Contributions to SAR Image Time Series Analysis

Ammar Mian

Ph.d defense - 26 September 2019

Supervisors: Jean-Philippe Ovarlez (ONERA & SONDRA), Guillaume Ginolhac (LISTIC) and Abdourahmane M. Atto (LISTIC)
Contents of the presentation

1. Introduction
   - Motivations
   - Data

2. Change detection based on a statistical framework
   - The problem
   - State of the art

3. Robust Covariance homogeneity testing for Change Detection
   - Tests using Complex Elliptical Symmetric distributions
   - Tests using deterministic compound-Gaussian modelling

4. Tests using Low-rank models

5. Clustering SAR ITS with Riemannian geometry: an opening

6. Conclusions
Contents of the presentation

1 **Introduction**
   - Motivations
   - Data

2 **Change detection based on a statistical framework**
   - The problem
   - State of the art

3 **Robust Covariance homogeneity testing for Change Detection**
   - Tests using Complex Elliptical Symmetric distributions
   - Tests using deterministic compound-Gaussian modelling

4 **Tests using Low-rank models**

5 **Clustering SAR ITS with Riemannian geometry: an opening**

6 **Conclusions**
Remote Sensing: big data analysis

Remote sensing allows to obtain image of the Earth’s surface for various applications such as **Change Detection**.

Huge increase in the number of available acquisitions:
- Sentinel-1: 12 days repeat cycle, since 2014
- TerraSAR-X: 11 days repeat cycle, since 2007
- UAVSAR, ... thousands of flight paths planned

**Problem**

→ There is a need to process the huge amount of data automatically!
SAR Image Time Series: changes analysis

Change detection is useful for various purposes: Activity monitoring

**Figure 1:** Terrasar-X images of the Burning-man festival
Change detection is useful for various purposes: Disaster assessment

**Figure 1:** Destruction map of Dorian Hurricane in the Bahamas using Change Detection over Sentinel-1 data
Synthetic aperture radar (SAR)

Principle of SAR

Advantages:
- All weather and illumination conditions (active technology)
- Very high-resolution (sub-meter) imaging

Comparison of optical and image
Multivariate data: natural or pre-processing

Feature selection

- Leverage **diversity** to improve the detection
- Requires to process **multivariate** pixels

Polarimetry ($p = 3$)

Wavelet decomposition (Spectro-angular diversity)

[1] Mian et al., 2019c

($p \geq 3$)

HR SAR image
Data dimensionality
Non-Gaussianity of HR Images

**Issue:** Data is non-Gaussian!

**Figure 2:** UAVSAR data (Courtesy NASA/JPL-Caltech)
Gaussian distribution fitting
Generalized Gaussian distribution fitting
Summary

In the context of this Ph.D we develop change detection methodologies which:

- leverage **diversity** in order to improve detection
  → It can be obtained through *wavelet transforms*  
  Chapter 2

- take into account the heterogeneity of HR SAR images
  → We develop methodologies that are **robust to non-Gaussianity**  
  Chapter 3

We also consider alternative problems to bring more information:

- about the time at which changes occur
  → Change-point detection problem  
  Chapter 4

- change patterns
  → Clustering problem  
  Chapter 5
Contents of the presentation

1. Introduction
   - Motivations
   - Data

2. Change detection based on a statistical framework
   - The problem
   - State of the art

3. Robust Covariance homogeneity testing for Change Detection
   - Tests using Complex Elliptical Symmetric distributions
   - Tests using deterministic compound-Gaussian modelling

4. Tests using Low-rank models

5. Clustering SAR ITS with Riemannian geometry: an opening

6. Conclusions
Statistical framework: principle

Interest of this approach
- Can account for physical modelling of the data/noise
- Strong theoretical guarantees from statistical literature
SAR image time series representation

Figure 3: Sliding windows $W_{1,T}$

<table>
<thead>
<tr>
<th>$p$</th>
<th>$N$</th>
<th>$T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>dimension of vectors</td>
<td>number of pixels on local window</td>
<td>number of dates in the time series</td>
</tr>
</tbody>
</table>
Parametric change detection

A probability model is assigned to the observations on the windows over time:

\[ x^t_k \sim p_{x_k}^t (x^t_k; \theta_t; \Phi_t). \]

The detection is done on some parameters of interest \( \theta_t \) while the remaining ones are the nuisance parameters \( \Phi_t \):

\[
\begin{align*}
H_0 : & \quad \theta_1 = \ldots = \theta_T = \theta_0 \quad \& \quad \Phi_1 \neq \ldots \neq \Phi_T, \\
H_1 : & \quad \exists (t, t') \in [1, T]^2, \theta_t \neq \theta_{t'} \quad \& \quad \Phi_1 \neq \ldots \neq \Phi_T.
\end{align*}
\] (1)

