Robust statistics for testing the homogeneity of covariance scale and shape

French Finish Workshop 07/12/2018

Ammar Mian

December 7, 2018

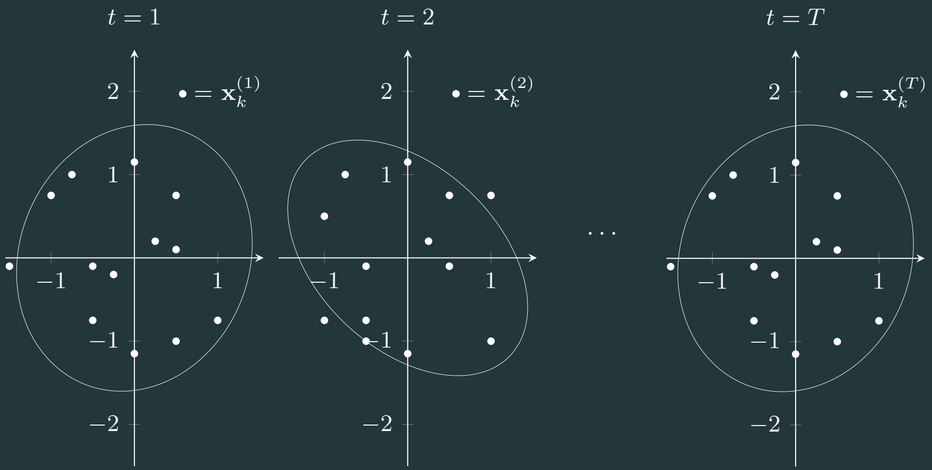
CentraleSupélec SONDRA, LISTIC

Table of contents

1. The problem of testing the homogeneity of covariances
   - Illustration
   - Statistics of decision under Gaussian model
   - Non robust behaviour
2. Testing homogeneity under Complex Elliptical model
   - Approach 1: Bootstrapping the Gaussian GLRT
   - Approach 2: GLRT using the elliptical model
   - Approach 2bis: GLRT under SIRV model
3. Application to SAR change detection

The problem of testing the homogeneity of covariances

The problem under consideration



How to compare these groups of observations ?

→ Idea: use the covariance matrix as a descriptor and test homogeneity.

The setting

- Denote by $\{\mathbf{X}_1, \ldots, \mathbf{X}_T\}$ a collection of T mutually independent samples of i.i.d p-dimensional complex vectors: $\mathbf{X}_t = [\mathbf{x}_1^{(t)}, \ldots, \mathbf{x}_N^{(t)}]$.
- We assume $\forall (k,t)$, $\mathbb{E}\{\mathbf{x}_k^{(t)}\} = \mathbf{0}_p$ and we denote $\boldsymbol{\Sigma}_t = \tau_t \boldsymbol{\xi}_t$ the shared covariance matrices among the elements of $\mathbf{X}_t$. $\boldsymbol{\xi}_t$ is the shape matrix $(\mathrm{Tr}(\boldsymbol{\xi}_t) = p)$ and τ_t is the scale.
- Suppose the observations are distributed under $\mathbb{CES}$ model:

$$p_{\mathbf{x}}(\mathbf{x}; \boldsymbol{\Sigma}, g) = \mathfrak{C}_{p,g} |\boldsymbol{\Sigma}|^{-1} g\left(\mathbf{x}^{\mathrm{H}} \boldsymbol{\Sigma}^{-1} \mathbf{x}\right) \tag{1}$$

We want to decide between the following assumptions:

$$\begin{cases} \mathrm{H}_0 : & \boldsymbol{\Sigma}_1 = \ldots = \boldsymbol{\Sigma}_T = \boldsymbol{\Sigma}_0 \,, \\ \mathrm{H}_1 : & \exists (t, t') \in [\![1, T]\!]^2, \, \boldsymbol{\Sigma}_t \neq \boldsymbol{\Sigma}_{t'} \end{cases}$$

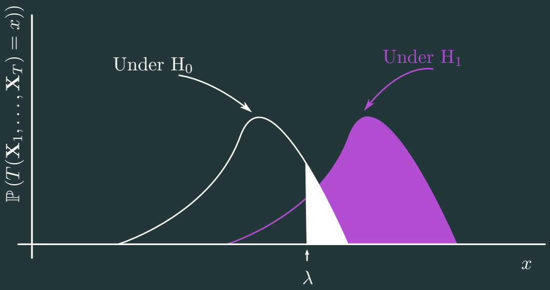
Statistic of decision

We want to obtain:

- a statistic of decision T: $$\begin{aligned}\mathbb{C}^{p\times N}\times\cdots\times\mathbb{C}^{p\times N} &\to \mathbb{R}^+\\ \mathbf{X}_1,\ldots,\mathbf{X}_T &\to T(\mathbf{X}_1,\ldots,\mathbf{X}_T)\end{aligned}$$
- a threshold λ

So that

$\mathbb{P}\left(T(\mathbf{X}_1,\ldots,\mathbf{X}_T) > \lambda/\mathrm{H}_1\right)$ is high while $\mathbb{P}\left(T(\mathbf{X}_1,\ldots,\mathbf{X}_T) > \lambda/\mathrm{H}_0\right)$ is low.



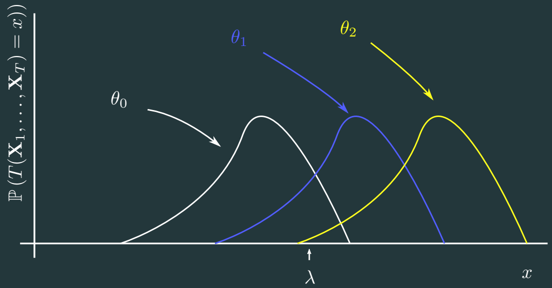
Constant False Alarm Rate (CFAR) property

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

Definition

A statistic T is said to be CFAR if for any set $(\boldsymbol{\theta}_0, \boldsymbol{\theta}_1)$,
$\mathbb{P}\left(T(\mathbf{X}_1, \ldots, \mathbf{X}_T; \boldsymbol{\theta}_0/\mathrm{H}_0) = x\right) = \mathbb{P}\left(T(\mathbf{X}_1, \ldots, \mathbf{X}_T; \boldsymbol{\theta}_1/\mathrm{H}_0) = x\right)$

Example of a non CFAR statistic:



Gaussian model i

Suppose $\forall t$, $\forall k$, $\mathbf{x}_k^{(t)} \sim \mathbb{C}\mathcal{N}(\mathbf{0}_p, \boldsymbol{\Sigma}_t)$ so that $p_{\mathbf{x}_k^{(t)};\boldsymbol{\Sigma}_t}(\mathbf{x}_k^{(t)}; \boldsymbol{\Sigma}_t) = \dfrac{1}{\pi^p |\boldsymbol{\Sigma}_t|} \operatorname{exptr}\left\{\mathbf{S}_k^{(t)} \boldsymbol{\Sigma}_t^{-1}\right\}$,

where $\mathbf{S}_k^{(t)} = \mathbf{x}_k^{(t)} {\mathbf{x}_k^{(t)}}^{\mathrm{H}}$.

