A ROBUST CHANGE DETECTOR FOR HIGHLY HETEROGENEOUS MULTIVARIATE IMAGES

Ammar Mian[1,3], *Jean-Philippe Ovarlez*[1,2], *Guillaume Ginolhac*[3] *and Abdourahmane M. Atto*[3]

[1] CentraleSupélec/SONDRA, Plateau du Moulon, 3 rue Joliot-Curie, F-91190 Gif-sur-Yvette, France
[2] ONERA, DEMR/TSI, Chemin de la Hunière, F-91120 Palaiseau, France
[3] LISTIC, Université de Savoie Mont-Blanc, F-74944, Annecy le Vieux, France

ABSTRACT

In this paper, we propose new detectors for Change Detection between two multivariate images. The data is supposed to follow a Compound Gaussian distribution. By using Likelihood Ratio Test (LRT) and Generalised LRT (GLRT) approaches, we derive our detectors. The CFAR behaviour has been studied and the simulations show that they outperform the classic Gaussian Detector when the data is highly heterogeneous.

Index Terms— CFAR Change Detection, Compound Gaussian, GLRT, Remote Sensing

1. INTRODUCTION

Change Detection (CD) is a classic problem in Remote Sensing. When two images of a same scene at different times are available, the aim is to detect zones on the image corresponding to an alteration in the scene. Many techniques exist in the literature depending on the nature of the data. The basic procedure is to select a quantity of interest on which a statistic is computed. Then a decision is done given the value of this statistic. For Synthetic Aperture Radar (SAR), we can for example use the pixel value and compute correlations between the two images. This is used in interferometric applications [1]. Another way can be modeling the pixels by a density probability and test the equality of its parameters through a likelihood ratio [2].

Sometimes available images are multivariate images such as polarimetric ones. In this case, each pixel consists of a vector of size $p = 3$ (each element corresponding to the backscattering power in a given mode of polarisation). For such data, several works [3, 4, 5, 6] have proposed to model the data as Gaussian. From this modeling, the detection test was derived using an assumption of covariance matrix equality.

When considering high-resolution images, the heterogeneous behaviour of the clutter has been studied and demonstrated [7, 8]. It has been shown that the Gaussian assumption does not fit the data well enough. A new family of distributions known as Elliptical distributions has been proposed to

model the data. Specifically, the Compound Gaussian ($\mathcal{CG}$) or Spherically-Invariant Random Vectors (SIRV) [9] which is a subclass of the elliptical family is interesting for that matter. Under the SIRV assumption, composite hypothesis testing the equality of covariance matrices was studied in [10, 11] for classification purposes. In [12], a similarity measure using SIRV model has been proposed for CD. But, to our knowledge, no work has proposed to use Likelihood Ratio Test (LRT) or Generalised LRT (GLRT) approaches to develop a test of detection specifically for bi-date CD in $\mathcal{CG}$ context.

The paper here aims at using approaches similar to ones that were used in robust target detection, but in the bi-date CD framework. LRT and GLRT are derived under a binary hypothesis problem. Then their Constant False Alarm Rate (CFAR) properties are studied. Finally, the performances of the new detectors are measured.
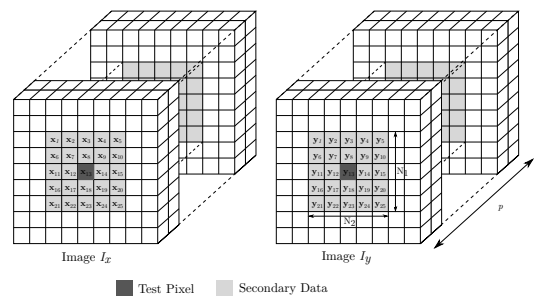
2. RELATION TO PRIOR WORK

Fig. 1. Description of the problem. A sliding window of $N = N_1 \times N_2$ observations $\mathbf{x}_{i=1\ldots N}$ and $\mathbf{y}_{i=1\ldots N}$ is used to compute the ratio for the test pixel.

The problem under study has been explored in the Gaussian assumption by testing the equality of covariance matrices. In [3], results of statistics literature on equality of Wishart matrices have been applied to CD.