**Problems**

- Specify a model and parameters of interest which are
  - A good fit to the empirical distribution
  - Robust to a large class of distributions and outliers
- Find a test statistic to obtain a rule of decision between the two alternatives.
Test statistic

We want to obtain:

- a statistic of decision $\hat{\Lambda}$:
  \[
  \mathbb{C}^{p \times N} \times \cdots \times \mathbb{C}^{p \times N} \rightarrow \mathbb{R}
  \]
  \[
  \mathcal{W}_{1,T} \rightarrow \hat{\Lambda}(\mathcal{W}_{1,T})
  \]

- a threshold $\lambda$

So that

\[
\mathbb{P}\left( \hat{\Lambda}(\mathcal{W}_{1,T}) > \lambda/H_1 \right) \text{ is high while } \mathbb{P}\left( \hat{\Lambda}(\mathcal{W}_{1,T}) > \lambda/H_0 \right) \text{ is low.}
\]
Constant False Alarm Rate (CFAR) property

Assume a parametric model: $\forall k, \forall t, x^t_k \sim p_{x;\theta}(x; \theta)$.

**Definition**

A statistic $\hat{\Lambda}$ is said to be CFAR if for any set $(\theta_0, \theta_1)$, we have:

$$P\left(\hat{\Lambda}(W_1, T; \theta_0/H_0) = x\right) = P\left(\hat{\Lambda}(W_1, T; \theta_1/H_0) = x\right)$$

Example of a non CFAR statistic:
Generalized Likelihood Ratio Test

Given the change detection decision problem, the GLRT is formulated as follows:

\[ \hat{\lambda} = \frac{\max_{\theta_1, \ldots, \theta_T, \Phi_1, \ldots, \Phi_T} p_{W_1, T}(W_1, T; \theta_1, \ldots, \theta_T, \Phi_1, \ldots, \Phi_T)}{\max_{\theta_0, \Phi_1, \ldots, \Phi_T} p_{W_1, T}(W_1, T; \theta_0, \Phi_1, \ldots, \Phi_T)} \overset{H_1}{\geq} \lambda. \]  

\text{Good invariance properties [Kay and Gabriel, 2003].}
Gaussian modelling

Definition

A vector \( \mathbf{x} \in \mathbb{C}^p \) is said to Gaussian distributed with mean parameter \( \mu \in \mathbb{C}^p \) and covariance parameter \( \Sigma \in \mathbb{S}_+^p \), denoted \( \mathbf{x} \sim \mathcal{CN}(\mu, \Sigma) \) if the Probability Distribution Function (PDF) of its distribution is the following:

\[
p_{\mathbf{x}}(\mathbf{x}; \mu, \Sigma) = \frac{1}{\pi^p |\Sigma|^{-1}} \exp \left\{ - (\mathbf{x} - \mu)^{\text{H}} \Sigma^{-1} (\mathbf{x} - \mu) \right\}.
\]  

(3)

Introduced by [Conradsen et al., 2003] which is a reference in the domain. The data is modelled by a Gaussian model as follows:

\[ \mathbf{x}_k^t \sim \mathcal{CN}(0_p, \Sigma_t). \]

Detection test: \( \theta_t = \Sigma_t \) \& \( \Phi_t = \emptyset \)
Prior works in Gaussian context

We have the following test:\(^1\):

### GLRT for covariance homogeneity test in Gaussian context

\[
\hat{\Lambda}_G = \frac{\frac{1}{T} \sum_{t=1}^{T} \hat{\Sigma}_t}{\prod_{t=1}^{T} \hat{\Sigma}_t} \begin{bmatrix} T^N \\ N \end{bmatrix} \begin{bmatrix} H_1 \\ H_0 \end{bmatrix} > \lambda
\]

where \( \forall t, \hat{\Sigma}_t = \frac{1}{N} \sum_{k=1}^{N} x_k^t (x_k^t)^H \).

\(^1\)Many other statistics based on other principles exist as described in [Ciuonzo et al., 2017].
Some properties of the statistic

**CFARness property**

The GLRT statistic is CFAR towards the covariance parameter.

**In simulation:** $x_k^t \sim \mathcal{CN}\left(0_p, (\rho^{\|i-j\|})_{ij}\right)$ with $10^5$ Monte-Carlo trials.
Non CFAR behaviour in non-Gaussian context: Experimental results

**In simulation:** \( x^t_k = \sqrt{\tau^t_k} z^t_k \) where \( z^t_k \sim \mathcal{CN}(0_p, (0.5|i-j|)_{ij}) \) and \( \tau^t_k \sim \Gamma(\mu, b) \) with \( p = 3, N = 10, T = 3 \cdot 10^4 \) Monte-Carlo trials.

\[
\begin{align*}
\mu &= 0.10, b = 0.30 \\
\mu &= 0.30, b = 0.30 \\
\mu &= 0.90, b = 0.30
\end{align*}
\]