Many statistic exists but the options can be reduced to [Ciuonzo et al., 2017]):

- the GLRT statistic:

$$\hat{\Lambda}_{\mathrm{G}} = \frac{\left|\hat{\boldsymbol{\Sigma}}_0^{\mathrm{SCM}}\right|^{TN}}{\prod_{t=1}^{T} \left|\hat{\boldsymbol{\Sigma}}_t^{\mathrm{SCM}}\right|^{N}} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \lambda, \tag{2}$$

where:

$$\forall t,\ \hat{\boldsymbol{\Sigma}}_t^{\mathrm{SCM}} = \frac{1}{N} \sum_{k=1}^{N} \mathbf{S}_k^{(t)} \text{ and } \hat{\boldsymbol{\Sigma}}_0^{\mathrm{SCM}} = \frac{1}{T} \sum_{t=1}^{T} \hat{\boldsymbol{\Sigma}}_t^{\mathrm{SCM}}. \tag{3}$$

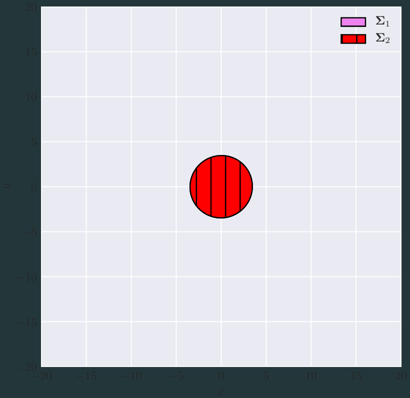
Gaussian model ii

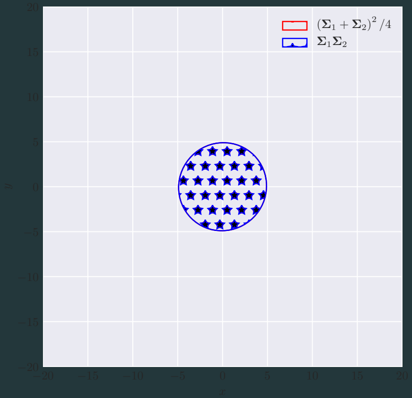
For the univariate case Bartlett noticed in [Bar, 1937], that for small sample size N, undesirable effects may arise which make the test biased.

A slight correction lead to an unbiased statistic:

$$\hat{\Lambda}_{\mathrm{B}} = \frac{\left|\hat{\boldsymbol{\Sigma}}_0^{\mathrm{SCM}}\right|^{T(N-1)}}{\prod_{t=1}^{T}\left|\hat{\boldsymbol{\Sigma}}_t^{\mathrm{SCM}}\right|^{N-1}} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \lambda, \tag{4}$$

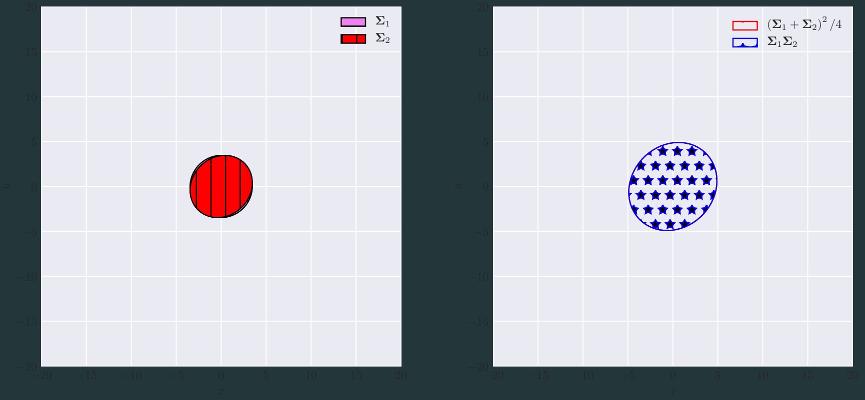
Visualisation of GLRT statistic for the case $T = p = 2$





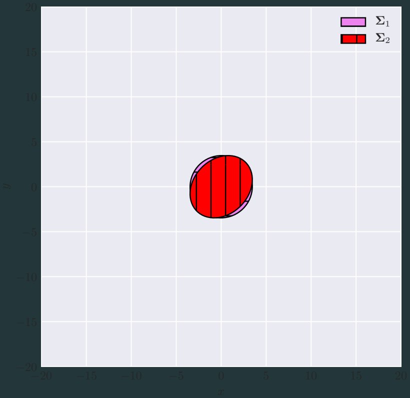
$$\mathbf{\Sigma}_1 = \begin{bmatrix} 1 & 0.01 \\ 0.01 & 1 \end{bmatrix} \quad \mathbf{\Sigma}_2 = \begin{bmatrix} 1 & 0.01 \\ 0.01 & 1 \end{bmatrix}$$

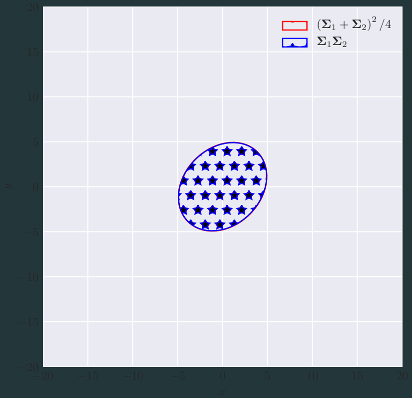
Visualisation of GLRT statistic for the case $T = p = 2$



$$\boldsymbol{\Sigma}_1 = \begin{bmatrix} 1 & 0.01 \\ 0.01 & 1 \end{bmatrix} \quad \boldsymbol{\Sigma}_2 = \begin{bmatrix} 1 & 0.12 \\ 0.12 & 1 \end{bmatrix}$$

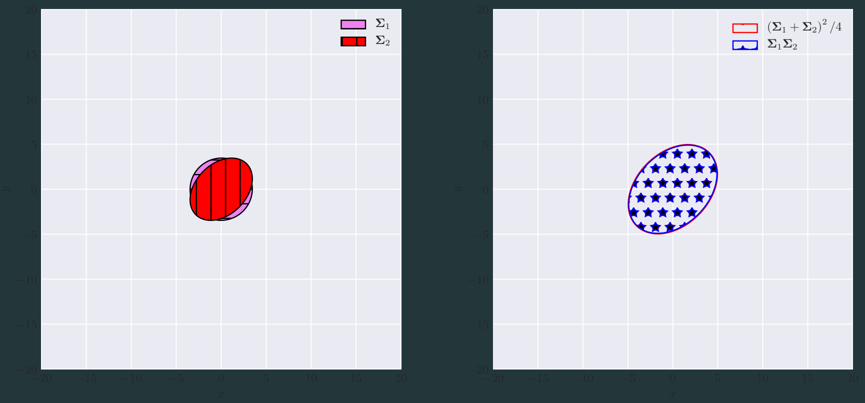
Visualisation of GLRT statistic for the case $T = p = 2$





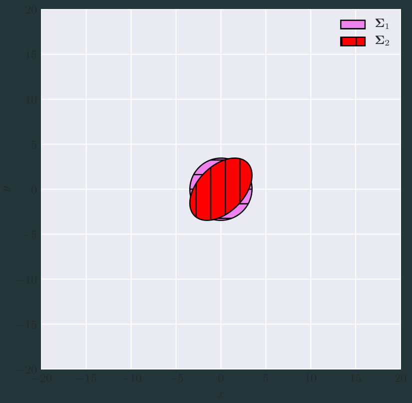
$$\boldsymbol{\Sigma}_1 = \begin{bmatrix} 1 & 0.01 \\ 0.01 & 1 \end{bmatrix} \quad \boldsymbol{\Sigma}_2 = \begin{bmatrix} 1 & 0.23 \\ 0.23 & 1 \end{bmatrix}$$

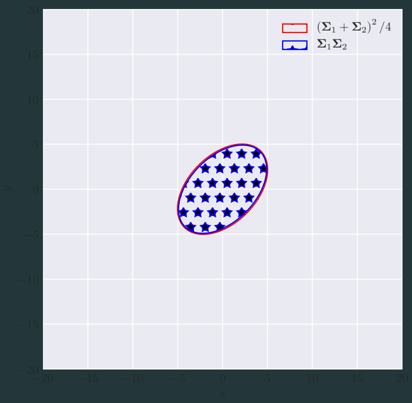
Visualisation of GLRT statistic for the case $T = p = 2$



$$\mathbf{\Sigma}_1 = \begin{bmatrix} 1 & 0.01 \\ 0.01 & 1 \end{bmatrix} \quad \mathbf{\Sigma}_2 = \begin{bmatrix} 1 & 0.34 \\ 0.34 & 1 \end{bmatrix}$$

