et al., 2011, Formont et al., 2011, Formont, 2013]

- Anomaly Detection (e.g. RXD [Reed and Yu, 1990]) in Hyperspectral Images: To detect all that is different from the background (Mahalanobis distance $\mathbf{z}_k\, \widehat{\mathbf{S}}_n^{-1}\, \mathbf{z}_k$) - Regulation of False Alarm. Application to radiance images.
- Detection of targets in Hyperspectral Images: To detect (GLRT) targets (characterized by a given spectral signature p) - Regulation of False Alarm. Application to reflectance images (after some atmospheric corrections).



DSO data 2010

RXD CDF



[*Manolakis 2002*]

[Manolakis et al., 2014]

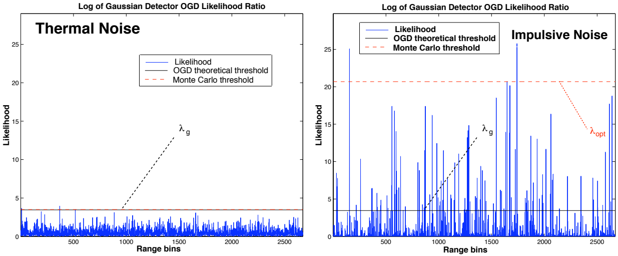


Figure: Failure of the Gaussian detector ($\lambda_g = -\log P_{fa}$): (left) Adjustment of the detection threshold, (right) K-distributed clutter with same power as the Gaussian noise

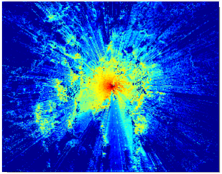
$\Rightarrow$ Bad performance of the conventional Gaussian detector in case of mis-modeling

$\Rightarrow$ Need/Use of non-Gaussian distributions

$\Rightarrow$ Need/Use of robust estimates

Going to adaptive detection

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



$\Rightarrow$ Covariance Matrix Estimation

Requirements:

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

End of Part A

Questions?

References I

Bertrand, J., Bertrand, P., and Ovarlez, J. P. (1994).
Frequency directivity scanning in laboratory radar imaging.
International Journal of Imaging Systems and Technology, 5(1):39–51.

Billingsley, J. B. (1993).
Ground clutter measurements for surface-sited radar.
Technical Report 780, MIT.

Formont, P. (2013).
Statistical and geometrical tools for the classication of highly textured polarimetric SAR images.
PhD thesis, University of Paris Sud.

Formont, P., Pascal, F., Vasile, G., Ovarlez, J. P., and Ferro-Famil, L. (2011).
Statistical classification for heterogeneous polarimetric SAR images.
IEEE Journal of Selected Topics in Signal Processing, 5(3):567–576.

Frontera-Pons, J., Pascal, F., and Ovarlez, J. P. (2017).
Adaptive nonzero-mean gaussian detection.
Geoscience and Remote Sensing, IEEE Transactions on, 55(2):1117–1124.

Frontera-Pons, J., Veganzones, M. A., Pascal, F., and Ovarlez, J. P. (2016).
Hyperspectral anomaly detectors using robust estimators.
IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9(2):720–731.

References II

Jakeman, E. (1980).
On the statistics of k-distributed noise.
Journal of Physics A: Mathematical and General, 13(1):31.

Kay, S. M. (1993).
Fundamentals of Statistical Signal Processing - Estimation Theory, volume 1.
Prentice-Hall PTR, Englewood CliJs, NJ.

Kay, S. M. (1998).
Fundamentals of Statistical Signal Processing - Detection Theory, volume 2.
Prentice-Hall PTR.

Kelly, E. J. (1986).
An adaptive detection algorithm.
Aerospace and Electronic Systems, IEEE Transactions on, 23(1):115–127.

Kelly, E. J. and Wishner, R. P. (1965).
Matched-filter theory for high-velocity, accelerating targets.
Military Electronics, IEEE Transactions on, 9(1):56–69.

Kraut, S. and Scharf, L. (1999).
The CFAR adaptive subspace detector is a scale-invariant GLRT.
Signal Processing, IEEE Transactions on, 47(9):2538–2541.

References III

Ledoit, O. and Wolf, M. (2004).
A well-conditioned estimator for large-dimensional covariance matrices.
Journal of multivariate analysis, 88(2):365–411.

Lee, J.-S. and Pottier, E. (2009).
Polarimetric Radar Imaging, From Basics to Applications.
CRC Press.

Manolakis, D., Truslow, E., Pieper, M., Cooley, T., and Brueggeman, M. (2014).
Detection algorithms in hyperspectral imaging systems: An overview of practical algorithms.
IEEE Signal Processing Magazine, 31(1):24–33.

Mensa, D. (1981).
High Resolution Radar Imaging.
Artech House, USA.

Ovarlez, J. P., Ginolhac, G., and Atto, A. M. (2017).
Multivariate linear time-frequency modeling and adaptive robust target detection in highly textured monovariate SAR image.
In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4029–4033.

Reed, I. and Yu, X. (1990).
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.

References IV

Rihaczek, A. W. (1969).
Principles of High Resolution Radar.
Mac-Graw-Hill.

Robey, F. C., Fuhrmann, D. R., Kelly, E. J., and Nitzberg, R. (1992).
A CFAR adaptive matched filter detector.
Aerospace and Electronic Systems, IEEE Transactions on, 28(1):208–216.

Scharf, L. L. and Friedlander, B. (1994).
Matched subspace detectors.
Signal Processing, IEEE Transactions on, 42(8):2146–2157.

Soumekh, M. (1994).
Fourier Array Imaging.
Prentice Hall, Englewood Cliffs.

Soumekh, M. (1999).
Synthetic Aperture Radar Signal Processing with MATLAB Algorithms.
John Wiley and Sons, New York.

Vasile, G., Ovarlez, J. P., Pascal, F., and Tison, C. (2010).
Coherency matrix estimation of heterogeneous clutter in high-resolution polarimetric SAR images.
Geoscience and Remote Sensing, IEEE Transactions on, 48(4):1809–1826.

References V

Vasile, G., Pascal, F., Ovarlez, J. P., Formont, P., and Gay, M. (2011).
Optimal parameter estimation in heterogeneous clutter for high-resolution polarimetric SAR data.
IEEE Geoscience and Remote Sensing Letters, 8(6):1046–1050.

Woodward, P. (1953).
Probability and Information Theory with Applications to Radar.
Pergamon, London (UK).