→ The Gaussian GLRT is **not CFAR** in the context of compound-Gaussian distributions!
Contents of the presentation

1 Introduction
   • Motivations
   • Data

2 Change detection based on a statistical framework
   • The problem
   • State of the art

3 Robust Covariance homogeneity testing for Change Detection
   • Tests using Complex Elliptical Symmetric distributions
   • Tests using deterministic compound-Gaussian modelling

4 Tests using Low-rank models

5 Clustering SAR ITS with Riemannian geometry: an opening

6 Conclusions
Complex Elliptical Symmetric modelling

Definition

A vector $\mathbf{x} \in \mathbb{C}^p$ is said to be Complex Elliptical Symmetric (CES) distributed with density generator function $g : \mathbb{R}^+ \rightarrow \mathbb{R}^+$, mean parameter $\mathbf{\mu} \in \mathbb{C}^p$ and scatter matrix parameter $\mathbf{\Sigma} \in \mathbb{S}_+^p$, denoted $\mathbf{x} \sim \mathcal{CE}(g, \mathbf{\mu}, \mathbf{\Sigma})$ if its PDF is of the following form:

$$p_x(x; \mathbf{\mu}, \mathbf{\Sigma}) = \mathcal{C}_{p,g} |\mathbf{\Sigma}|^{-1} g\left\{ (\mathbf{x} - \mathbf{\mu})^H \mathbf{\Sigma}^{-1} (\mathbf{x} - \mathbf{\mu}) \right\}.$$ (5)

<table>
<thead>
<tr>
<th>$g(t)$</th>
<th>$\mathcal{C}_{p,g}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\exp(-t)$</td>
<td>$\pi^{-p}$</td>
</tr>
<tr>
<td>$\exp(-t^s/b)$</td>
<td>$\frac{s\Gamma(p)b^{-p/s}}{\pi^p\Gamma(p/s)}$</td>
</tr>
<tr>
<td>$(1 + t/d)^{-(d+p)}$</td>
<td>$\frac{\Gamma(p + d)}{\pi^m d^p \Gamma(d)}$</td>
</tr>
<tr>
<td>$ts^{-1} \exp(-t^s/b)$</td>
<td>$\frac{s\Gamma(p)b^{-(p+s-1)/s}}{\pi^p \Gamma((p + s - 1)/s)}$</td>
</tr>
<tr>
<td>$\sqrt{1^{\nu-p}} K_{\nu-p} (2\sqrt{\nu t})$</td>
<td>$\frac{2^{\nu(\nu+p)/2}}{\pi^p \Gamma(\nu)}$</td>
</tr>
</tbody>
</table>

[Ollila et al., 2012] proposed to use elliptical distributions for modelling the clutter of HR SAR images:

$$\mathbf{x}^t_k \sim \mathcal{CE}(0_p, g, \mathbf{\Sigma}_t).$$
Some attempts using Robust estimation theory

In [Formont et al., 2011], it was proposed to bootstrap the Gaussian GLRT using the Tyler’s estimator. If we have \( N \) observations \( \{x_1, \ldots, x_N\} \) such that \( x_k \sim \mathcal{C}(g, 0_p, \Sigma) \) where:

\[
\Sigma = \sigma \xi,
\]

where \( \text{Tr}(\xi) = p \).

An estimator of \( \xi \) is:

**Tyler estimator**

\[
\hat{\xi}_t^{\text{TE}} = \frac{p}{N} \sum_{k=1}^{N} x_k^t x_k^H \frac{x_k^t x_k^H}{\xi_t^{\text{TE}} - 1 x_k^t}.
\]
Some attempts using Robust estimation theory ii

The test statistic proposed is:

\[
\hat{\Lambda}_{G,TE} = \frac{0.5(\hat{\xi}_1^{TE} + \hat{\xi}_2^{TE})}{\hat{\xi}_1^{TE} \hat{\xi}_2^{TE}} \overset{N}{\sim} \begin{pmatrix} \frac{N}{2} & \frac{N}{2} \\ \hat{\xi}_1^{TE} & \hat{\xi}_2^{TE} \end{pmatrix} \overset{H_0}{\gtrless} \lambda
\]

Problems

- The statistic is no longer CFAR matrix due to normalisation: \( \text{Tr} \left( \hat{\xi}_t^{TE} \right) = p \).
- It omits the relative scale between the matrices.

→ We need new statistics taking into account those aspect at the design stage!
Attempt of GLRT under \( CE \) model

Rather than a 2-step methodology, let’s compute the GLRT directly:

(proposed in Chapter 3) Detection test: \( \theta_t = \Sigma_t \) \& \( \Phi_t = \emptyset \)

**Problem**

We need to know the density generator \( g \) entirely in order to use the statistic. For most applications, \( g \) is **unknown** and we can’t guarantee that it stays the same over time.
Partial solution

Idea

Consider the normalized observations: \( \{z_k^t = x_k^t / \|x_k^t\|_2 \mid 1 \leq k \leq N, 1 \leq t \leq T\} \). They are known to be \( \mathcal{CAE} \) distributed.