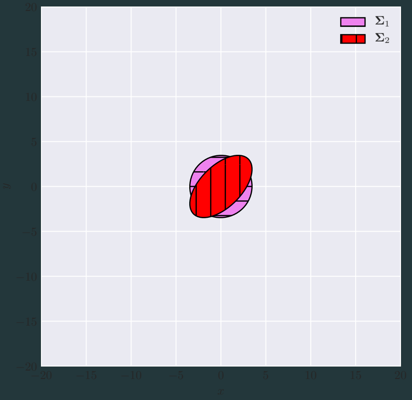
Visualisation of GLRT statistic for the case $T = p = 2$

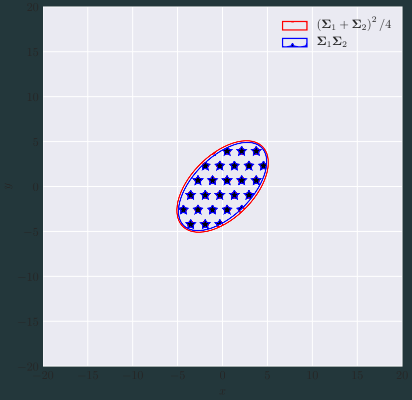




$$\Sigma_1 = \begin{bmatrix} 1 & 0.01 \\ 0.01 & 1 \end{bmatrix} \quad \Sigma_2 = \begin{bmatrix} 1 & 0.45 \\ 0.45 & 1 \end{bmatrix}$$

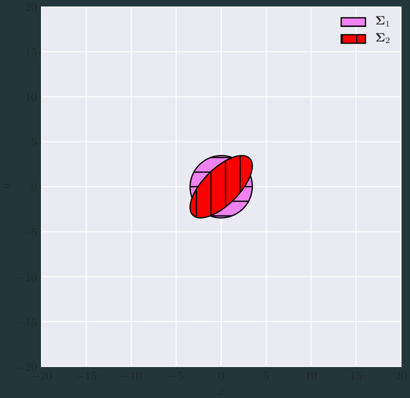
Visualisation of GLRT statistic for the case $T = p = 2$

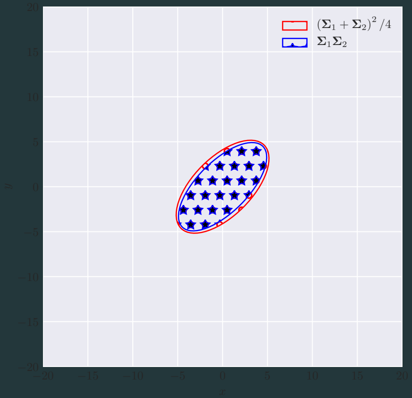




$$\boldsymbol{\Sigma}_1 = \begin{bmatrix} 1 & 0.01 \\ 0.01 & 1 \end{bmatrix} \quad \boldsymbol{\Sigma}_2 = \begin{bmatrix} 1 & 0.55 \\ 0.55 & 1 \end{bmatrix}$$

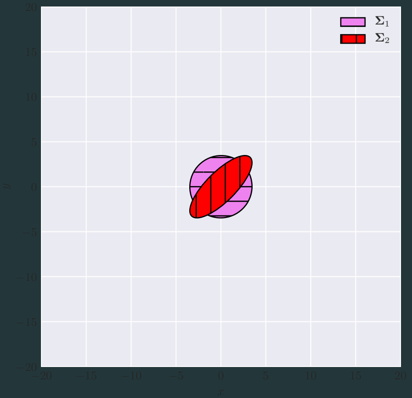
Visualisation of GLRT statistic for the case $T = p = 2$

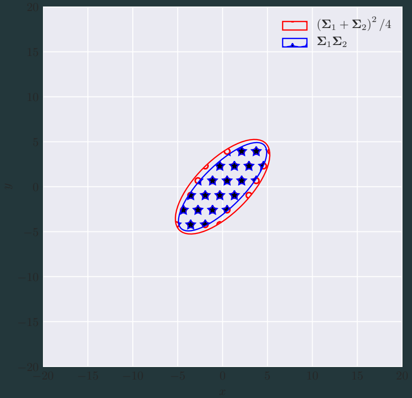




$$\boldsymbol{\Sigma}_1 = \begin{bmatrix} 1 & 0.01 \\ 0.01 & 1 \end{bmatrix} \quad \boldsymbol{\Sigma}_2 = \begin{bmatrix} 1 & 0.66 \\ 0.66 & 1 \end{bmatrix}$$

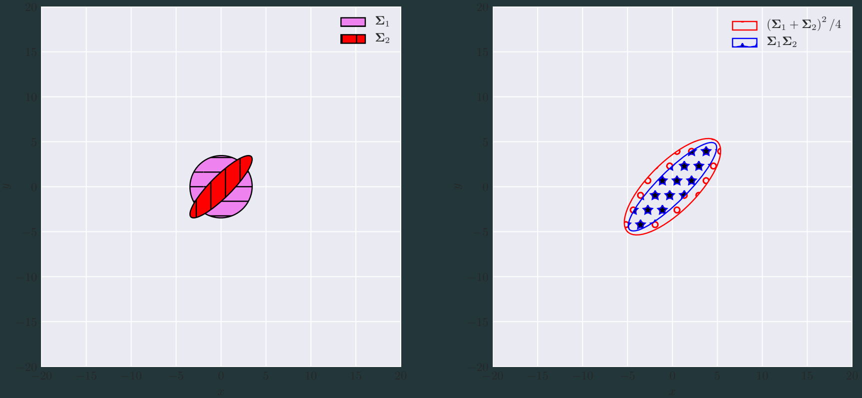
Visualisation of GLRT statistic for the case $T = p = 2$





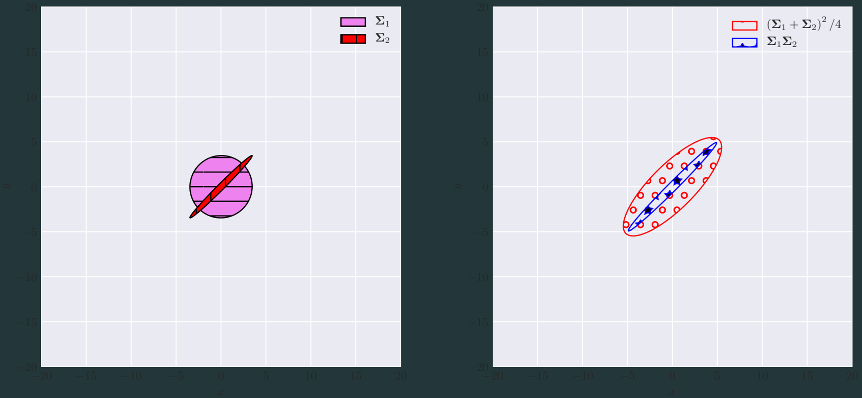
$$\boldsymbol{\Sigma}_1 = \begin{bmatrix} 1 & 0.01 \\ 0.01 & 1 \end{bmatrix} \quad \boldsymbol{\Sigma}_2 = \begin{bmatrix} 1 & 0.77 \\ 0.77 & 1 \end{bmatrix}$$

Visualisation of GLRT statistic for the case $T = p = 2$



$$\boldsymbol{\Sigma}_1 = \begin{bmatrix} 1 & 0.01 \\ 0.01 & 1 \end{bmatrix} \quad \boldsymbol{\Sigma}_2 = \begin{bmatrix} 1 & 0.88 \\ 0.88 & 1 \end{bmatrix}$$

Visualisation of GLRT statistic for the case $T = p = 2$



$$\boldsymbol{\Sigma}_1 = \begin{bmatrix} 1 & 0.01 \\ 0.01 & 1 \end{bmatrix} \quad \boldsymbol{\Sigma}_2 = \begin{bmatrix} 1 & 0.99 \\ 0.99 & 1 \end{bmatrix}$$

Other statistics

- the t_1 statistic which is obtained from Terrell [Terrell, 2002] or Rao [Radhakrishna Rao, 1948] tests:

$$\hat{\Lambda}_{\mathrm{t}_1} = N \sum_{t=1}^{T} \mathrm{Tr}\left[\left(\hat{\mathbf{\Sigma}}_t^{\mathrm{SCM}}\left(\hat{\mathbf{\Sigma}}_0^{\mathrm{SCM}}\right)^{-1} - \mathbf{I}_p\right)^2\right] \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \lambda. \tag{5}$$