Suppose we have two multivariate images $I_{\mathbf{x}}$ and $I_{\mathbf{y}}$ and we want to detect local changes in the data. We have a window of $N = N_1 \times N_2$ observations around the test pixel. We note p, the size of each vector and define $W_{\mathbf{x}} = \{\mathbf{x}_i\}_{i=1\ldots N}$,

$W_{\mathbf{y}} = \{\mathbf{y}_i\}_{i=1...N}$ the windows on images $I_{\mathbf{x}}$ and $I_{\mathbf{y}}$ respectively. Figure 1 gives an illustration of the local data selection. The pixels are modeled as multivariate complex Gaussian vectors ($\mathcal{CN}$). The pixels $\mathbf{x}_{i=1...N}$, are supposed independent and sharing the same covariance matrix. We make the same assumptions for $\mathbf{y}_{i=1...N}$. The change detection is done through as a binary hypothesis testing the equality of covariance matrices between the two images:
$\forall(\mathbf{x}_i, \mathbf{y}_i) \in W_{\mathbf{x}} \times W_{\mathbf{y}}$,

$$\begin{cases} H_0 : \mathbf{x}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma_x}) \text{ and } \mathbf{y}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma_x}) \\ H_1 : \mathbf{x}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma_x}) \text{ and } \mathbf{y}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma_y}) \end{cases} . \quad (1)$$

The GLRT detector under these hypotheses is [3]:

$$\hat{\Lambda}_G = \frac{|\hat{\mathbf{\Sigma}}_{\mathbf{x}} + \hat{\mathbf{\Sigma}}_{\mathbf{y}}|^{2N}}{|\hat{\mathbf{\Sigma}}_{\mathbf{x}}|^N |\hat{\mathbf{\Sigma}}_{\mathbf{y}}|^N} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda, \ \hat{\mathbf{\Sigma}}_{\mathbf{x}} = \sum_{i=1}^{N} \frac{\mathbf{x}_i \mathbf{x}_i^H}{N}, \ \hat{\mathbf{\Sigma}}_{\mathbf{y}} = \sum_{i=1}^{N} \frac{\mathbf{y}_i \mathbf{y}_i^H}{N} .$$

$$(2)$$

Here, $\hat{\mathbf{\Sigma}}_{\mathbf{x}}$ and $\hat{\mathbf{\Sigma}}_{\mathbf{y}}$ correspond to the Sample Covariance Matrices (SCM) of the data. Recent works on time series of SAR images such as [13, 14] still use this modeling to derive tests of detection. The contributions of the present paper are:

- We make a $\mathcal{CCG}$ assumption and derive the LRT and GLRT of the bi-date CD problem.
- We study the CFAR properties of both detectors.

3. DETECTORS UNDER $\mathcal{CCG}$ ASSUMPTION

To model the highly heterogeneous behaviour of the data, we model the probability distribution as a multivariate $\mathcal{CCG}$ with zero mean, covariance matrices $\mathbf{\Sigma_x}, \mathbf{\Sigma_y}$ and *deterministic* textures $\tau_{\mathbf{x}_i}, \tau_{\mathbf{y}_i}$:

$$p_{\mathbf{x}_i}(\mathbf{x}_i / \mathbf{\Sigma_x}, \tau_{\mathbf{x}_i}) = \frac{1}{\pi^p |\mathbf{\Sigma_x}| \tau_{\mathbf{x}_i}^p} \exp\left(-\frac{\mathbf{x}_i^H \mathbf{\Sigma_x}^{-1} \mathbf{x}_i}{\tau_{\mathbf{x}_i}}\right). \quad (3)$$