If \( x \sim \mathcal{CE}(0_p, g, \tau\xi) \) with \( \text{Tr}(\xi) = p \). Then the PDF of \( z \) is:

\[
p_z^{\mathcal{CAE}}(z; \xi) = \mathcal{C}_p^{-1}|\xi|^{-1}(z^H \xi^{-1} z)^{-p}.
\]

(proposed) Detection test: \( \theta = \xi_t \quad \& \quad \Phi_t = \emptyset \)
The derivation of the GLRT leads to the following result:

**Proposition**

\[
\Lambda_{CAE} = \left| \begin{array}{c} \hat{\xi}_0^{TE} \\ T \end{array} \right|^{-N} \prod_{t=1}^{T} \prod_{k=1}^{N} \left( \frac{q(\hat{\xi}_t^{TE}, x_t^k)}{q(\hat{\xi}_t^{TE}, x_t^k)} \right)^{P} \frac{p(\hat{\xi}_0^{TE}; x_t^k)}{p(\hat{\xi}_t^{TE}; x_t^k)} \overset{H_1}{\gtrless} \lambda, \quad (7)
\]

where: 
\[\hat{\xi}_t^{TE} = f_t^{TE}(\hat{\xi}_t^{TE}), \hat{\xi}_0^{TE} = \frac{1}{T} \sum_{t=1}^{T} f_t^{TE}(\hat{\xi}_0^{TE}) \]

\[q(\xi, x) = x^H \xi^{-1} x \text{ and} \]

\[f_t^{TE}(\xi) = \frac{p}{N} \sum_{k=1}^{N} \frac{x_t^k x_t^k^H}{q(\xi, x_t^k)}. \quad (8)\]
### Properties

**Convergence of fixed-point estimates**

The algorithm $\hat{\xi}_0^{\text{TE}}(i+1) = \begin{cases} I_p & \text{if } i = 0 \\ \frac{1}{T} \sum_{t=1}^{T} f_t^{\text{TE}} \left( \hat{\xi}_0^{\text{TE}}(i) \right) & \text{otherwise} \end{cases}$ always converges under some regularity conditions.

Proof: Tyler estimator’s convergence is well studied [Kent and Tyler, 1988, Pascal et al., 2008].

**CFARness**

$\hat{\Lambda}_{\text{CAE}}$ is CFAR towards the shape matrix parameter under any $\mathcal{C} \mathcal{E}$ distribution.

Proof: The test statistic is invariant for the group of transformation $\mathcal{G} = \left\{ G z_k^{(t)} | t \in [1, T], k \in [1, N], G \in \mathbb{S}_H^p \right\}$. 

Ph.D defense Ammar Mian, 26 September 2019
Robust behaviour: Experimental results

In simulation: \( x_k^t = \sqrt{\tau_k^t} z_k^t \) where \( z_k^t \sim \mathcal{CN}(0_p, (0.5|i-j|)_{ij}) \) and \( \tau_k^t \sim \Gamma(\mu, b) \) with \( p = 3, N = 10, T = 3 \times 10^4 \) Monte-Carlo trials.

Problem
We still have discarded the **scale** information!
Deterministic compound-Gaussian modelling

**Definition**

A set of vectors \( \{x_1, \ldots, x_N\} \in \mathbb{C}^{p \times N} \) is said to follow a deterministic compound-Gaussian model with texture parameter \( \tau = [\tau_1, \ldots, \tau_N]^T \) and shape matrix parameter \( \xi \in S_p^p \) with \( \text{Tr}(\xi) = p \), denoted \( \mathcal{CG}(\tau, \xi) \) if each observation is distributed as follows:

\[
\mathbf{x}_k \sim \mathcal{CN}(0_p, \tau_k \xi)
\]

This model is often found in radar applications [Greco and De Maio, 2016, Pascal et al., 2004]. In the context of SAR ITS we can use the following formulation:

\[
\mathbf{x}_k^t \sim \mathcal{CG}(\tau_t, \xi_t).
\]
Deterministic compound-Gaussian modelling

Intuition behind texture:

(proposed) Detection tests:

\( \theta_t = \{ \tau_t, \xi_t \} \quad \& \quad \Phi_t = \emptyset \)

\( \theta_t = \{ \xi_t \} \quad \& \quad \Phi_t = \{ \tau_t \} \) (equivalent to \( \mathcal{CAE} \))

\( \theta_t = \{ \tau_t \} \quad \& \quad \Phi_t = \{ \xi_t \} \)
GLRT for shape matrix and texture \( (\theta_t = \{\tau_t, \xi_t\} \& \Phi_t = \emptyset) \)