- the Wald statistic [Wald, 1943]:

$$\begin{aligned}\hat{\Lambda}_{\mathrm{Wald}} =& N \sum_{t=2}^{T} \mathrm{Tr}\left[\left(\mathbf{I}_p - \hat{\mathbf{\Sigma}}_1^{\mathrm{SCM}}(\hat{\mathbf{\Sigma}}_t^{\mathrm{SCM}})^{-1}\right)^2\right] \\ & - q\left(N \sum_{t=1}^{T} (\hat{\mathbf{\Sigma}}_t^{\mathrm{SCM}})^{-T} \otimes (\hat{\mathbf{\Sigma}}_t^{\mathrm{SCM}})^{-1}, \mathrm{vec}\left(\sum_{t=2}^{T} \mathbf{\Upsilon}_t\right)\right) \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \lambda,\end{aligned} \tag{6}$$

$$\text{where } \mathbf{\Upsilon}_t = N\left((\hat{\mathbf{\Sigma}}_t^{\mathrm{SCM}})^{-1} - (\hat{\mathbf{\Sigma}}_t^{\mathrm{SCM}})^{-1}\hat{\mathbf{\Sigma}}_1^{\mathrm{SCM}}(\hat{\mathbf{\Sigma}}_t^{\mathrm{SCM}})^{-1}\right).$$

$$q(\mathbf{x}, \mathbf{\Sigma}) = \mathbf{x}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\mathbf{x}$$

CFARness properties:

Proposition

The GLRT, Bartlett, t_1 and Wald statistic are CFAR matrix.

Proof: The statistics are invariant for the group of transformation
$\mathcal{G} = \left\{ \mathbf{G}\,\mathbf{x}_k^{(t)} | t \in [\![1, T]\!],\, k \in [\![1, N]\!],\, \mathbf{G} \in \mathbb{S}_{\mathbb{H}}^{p} \right\}.$

Assymptotic distribution (F-approximation [Box, 1949]):

Proposition

Under H_0 *hypothesis, we have:* $2(1-c)\ln(\hat{\Lambda}_{\mathrm{B}}) \sim \chi^2\left((T-1)p(p+1)\right)$, *where*

$$c = \frac{T^2 - 1}{T(N-1)} \times \frac{2p^2 + 3p - 1}{6(T-1)(p+1)}$$

False alarm/threshold relationship:

Assymptotic expansion

Under H_0 hypothesis:

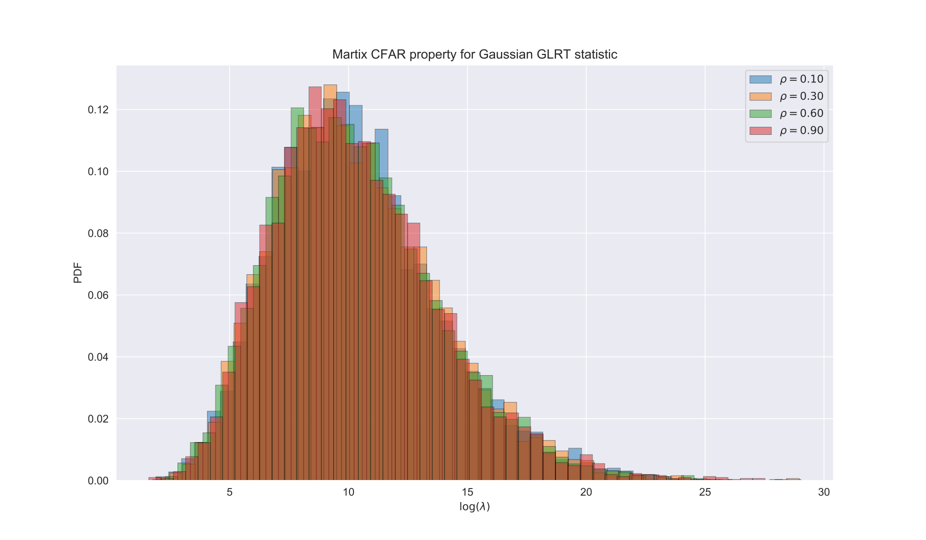
$$P\left\{2\rho\log(\hat{\Lambda}_{\mathrm{G}}) \leq z\right\} \approx P\left\{\chi^2(f^2) \leq z\right\} + \omega_2\left[P\left\{\chi^2(f^2+4) \leq z\right\} - P\left\{\chi^2(f^2) \leq z\right\}\right]$$

$$f = (T-1)p^2,\ \rho = 1 - \frac{(2p^2-1)}{6(T-1)p}\left(\frac{T}{N} - \frac{1}{NT}\right),$$

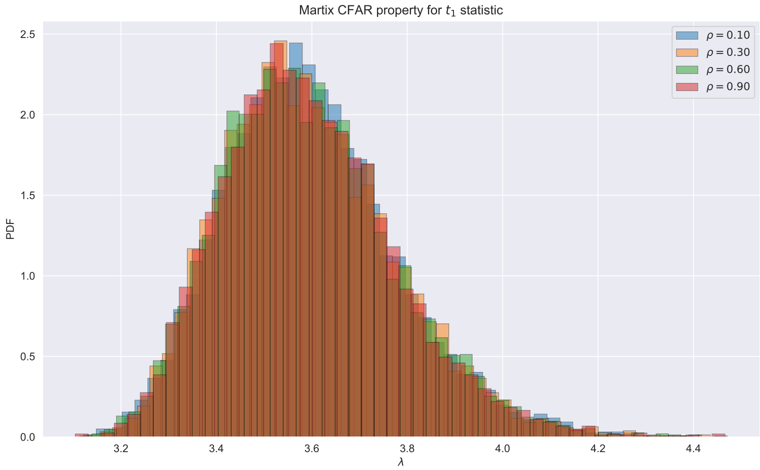
$$\omega_2 = \frac{p^2(p^2-1)}{24\rho^2}\left(\frac{T}{N^2} - \frac{1}{(NT)^2}\right) - \frac{p^2(T-1)}{4}\left(1-\frac{1}{\rho}\right)^2$$

Proof: See [Anderson, 2003] for real case or [Conradsen et al., 2003] for complex one.

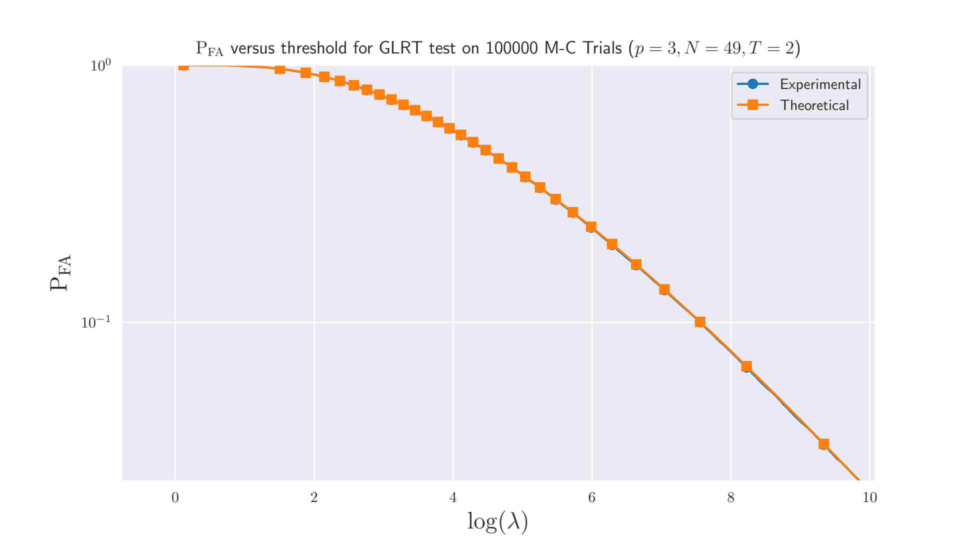
Experimental validation of matrix CFAR property



Experimental validation of matrix CFAR property

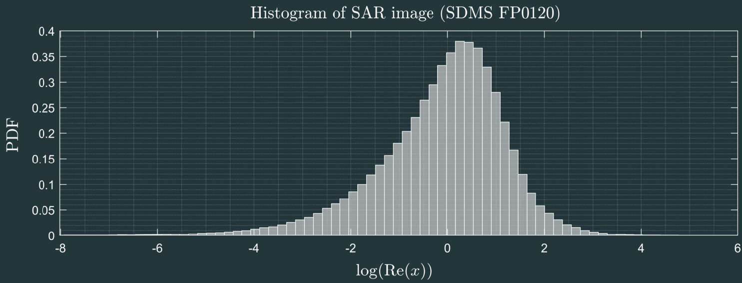


Experimental validation of F-approximation



Non Gaussian data

In some case the Gaussian assumption is not accurate enough:



Suppose now that the Gaussian model does not match the distribution of the data. What is the behavior of the Gaussian-derived statistics in the context of $\mathbb{CES}$?