The change detection problem here can be written:
$\forall(\mathbf{x}_i, \mathbf{y}_i) \in W_{\mathbf{x}} \times W_{\mathbf{y}}$,

$$\begin{cases} H_0 : \mathbf{x}_i \sim \mathcal{CCG}(\mathbf{0}, \tau_{\mathbf{x}_i}, \mathbf{\Sigma_x}) \text{ and } \mathbf{y}_i \sim \mathcal{CCG}(\mathbf{0}, \tau_{\mathbf{y}_i}, \mathbf{\Sigma_x}) \\ H_1 : \mathbf{x}_i \sim \mathcal{CCG}(\mathbf{0}, \tau_{\mathbf{x}_i}, \mathbf{\Sigma_x}) \text{ and } \mathbf{y}_i \sim \mathcal{CCG}(\mathbf{0}, \tau_{\mathbf{y}_i}, \mathbf{\Sigma_y}) \end{cases} . \quad (4)$$

We note here that we do not make any assumptions on the textures and focus solely on the covariance matrices. Alternative hypotheses using constraints on the textures between the two images will be investigated in a forthcoming paper.

To simplify the notations we note $\boldsymbol{\tau_x} = [\tau_{\mathbf{x}_1}, \ldots, \tau_{\mathbf{x}_N}]^T$ and $\boldsymbol{\tau_y} = [\tau_{\mathbf{y}_1}, \ldots, \tau_{\mathbf{y}_N}]^T$. We also use $p_{W_{\mathbf{x}}, W_{\mathbf{y}}}(W_{\mathbf{x}}, W_{\mathbf{y}})$ when we refer to $p_{\mathbf{x}_1, ..., \mathbf{x}_N, \mathbf{y}_1, ..., \mathbf{y}_N}(\mathbf{x}_1, ..., \mathbf{x}_N, \mathbf{y}_1, ..., \mathbf{y}_N)$, the joint distribution of all observations.

3.1. Derivation of the LRT

A classic scheme in hypothesis testing is the so-called LRT. It assumes that the distributions are perfectly known. Here, it implies the knowledge of the matrices $\mathbf{\Sigma_x}, \mathbf{\Sigma_y}$ and the textures $\boldsymbol{\tau_x}, \boldsymbol{\tau_y}$. The idea is to compute the ratio of the likelihood

of the data under hypothesis H_1 and the likelihood of the data under hypothesis H_0:

$$lrt = \frac{p_{W_{\mathbf{x}}, W_{\mathbf{y}}}(W_{\mathbf{x}}, W_{\mathbf{y}}/H_1, \mathbf{\Sigma_x}, \mathbf{\Sigma_y}, \boldsymbol{\tau_x}, \boldsymbol{\tau_y})}{p_{W_{\mathbf{x}}, W_{\mathbf{y}}}(W_{\mathbf{x}}, W_{\mathbf{y}}/H_0, \mathbf{\Sigma_x}, \boldsymbol{\tau_x}, \boldsymbol{\tau_y})}. \quad (5)$$

The resulting value has then to be compared to a threshold value. The Neyman Pearson lemma [15] ensures that the LRT is the uniformly most powerful test (in the sense of maximising the probability of detection) at a given significance level (*a.k.a.* Probability of False Alarm).

Proposition 3.1 (Likelihood Ratio). *The Likelihood Ratio under the binary hypothesis presented in* (4) *is the following:*

$$\Lambda_{lrt} = \frac{|\mathbf{\Sigma_x}|^N}{|\mathbf{\Sigma_y}|^N} \exp\left(\text{Tr}\left((\mathbf{\Sigma_x}^{-1} - \mathbf{\Sigma_y}^{-1}) \sum_{i=1}^{N} \frac{\mathbf{y}_i \mathbf{y}_i^H}{\tau_{\mathbf{y}_i}} \right) \right) \underset{H_0}{\overset{H_1}{\gtrless}} \lambda.$$

$$(6)$$

Proof. Straightforward calculus. $\square$

When the parameters are unknown, Maximum Likelihood Estimates (MLE) of the unknown parameters are plugged into the LRT.