**Proposition**

\[
\hat{\Lambda}_{MT} = \frac{M^T_{0}}{\prod_{t=1}^{T} |\hat{\xi}_t|} \prod_{k=1}^{N} T_{p} \left( \sum_{t=1}^{T} q \left( \hat{\xi}_{0}^{MT}, x_{t}^{k} \right) \right)^{p} \prod_{t=1}^{T} \left( q \left( \hat{\xi}_t^{TE}, x_{t}^{k} \right) \right) \]

\[10\]

where:

\[
\hat{\xi}_{0}^{MT} = \hat{f}_{N, T} \left( \hat{\xi}_{0}^{MT} \right) = \frac{p}{N} \sum_{k=1}^{N} \frac{\sum_{t=1}^{T} x_{k}^{t}(x_{k}^{t})^H}{\sum_{t=1}^{T} q \left( \xi_{0}^{MT}, x_{t}^{k} \right)}.
\]

\[11\]
GLRT for texture only ($\theta_t = \{\tau_t\} \& \Phi_t = \xi_t$)

**Proposition**

$$\hat{\Lambda}_{\text{Tex}} = \prod_{t=1}^{T} \left[ \frac{\hat{\xi}_t^{\text{Tex}}}{\hat{\xi}_t^T} \right]^{\frac{N}{T}} \prod_{k=1}^{N} \left[ \frac{\hat{\xi}_t^T}{\hat{\xi}_t} \right]^{\frac{N}{T}} \left( \sum_{t=1}^{T} q \left( \hat{\xi}_t^{\text{Tex}} , x^t_k \right) \right)^{Tp} \prod_{t=1}^{T} \left( q \left( \hat{\xi}_t^{\text{TE}} , x^t_k \right) \right)^{Tp} \quad \text{H}_1 \geq \lambda , \quad \text{p} \quad \text{H}_0$$  \hspace{1cm} (12)

where:

$$\hat{\xi}_t^{\text{Tex}} = f_{N,T,t} \left( \hat{\xi}_1^{\text{Tex}} , \ldots , \hat{\xi}_T^{\text{Tex}} \right) = \frac{Tp}{N} \sum_{k=1}^{N} \frac{x^t_k (x^t_k)^H}{\sum_{t'=1}^{T} q \left( \hat{\xi}_{t'}^{\text{Tex}} , x^t_k \right)} \quad . \quad \text{H}_1$$  \hspace{1cm} (13)

→ We proposed an alternate fixed-point implementation for this estimator.
Properties: convergence of estimates

**Global maxima**

\( \hat{\xi}_0^{MT} \) (resp. \( \hat{\xi}_t^{Tex} \)) is the argument to the global maxima of the deterministic compound-Gaussian likelihood under \( H_0 \) of problem \( \theta_t = \{\tau_t, \xi_t\} \& \Phi_t = \emptyset \) (resp. \( \theta_t = \{\tau_t\} \& \Phi_t = \{\xi_t\} \)).

Proof: Using g-convexity on the manifold \( \mathcal{M} = S^p_{\mathbb{H}} \) as studied in [Wiesel, 2012].

\[ f(q_0) \]
\[ f(q_1) \]

+ We showed that a fixed-point algorithm implementation of \( \hat{\xi}_0^{MT} \) converges under some minor regularity conditions.
Experimental validation of convergence

In simulation: $\mathbf{x}_k^t = \sqrt{\tau_k^t} \mathbf{z}_k^t$ where $\mathbf{z}_k^t \sim \mathcal{C}\mathcal{N} \left( \mathbf{0}_p, (\rho_l^{i-j})_{ij} \right)$ and $\tau_k^t \sim \Gamma(0.3, 0.1)$

with $p = 3$, $N = 10$, $T = 3 \times 10^4$ Monte-Carlo trials.

$\rho_1 = 0.08$, $\rho_2 = 0.9$, $\rho_3 = 0.1$. 
Properties: CFARness

CFARness towards shape matrix

\( \hat{\Lambda}_{MT} \) is CFAR matrix while \( \hat{\Lambda}_{Tex} \) is not due to trace normalization.

Simulation: \( x_k^t \sim \mathcal{CN} \left( 0_p, (\rho|i-j|)_{ij} \right) \).
Properties: CFARness

CFARness towards texture parameters

\( \hat{\Lambda}_{MT} \) and \( \hat{\Lambda}_{Tex} \) are CFAR texture.

Simulation: \( x_k^t = \sqrt{\tau_k} z_k^t \) where \( z_k^t \sim \mathcal{CN}(0_p, (0.3|i-j|)_{ij}) \) and \( \forall (k, t), \tau_k^t = \tau_k \)
Application to real SAR data: UAVSAR (NASA) dataset

Data description:

- Polarimetric data: $p = 3$
- Dimensions: 2360px 600px
- Resolution: 1.67 m (Range) and 0.60 m (Azimuth)
Left: Scene 1. Right: Scene 2.