Non robust behaviour

One result:

Proposition (Real case) [Yanagihara et al., 2005]

Assuming an elliptical model with an equality of kurtosis between the T groups of observations (homokurticity), the assymptotic distribution of $\hat{\Lambda}_{\mathrm{B}}$ is:

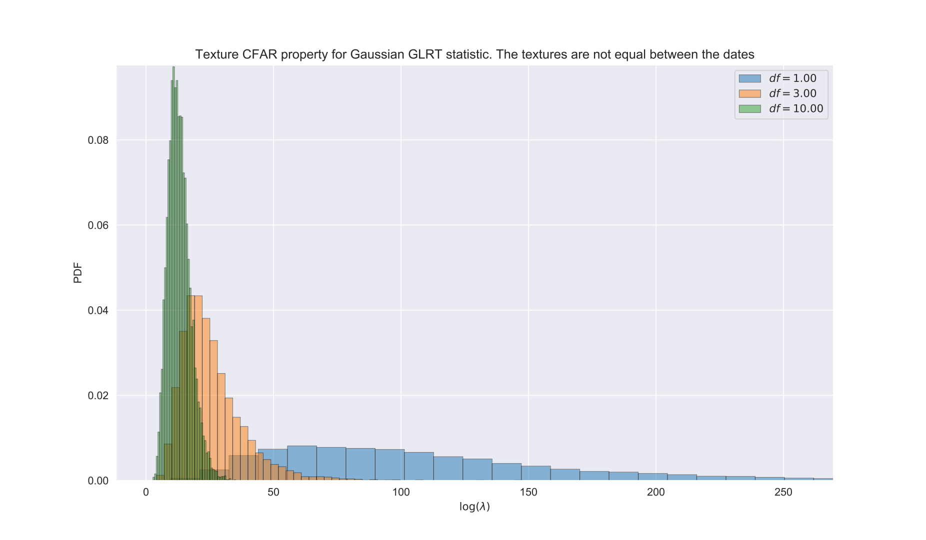
$$(1+\kappa_p)\left\{\left[1+\frac{p\kappa_p}{2(1+\kappa_p)}\right]\chi^2(T-1)+\chi^2\left((T-1)(p-1)\frac{p+2}{2}\right)\right\}, \tag{7}$$

where κ_p is the common kurtosis of the T groups of samples.

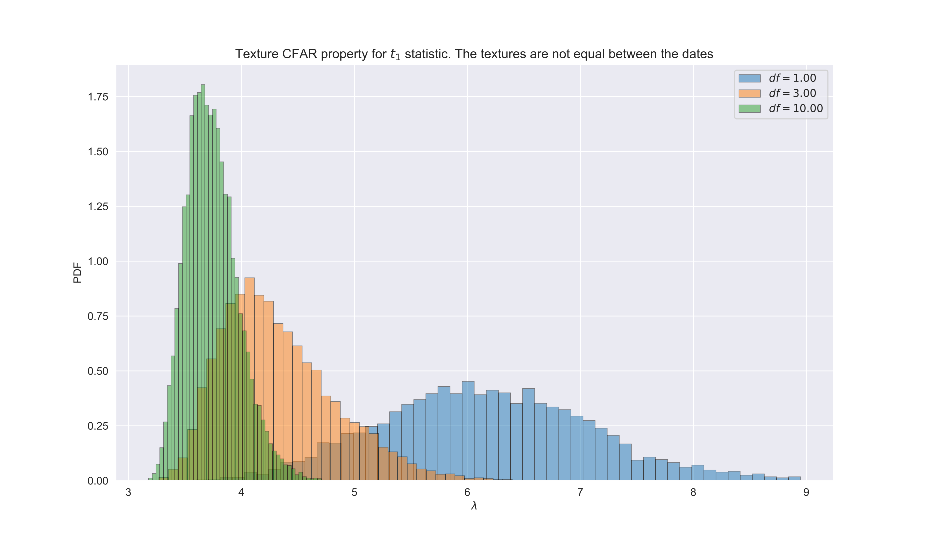
For the Gaussian case we keep the chi-square approximation but in elliptical case, the kurtosis depends on the distribution.

$\rightarrow$ The statistic is not CFAR.

Non Robust behaviour: Experimental results on a Student-t distribution



Non Robust behaviour: Experimental results on a Student-t distribution



Testing homogeneity under Complex Elliptical model

Approach 1: Bootstrapping the Gaussian GLRT

Since the distribution of Gaussian-derived detector is known it is possible to correct it in order to obtain one keeping the CFAR property in $\mathbb{CES}$ context.

Hallin proposed in [Hallin and Paindaveine, 2009] to use the following test:

Pseudo-Gaussian test (homokurtic real case)

$$\mathcal{Q}_{\mathcal{N}} = 2 \sum_{1 \leq t < t' \leq T} \mathcal{Q}_{\mathcal{N};t,t'} \geq \chi^2 \left(\frac{(T-1)p(p+1)}{2} \right)_{1-\mathrm{P_{FA}}}, \text{where} \tag{8}$$

$$\mathcal{Q}_{\mathcal{N};t,t'} = \frac{1}{4(1+\hat{\kappa}_p)} \left\{ \mathrm{Tr}\left[(\hat{\mathbf{\Sigma}}_0^{\mathrm{SCM}})^{-1} (\hat{\mathbf{\Sigma}}_t^{\mathrm{SCM}} - \hat{\mathbf{\Sigma}}_{t'}^{\mathrm{SCM}})^2 \right] - \right.$$
$$\left. \frac{\hat{\kappa}_p}{(p+2)\hat{\kappa}_p + 2} \mathrm{Tr}^2 \left[(\hat{\mathbf{\Sigma}}_0^{\mathrm{SCM}})^{-1} (\hat{\mathbf{\Sigma}}_t^{\mathrm{SCM}} - \hat{\mathbf{\Sigma}}_{t'}^{\mathrm{SCM}}) \right] \right\},$$

$\hat{\kappa}_p = p(p+1)/2 \sum_{t=1}^{T} \sum_{k=1}^{N} d^4(\mathbf{x}_k^{(t)}, \hat{\mathbf{\Sigma}}_0^{\mathrm{SCM}}) - 1$ and $d(\mathbf{x}, \mathbf{\Sigma}) = \|\mathbf{\Sigma}^{-1/2}\mathbf{x}\|$

$\rightarrow$ SCM not very robust to outliers compared to others (M-estimators, Tyler estimator).

Approach 2: GLRT using the elliptical model

We compute the following statistic:

$$\hat{\Lambda} = \frac{\max\limits_{\boldsymbol{\Sigma}_1,\ldots,\boldsymbol{\Sigma}_T} p_{\mathbf{X}_1,\ldots,\mathbf{X}_T;\boldsymbol{\Sigma}_1,\ldots,\boldsymbol{\Sigma}_T}\left(\mathbf{X}_1,\ldots,\mathbf{X}_T;\boldsymbol{\Sigma}_1,\ldots,\boldsymbol{\Sigma}_T/\mathrm{H}_1\right)}{\max\limits_{\boldsymbol{\Sigma}_0} p_{\mathbf{X}_1,\ldots,\mathbf{X}_T;\boldsymbol{\Sigma}_0}\left(\mathbf{X}_1,\ldots,\mathbf{X}_T;\boldsymbol{\Sigma}_0/\mathrm{H}_0\right)} \tag{9}$$

which yields:

$$\hat{\Lambda}^g_{\mathcal{R}} = \frac{\left|\hat{\boldsymbol{\Sigma}}_0^{\mathrm{M}}\right|^{TN}}{\prod\limits_{t=1}^{T}\left|\hat{\boldsymbol{\Sigma}}_t^{\mathrm{M}}\right|^{N}} \prod_{t=1}^{T}\prod_{k=1}^{N} \frac{g\left({\mathbf{x}_k^{(t)}}^{\mathrm{H}}\{\hat{\boldsymbol{\Sigma}}_0^{\mathrm{M}}\}^{-1}\mathbf{x}_k^{(t)}\right)}{g\left({\mathbf{x}_k^{(t)}}^{\mathrm{H}}\{\hat{\boldsymbol{\Sigma}}_t^{\mathrm{M}}\}^{-1}\mathbf{x}_k^{(t)}\right)} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \lambda, \tag{10}$$

where:

$$\hat{\boldsymbol{\Sigma}}_t^{M} = f_t(\hat{\boldsymbol{\Sigma}}_t^{M}),\ \hat{\boldsymbol{\Sigma}}_0^{M} = \frac{1}{T}\sum_{t=1}^{T} f_t(\hat{\boldsymbol{\Sigma}}_0^{M}) \text{ and } f_t(\boldsymbol{\Sigma}) = \frac{1}{N}\sum_{k=1}^{N} \frac{-g'\left({\mathbf{x}_k^{(t)}}^{\mathrm{H}}\{\hat{\boldsymbol{\Sigma}}\}^{-1}\mathbf{x}_k^{(t)}\right)}{g\left({\mathbf{x}_k^{(t)}}^{\mathrm{H}}\{\hat{\boldsymbol{\Sigma}}\}^{-1}\mathbf{x}_k^{(t)}\right)} \mathbf{x}_k^{(t)}{\mathbf{x}_k^{(t)}}^{\mathrm{H}}$$

Proposition

For any function $u = g'/g$ respecting some regularity conditions (Maronna's), $\hat{\Sigma}_t$ and $\hat{\Sigma}_0$ can be computed using a fixed-point algorithm.

Sketch of the proof:

- The log-likelihood is g-convex and the maximum occurs in the interior of $\mathbb{S}^p_{\mathbb{H}}$. The proposed estimators correspond to the arguments to this global maximum.
- Using Maronna's conditions, the convergence of the recursive algorithm can be shown.

Proposition

The distribution of $2\log(\hat{\Lambda}^g_{\mathcal{R}})$ under H_0 is assymptotically that of a $\chi^2\left((T-1)p(p+1)\right)$

Proof: Property inherited from the GLRT (Wilk's Phenomenon)

$\rightarrow$ Problem: g needs to be known to compute the statistic.

Limit case

A limit case can be obtained using $u(x) = 1/x$. The statistic reads:

$$\hat{\Lambda}_{\mathcal{R}}^{1/x^p} = \frac{\left|\hat{\boldsymbol{\Sigma}}_0^{\mathrm{M}}\right|^{TN}}{\prod_{t=1}^{T}\left|\hat{\boldsymbol{\Sigma}}_t^{\mathrm{M}}\right|^{N}} \prod_{t=1}^{T}\prod_{k=1}^{N} \frac{\left(\mathbf{x}_k^{(t)\mathrm{H}}\{\hat{\boldsymbol{\Sigma}}_t^{\mathrm{M}}\}^{-1}\mathbf{x}_k^{(t)}\right)^p}{\left(\mathbf{x}_k^{(t)\mathrm{H}}\{\hat{\boldsymbol{\Sigma}}_0^{\mathrm{M}}\}^{-1}\mathbf{x}_k^{(t)}\right)^p} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \lambda, \tag{11}$$

Even though Maronna's conditions are not respected, the previous results still apply.

Proposition [Mian et al., 2018b]

$\hat{\Lambda}_{\mathcal{R}}^{1/x}$ is valid to test an homogeneity in the shape matrix of any $\mathbb{CES}$ distribution.

$\rightarrow$ In order to test the scale, we must consider another approach

Approach 2bis: GLRT under SIRV model [Mian et al., 2018a]

We consider the Spherically Invariant Random Vectors (SIRV) model:

$\mathbf{x}_k^{(t)} \sim \sqrt{\tau_k^{(t)}}\mathbb{C}\mathcal{N}(\mathbf{0}_p, \boldsymbol{\xi}_t)$, where $\tau_k^{(t)}$ are assumed to be *deterministic* and $\mathrm{Tr}(\boldsymbol{\xi})$=p.

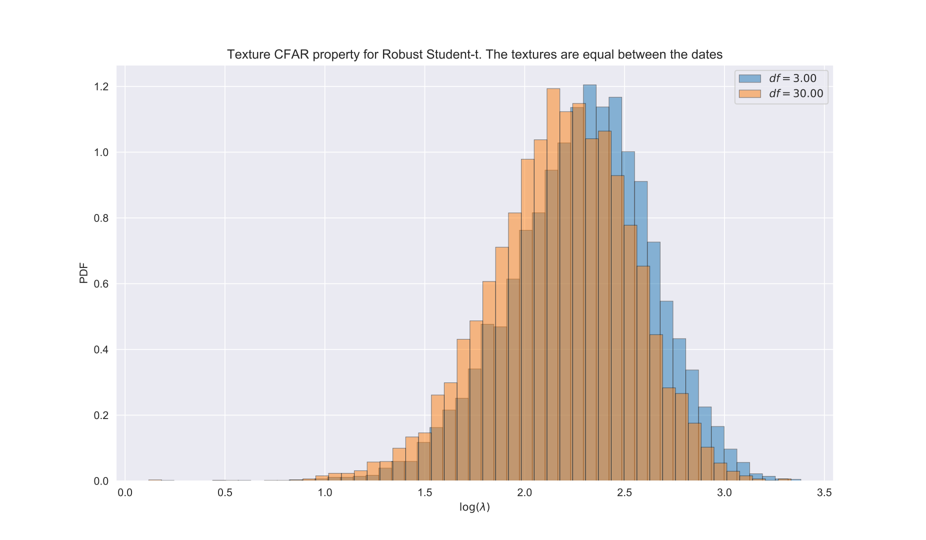
→ We can integrate these unknown parameters in the GLRT:

$$\hat{\Lambda}_{\mathrm{MT}} = \frac{\left|\hat{\boldsymbol{\Sigma}}_0^{\mathrm{MT}}\right|^{TN}}{\prod_{t=1}^{T}\left|\hat{\boldsymbol{\Sigma}}_t^{\mathrm{MT}}\right|^{N}} \prod_{k=1}^{N} \frac{\left(\sum_{t=1}^{T} \mathbf{x}_k^{(t)\mathrm{H}} \{\hat{\boldsymbol{\Sigma}}_t^{\mathrm{MT}}\}^{-1} \mathbf{x}_k^{(t)}\right)^{Tp}}{\prod_{t=1}^{T}\left(\mathbf{x}_k^{(t)\mathrm{H}} \{\hat{\boldsymbol{\Sigma}}_0^{\mathrm{MT}}\}^{-1} \mathbf{x}_k^{(t)}\right)^{p}} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \lambda, \text{ where :} \tag{12}$$