Ph.d defense Ammar Mian, 26 September 2019
Concluding remarks

What we have achieved

- Robust shape matrix testing under $\mathcal{CE}$ modelling.
- Robust scale and shape matrix testing under $\mathcal{CCG}$ modelling.
  → Improved performance for real SAR images

Other things to consider

Performance of detection improves when $p$ increases (diversity) [Mian et al., 2017].

Drawback:

- $N$ has to be bigger as well!
- No spatial change hypothesis is challenged.

Solutions: Regularisation or consider structured matrices.
Contents of the presentation

1 Introduction
   • Motivations
   • Data

2 Change detection based on a statistical framework
   • The problem
   • State of the art

3 Robust Covariance homogeneity testing for Change Detection
   • Tests using Complex Elliptical Symmetric distributions
   • Tests using deterministic compound-Gaussian modelling

4 Tests using Low-rank models

5 Clustering SAR ITS with Riemannian geometry: an opening

6 Conclusions
Low-rank models

Gaussian context

\[ x \sim \mathcal{CN}(0_p, \Sigma_t + \sigma^2 I) \text{ where } \text{rank}(\Sigma_t) = R < p. \]

\[ \theta_t = \Sigma_t \quad \& \quad \Phi_t = \emptyset. \]

CCG context

Assuming \( x \sim \mathcal{CCG}(\tau_t, \xi_t + \sigma^2 I) \) where \( \text{rank}(\xi_t) = R < p. \)

\[ \theta_t = \{\tau_t, \xi_t\} \quad \& \quad \Phi_t = \emptyset. \]

Work done in collaboration with Rayen Ben Abdallah and Arnaud Breloy of LEME, Paris Nanterre University.
Optimization considerations: Gaussian case \( \theta_t = \Sigma_t \) \& \( \Phi_t = \emptyset \)

Expression of the GLRT: no closed form

\[
\hat{\Lambda}_{LRG} = \frac{\mathcal{L} \left( \mathcal{W}_1, T / H_1; \mathcal{T}_R \{ \hat{\Sigma}_1 \}, \ldots, \mathcal{T}_R \{ \hat{\Sigma}_T \} \right)}{\mathcal{L} \left( \mathcal{W}_1, T / H_0; \mathcal{T}_R \{ \hat{\Sigma}_0 \} \right)} \underset{H_1}{\overset{H_0}{\geq}} \lambda. \tag{14}
\]

where \( \hat{\Sigma}_t = \frac{1}{N} \sum_{k=1}^{N} x_k^t (x_k^t)^H \) and for \( \Sigma \overset{\text{EVD}}{=} V \Lambda V^H, \mathcal{T}_R \{ \Sigma \} \overset{\text{EVD}}{=} V \tilde{\Lambda} V^H \) with

\[
[\tilde{\Lambda}]_{i,i} = \begin{cases} 
\max([\Lambda]_{i,i}, \sigma^2) & i \leq R \\
\sigma^2 & i > R 
\end{cases} \tag{15}
\]
Optimization considerations: \( CCG \) case \((\theta_t = \{\tau_t, \xi_t\} \quad \& \quad \Phi_t = \emptyset.)\)

Same methodology for this case but the optimization of the likelihood is **non-convex**. We propose the following approach:

**Low-rank Compound-Gaussian GLRT**

\[
\hat{\Lambda}_{LRCG} = \frac{\mathcal{L} \left( W_1, T / H_1 ; \{\hat{\xi}_t, \{\hat{\tau}^t_k\}\} \right)}{\mathcal{L} \left( W_1, T / H_0 ; \{\hat{\xi}_0, \{\hat{\tau}^0_k\}\} \right)} \geq \frac{H_1}{H_0} \lambda.
\]

where the set of parameters \( \hat{\xi}_0, \hat{\xi}_t, \tau^t_k, \tau^0_k \) are estimated through a block Majorization-Minimization of the negative log-likelihood.