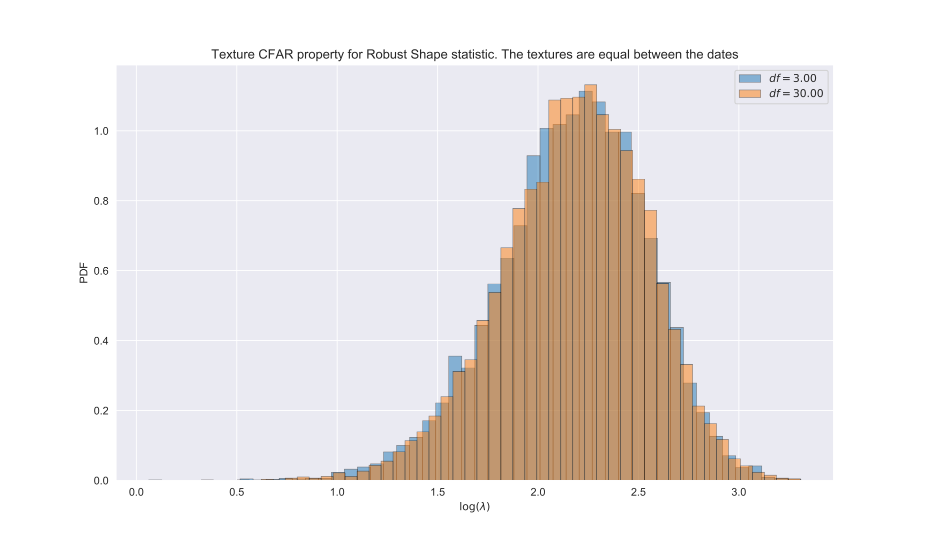
$$\hat{\boldsymbol{\Sigma}}_t^{\mathrm{MT}} = \frac{p}{N}\sum_{k=1}^{N} \frac{\mathbf{x}_k^{(t)}\mathbf{x}_k^{(t)\mathrm{H}}}{\mathbf{x}_k^{(t)\mathrm{H}}\{\hat{\boldsymbol{\Sigma}}^{\mathrm{MT}}\}^{-1}\mathbf{x}_k^{(t)}} \text{ and } \hat{\boldsymbol{\Sigma}}_0^{\mathrm{MT}} = \frac{p}{N}\sum_{k=1}^{N} \frac{\sum_{t=1}^{T}\mathbf{x}_k^{(t)}\mathbf{x}_k^{(t)\mathrm{H}}}{\sum_{t=1}^{T}\mathbf{x}_k^{(t)\mathrm{H}}\{\hat{\boldsymbol{\Sigma}}^{\mathrm{MT}}\}^{-1}\mathbf{x}_k^{(t)}}$$

→ Convergence, uniqueness, CFARness are shown but no assymptotic distribution for now

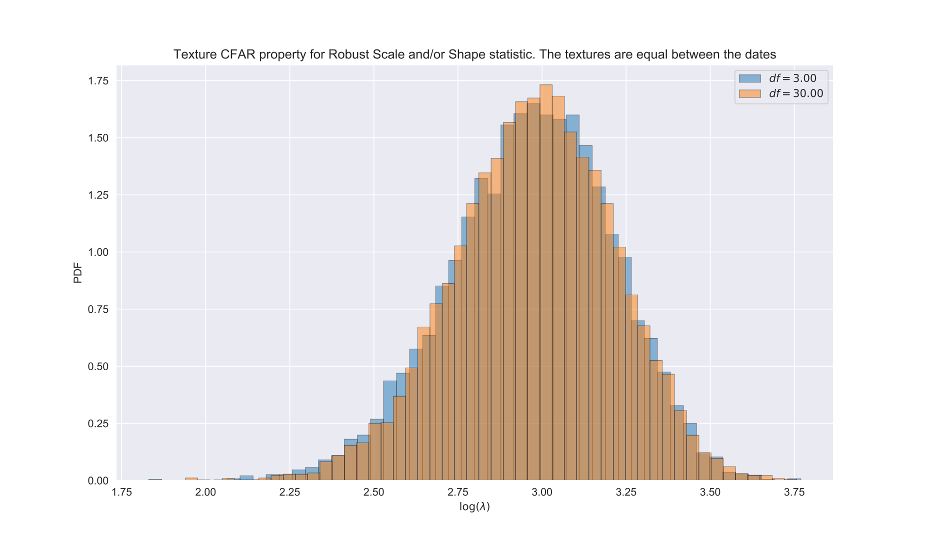
Robust behaviour: Experimental results for Student t GLRT



Robust behaviour: Experimental results for limit case

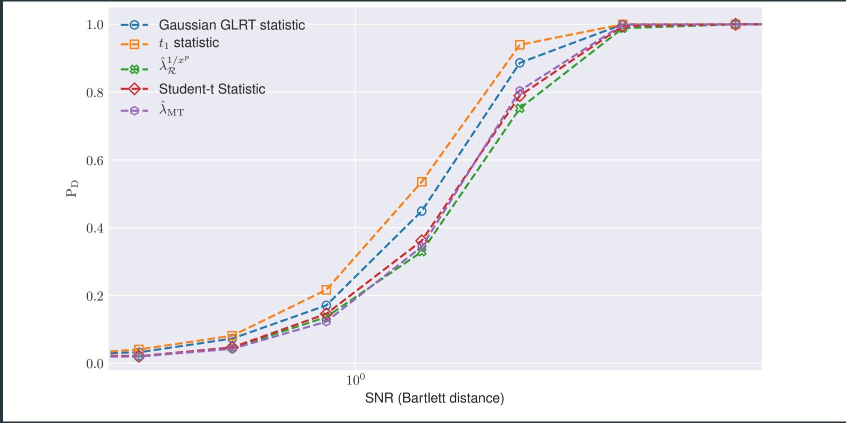


Robust behaviour: Experimental results for SIRV GLRT



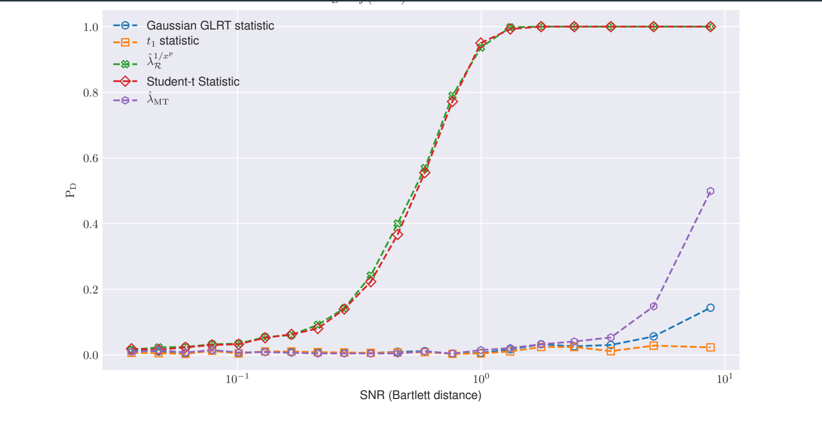
Some performance

$T = 10$, $p = 10$, $N = 25$, $\boldsymbol{\Sigma}_{t=1,\dots,5} = \boldsymbol{\xi}$ and $\boldsymbol{\Sigma}_{t=5,\dots,10} = \boldsymbol{\xi}'$, Gaussian data and change in shape. $\mathrm{P_{FA}} = 10^{-2}$.



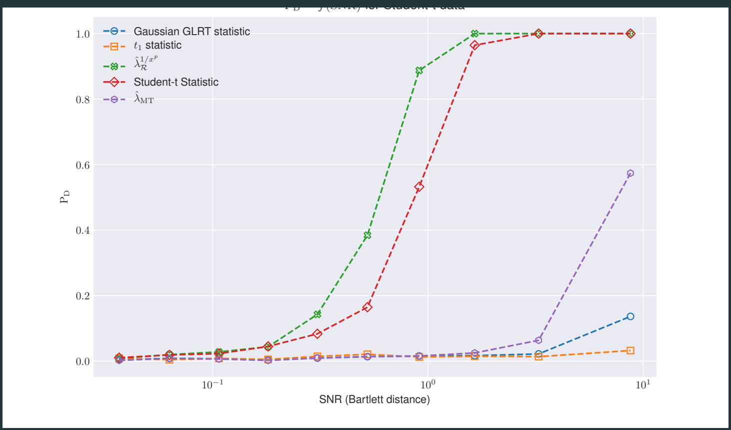
Some performance

$T = 10$, $p = 10$, $N = 25$, $\mathbf{\Sigma}_{t=1,\dots,5} = \boldsymbol{\xi}$ and $\mathbf{\Sigma}_{t=5,\dots,10} = \boldsymbol{\xi}'$, Student-t data $df = 2$ and change in shape. $\mathrm{P_{FA}} = 10^{-2}$.