\( \rightarrow \) Convergence insured to a local optima.
Optimization for $\mathbf{C C G}$ case

**Algorithm 1** BCD for MLEs under $H_1$

Initialise $\xi_t = \mathbf{I}_p$

repeat

$$\tau_k^t = \frac{\left((x_k^t)^H \Sigma_t^{-1} x_k^t \right)}{p}$$

$$\xi_t = \mathcal{T}_R \left\{ \frac{1}{N} \sum_{k=1}^K x_k^t (x_k^t)^H \right\}$$

until convergence

Output: $\left\{ \{\hat{\xi}_t\}, \{\hat{\tau}_k^t\} \right\}$

**Algorithm 2** BCD for MLE under $H_0$

Initialise $\xi_0 = \mathbf{I}_p$

repeat

$$\tau_k^0 = \frac{\left(\sum_{t=1}^T (x_k^t)^H \Sigma_0^{-1} x_k^t \right)}{Tp}$$

$$\xi_0 = \mathcal{T}_R \left\{ \frac{1}{N} \sum_{k=1}^N \sum_{t=1}^T \frac{x_k^1 (x_k^1)^H}{T \tau_k^0} \right\}$$

until convergence

Output: $\left\{ \hat{\xi}_0, \{\hat{\tau}_k^0\} \right\}$
**CFARness: Gaussian case**

**Simulation setup:** \( p = 20, N = 50, T = 10, R = 5 \)
\[ x_k^t \sim \mathcal{CN} \left( 0_p, \mathcal{T}_R\{(\rho|i-j|i)|j\} \right) \] with \( 10^4 \) Monte-Carlo trials.
**CFARness: CCG case**

**Simulation setup:** \( p = 20, N = 50, T = 10, R = 5 \)

\( \mathbf{x}_k^t \sim \mathcal{N} \left( 0_p, \mathcal{T}_R \{ (\rho^{i-j})_{ij} \} \right) \) with \( 10^4 \) Monte-Carlo trials.
Experimental results: Dataset

Description

- Polarimetric data $\rightarrow$ wavelet decomp. [Mian et al., 2017] $\rightarrow p = 12$ dim. pixels
- Image size: 2360px$\times$600px
- Resolution: 1.67 m (Range) and 0.60 m (Azimuth)
- CD ground truth from [Nascimento et al., 2019]
Results with a $5 \times 5$ sliding windows: Gaussian detectors
Results with a $5 \times 5$ sliding windows: Robust detectors
**Performance curves**

**Figure 4:** Probability of detection $P_D$ versus probability of false alarm $P_{FA}$ with $(p = 12, N = 25, R = 3)$

**Figure 5:** $P_D$ versus the size of window at $P_{FA} = 5\%$ with $(p = 12, R = 3)$
Concluding remarks and perspectives

**Conclusions**

We have tests assuming low-rank structure of the matrix that:

- Reduced considerably the number of false alarms.
- Perform well with lower window size.

**Perspectives**

- Rank estimation strategies [Stoica and Selen, 2004, Terreaux et al., 2018]
- CFAR test statistic in Low-rank?
  → Random Matrix theory correction [Vallet et al., 2019].
- Other structures: Persymmetry, Toeplitz, etc
Contents of the presentation

1 Introduction
   • Motivations
   • Data

2 Change detection based on a statistical framework
   • The problem
   • State of the art

3 Robust Covariance homogeneity testing for Change Detection
   • Tests using Complex Elliptical Symmetric distributions
   • Tests using deterministic compound-Gaussian modelling

4 Tests using Low-rank models

5 Clustering SAR ITS with Riemannian geometry: an opening

6 Conclusions
Illustrative example
The SAR ITS clustering problem

**Problem:** Can we cluster pixels according to their temporal evolution?

- Pattern of changes
- Type of objects who change

**Idea**

Consider a time-series clustering problem!

**We need:** a **feature** to represent the time series and a **distance**.

Work done in collaboration with **Florent Bouchard** of University Savoie Mont-Blanc.
→ We need a descriptive feature and a distance on the feature space!
Multivariate SAR clustering: state of the art

**Spatial clustering:** cluster patches \( \{X_i | 1 \leq i \leq M\} \) where each set is defined as \( X_i = \{x_k \in \mathbb{C}^p | 1 \leq k \leq N\} \)

**Wishart Classifier ([Jong-Sen Lee et al., 1999])**

**Feature:** \( \hat{\Sigma} = N^{-1} \sum_{k=1}^{N} x_k x_k^H \).

\[
d_{\mathcal{W}}(\Sigma, C) = \log|C| - \log|\Sigma| + \text{tr}(C^{-1}\Sigma) \tag{16}\]

**Robust distance [Vasile et al., 2010]**

**Feature:** Tyler estimator.

\[
d_{\mathcal{STRV}}(\Sigma, C) = \log|C| - \log|\Sigma| + \frac{p}{N} \sum_{k=1}^{N} \frac{q(C, x_k)}{q(\Sigma, x_k)} \tag{17}\]
Clustering on temporal features

\{X_i | 1 \leq i \leq N\} where each set
\[ X_i = \{x^t \in \mathbb{C}^p | 1 \leq t \leq T\} \]

**Temporal features**

\[ \theta = \{\hat{\xi}^{TE}, \hat{\tau}\}, \]

where

\[ \hat{\xi}^{TE} = \frac{p}{N} \sum_{t=1}^{T} \frac{x^t(x^t)^H}{\{\hat{\xi}^{TE}\}^{-1}(x^t)}, \]
\[ \hat{\tau} = p^{-1}(x^t)^H \{\hat{\xi}^{TE}\}^{-1}(x^t) \]

→ We need a distance!
Riemannian Geometry for clustering

We consider **Riemannian Geometry** on the manifold

\[ \mathcal{M}_{p,T} = S^{p}_{H,J} \times (\mathbb{R}^+)^T, \]

where \( S^{p}_{H,J} = \{ \xi \in S^{p}_{H} | |\xi| = 1 \} \).

Riemannian framework have been shown to be useful for clustering when the feature lie in a manifold [Formont et al., 2011, Berthomieu et al., 2017].

**Problem**

We have to develop the geometry on this new manifold!
Proposition

The natural distance between two points \( \theta_0 = (\xi_0, \tau_0) \) and \( \theta_1 = (\xi_1, \tau_1) \) belonging to \( \mathcal{M}_{p,N} \) is given by:

\[
\delta^2_{\mathcal{M}_{p,N}} = \delta^2_{\mathcal{S}_{H,|\cdot|}} (\xi_0, \xi_1) + \delta^2_{(\mathbb{R}^+)^N} (\tau_0, \tau_1).
\]

where:

\[
\delta^2_{\mathcal{S}_{H,|\cdot|}} (\xi_0, \xi_1) = \| \log(\xi_0^{1/2} \xi_1 \xi_0^{-1/2}) \|_2^2.
\]

\[
\delta^2_{(\mathbb{R}^+)^N} (\tau_0, \tau_1) = \| \log(\tau_0^{-1} \odot \tau_1) \|_2^2.
\]

\( \rightarrow \) No crossed term between shape matrix and texture part! (choice of normalization)
Mean on $\mathcal{M}_{p,N}$

Riemannian mean

$$\theta_{\text{mean}} = \arg\min_{\theta \in \mathcal{M}_{p,N}} \sum_{m=1}^{M} \delta^2_{\mathcal{M}_{p,N}}(\theta, \theta_m)$$

Can be done separately:

- Concerning $S^p_{\mathbb{H}, \bullet}$, no closed-form [Moakher, 2005]:

$$\xi_{\text{mean}} = \arg\min_{\xi \in S^p_{\mathbb{H}, \bullet}} \sum_{m=1}^{M} \delta^2_{S^p_{\mathbb{H}, \bullet}}(\xi, \xi_m).$$

- Concerning $(\mathbb{R}^+)^N$, the geometric mean is:

$$\tau_{\text{mean}} = \left(\odot_{m=1}^{M} \tau_m\right)^{1/M}.$$
Application to SAR data

UAVSAR data (Courtesy NASA/JPL-Caltech): polarimetry ($p = 3$) and $T = 17$ images.
Wishart Classifier
Wishart Classifier with Riemannian mean
SIRV distance with Riemannian mean
Riemannian distance and mean
Contents of the presentation

1 Introduction
   ● Motivations
   ● Data

2 Change detection based on a statistical framework
   ● The problem
   ● State of the art

3 Robust Covariance homogeneity testing for Change Detection
   ● Tests using Complex Elliptical Symmetric distributions
   ● Tests using deterministic compound-Gaussian modelling

4 Tests using Low-rank models

5 Clustering SAR ITS with Riemannian geometry: an opening

6 Conclusions
**Conclusions of the Ph.d**

**What we have done**

- Found a way to construct diversity in monovariate HR SAR images
- Improved change detection algorithms for multivariate HR SAR images
- Considered change-point estimation problems as well
- Developed a framework for Riemannian clustering of time series (still working on it)

**Several collaborations**

- with **Lucien Bacharach** of SATIE, ENS Paris-Saclay and **Alexandre Renaux** of L2S, CentraleSupélec, on lower bounds for change-point
- with **Rayen Ben Abdallah** and **Arnaud Breloy** of LEME, Paris Nanterre on low-rank change detection
- with **Florent Bouchard** of LISTIC, Université Savoie Mont-Blanc on Robust Riemannian clustering
- with **Jialun Zhou, Salem Said and Yannick Berthomieux** of IMS, Bordeaux on Riemannian stochastic optimization
### Change detection and change-point estimation

- Consider other structures and improve current test statistics.
- Integrate temporal or spatial correlation in the design:
  - Typically done in Finance applications
- Improve lower-bound to take into account non-Gaussian modelling.

### Time series clustering

- Validate current methodology
- Consider a multivariate Dynamic Time Warping (DTW) framework
- Consider distance between segments on manifolds
Thanks for your attention!
Journals:


Conferences proceedings:


An invariance property of the generalized likelihood ratio test.  

Maximum likelihood estimation for the wrapped Cauchy distribution.  

Robust low-rank change detection for SAR image time series.  

New robust statistics for change detection in time series of multivariate SAR images.  