Some performance (mismatch)

$T = 10$, $p = 10$, $N = 25$, $\boldsymbol{\Sigma}_{t=1,\dots,5} = \boldsymbol{\xi}$ and $\boldsymbol{\Sigma}_{t=5,\dots,10} = \boldsymbol{\xi}'$, Student-t data $df = 2$ but supposed 5 for statistic and change in shape. $\mathrm{P_{FA}} = 10^{-2}$.



Application to SAR change detection

Problematic

- Survey large zones of interest over a long period of time.
- Using Remote sensing methods: Satellite Time Series of SAR (Synthethic Aperture Radar) Images.
- Applications: Infrastructure monitoring, Traffic surveillance, Long-tem Change analysis, etc...

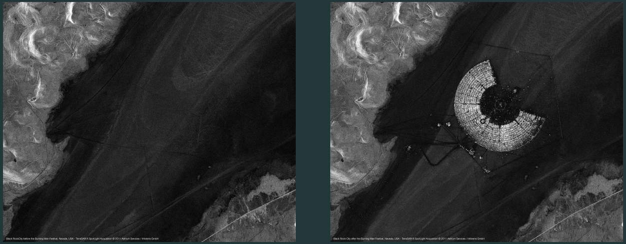
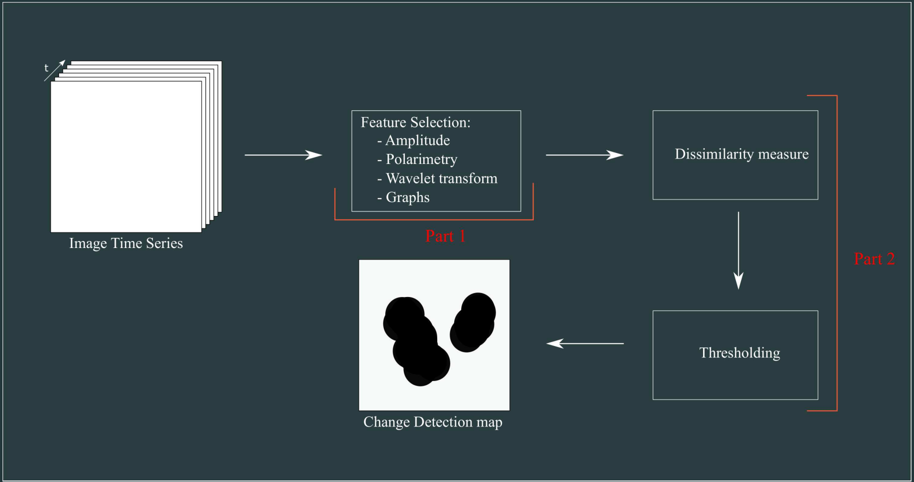


Figure 1: TerraSAR-X Images Burning Man festival

Problematic: Unsupervised Change Detection Framework

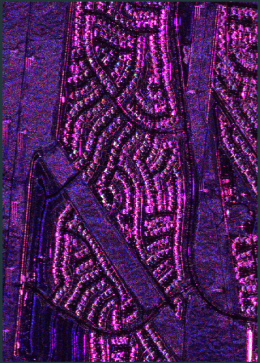
Pixel-based techniques [Hussain et al., 2013]:



Example on UAVSAR (NASA) dataset

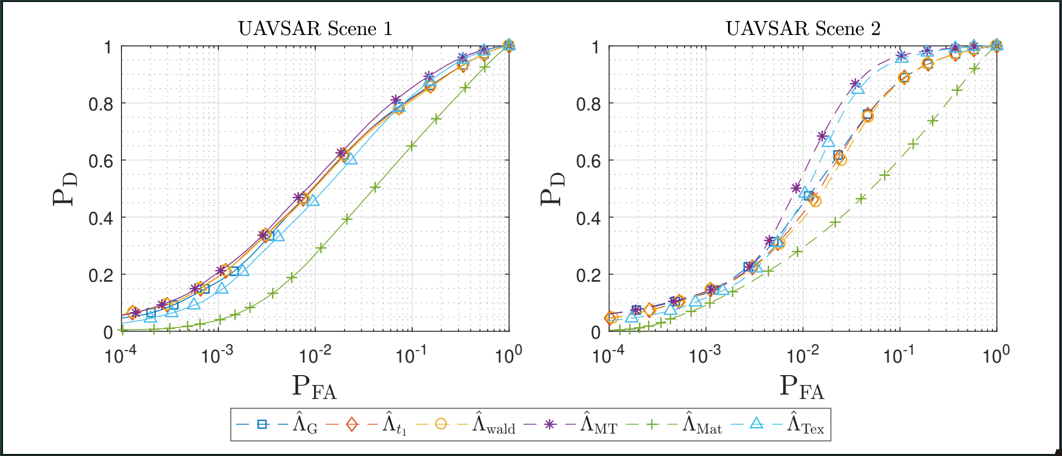
- Polarimetric data: $p = 3$
- Dimensions: 2360px 600px
- Resolution: 1.67 m (Range) and 0.60 m (Azimuth)

April 23, 2009



May 1, 2011

Example on UAVSAR (NASA) dataset ($p = 3$, $N = 121$) [Mian et al., 2018a]



References i

(1937).
Properties of sufficiency and statistical tests.
Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 160(901):268–282.

Anderson, T. (2003).
***An Introduction to Multivariate Statistical Analysis*.**
Wiley Series in Probability and Statistics. Wiley.

Box, G. E. (1949).
A general distribution theory for a class of likelihood criteria.
Biometrika, 36(3/4):317–346.

Ciuonzo, D., Carotenuto, V., and Maio, A. D. (2017).
On multiple covariance equality testing with application to SAR change detection.
IEEE Transactions on Signal Processing, 65(19):5078–5091.

Conradsen, K., Nielsen, A. A., Schou, J., and Skriver, H. (2003).
A test statistic in the complex wishart distribution and its application to change detection in polarimetric sar data.
IEEE Transactions on Geoscience and Remote Sensing, 41(1):4–19.

Hallin, M. and Paindaveine, D. (2009).
Optimal tests for homogeneity of covariance, scale, and shape.
Journal of Multivariate Analysis, 100(3):422 – 444.

Hussain, M., Chen, D., Cheng, A., Wei, H., and Stanley, D. (2013).
Change detection from remotely sensed images: From pixel-based to object-based approaches.
ISPRS Journal of Photogrammetry and Remote Sensing, 80:91 – 106.

Mian, A., Ginolhac, G., Ovarlez, J., and Atto, A. (2018a).
New robust statistics for change detection in time series of multivariate sar images.
IEEE Transactions on Signal Processing, pages 1–1.

Mian, A., Ovarlez, J.-P., Ginolhac, G., and Atto, A. M. (2018b).
Robust detection and estimation of Change-Points in a time series of multivariate images.
In *2018 26th European Signal Processing Conference (EUSIPCO) (EUSIPCO 2018)*, Roma, Italy.

Radhakrishna Rao, C. (1948).
Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation.
Mathematical Proceedings of the Cambridge Philosophical Society, 44(1):50–57.

Terrell, G. R. (2002).
The gradient statistic.
Computing Science and Statistics, 34(34):206–215.

Wald, A. (1943).
Tests of statistical hypotheses concerning several parameters when the number of observations is large.
Transactions of the American Mathematical Society, 54(3):426–482.

Yanagihara, H., Tonda, T., and Matsumoto, C. (2005).
The effects of nonnormality on asymptotic distributions of some likelihood ratio criteria for testing covariance structures under normal assumption.
Journal of Multivariate Analysis, 96(2):237–264.