Proposition 3.2 (2-step lrt). *The detector obtained when replacing unknown matrices and textures in ratio* (6) *is:*

$$\hat{\Lambda}_{lrt} = \frac{\left|\hat{\mathbf{\Sigma}}_{\mathbf{x}}\right|^N}{\left|\hat{\mathbf{\Sigma}}_{\mathbf{y}}\right|^N} \exp\left(p \sum_{i=1}^{N} \frac{\mathbf{y}_i^H \hat{\mathbf{\Sigma}}_{\mathbf{x}}^{-1} \mathbf{y}_i}{\mathbf{y}_i^H \hat{\mathbf{\Sigma}}_{\mathbf{y}}^{-1} \mathbf{y}_i} \right) \underset{H_0}{\overset{H_1}{\gtrless}} \lambda,$$

where

$$(7)$$

$$\hat{\mathbf{\Sigma}}_{\mathbf{x}} = \frac{p}{N} \sum_{i=1}^{N} \frac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H \hat{\mathbf{\Sigma}}_{\mathbf{x}}^{-1} \mathbf{x}_i}, \ \hat{\mathbf{\Sigma}}_{\mathbf{y}} = \frac{p}{N} \sum_{i=1}^{N} \frac{\mathbf{y}_i \mathbf{y}_i^H}{\mathbf{y}_i^H \hat{\mathbf{\Sigma}}_{\mathbf{y}}^{-1} \mathbf{y}_i}.$$

Proof. Straigthforward. See [16] for the estimators of covariance matrices under $\mathcal{CCG}$ assumption. $\square$

Discussion: It is interesting to note that when the parameters are known, the lrt uses solely observations on one image I_y. The ratio test whether the data $W_{\mathbf{y}}$ is better described by the covariance matrix $\mathbf{\Sigma_y}$ than $\mathbf{\Sigma_x}$. When the covariance matrices are unknown, classic Tyler's fixed point estimators are used [17]. We remind that an infinite number of matrices are solution to the fixed point equation (if $\hat{\mathbf{M}}$ is solution, then so is $a \hat{\mathbf{M}}, \forall a \in \mathbb{R}^*$). Classically to resolve this, a normalization constraint is imposed on the matrix ($\text{Tr}(\hat{\mathbf{M}}) = p$). Contrarily to the Gaussian case, the estimation is more robust to arbitrary change in the power of the scatterers. We focus more on the correlation structure between the different channels of the images. This also means that a punctual change (on a pixel), even corresponding to the arrival of a powerful target, is less likely to be detected. Indeed, contrarily to the SCM, such a change is considered as an outlier and would not impact much the estimated covariance matrix. We expect this detector to provide enhancements in situations where the change is distributed with a low SNR.

3.2. Derivation of the GLRT

When the parameters of the distribution are unknown, another way around is to use the GLRT. This method differs from the previous one in that the estimation of the parameters is not done using the likelihood of the observations but rather the likelihood under both hypotheses:

$$glrt = \frac{\max\limits_{\boldsymbol{\Sigma_x},\boldsymbol{\Sigma_y},\boldsymbol{\tau_x},\boldsymbol{\tau_y}} p_{W_\mathbf{x},W_\mathbf{y}}(W_\mathbf{x},W_\mathbf{y}/H_1,\boldsymbol{\Sigma_x},\boldsymbol{\Sigma_y},\boldsymbol{\tau_x},\boldsymbol{\tau_y})}{\max\limits_{\boldsymbol{\Sigma_x},\boldsymbol{\tau_x}} p_{W_\mathbf{x},W_\mathbf{y}}(W_\mathbf{x},W_\mathbf{y}/H_0,\boldsymbol{\Sigma_x},\boldsymbol{\tau_x})}. \tag{8}$$

Proposition 3.3 (Generalised Likelihood Ratio). *The Generalised Likelihood Ratio under binary hypothesis presented in (4) is the following:*

$$\hat{\Lambda}_{glrt} = \frac{\left|\hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_0}}\right|^{2N}}{\left|\hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_1}}\right|^N \left|\hat{\boldsymbol{\Sigma}}_\mathbf{y}\right|^N} \prod_{i=1}^N \frac{\left(\mathbf{x}_i^H \hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_0}}^{-1} \mathbf{x}_i\right)^p \left(\mathbf{y}_i^H \hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_0}}^{-1} \mathbf{y}_i\right)^p}{\left(\mathbf{x}_i^H \hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_1}}^{-1} \mathbf{x}_i\right)^p \left(\mathbf{y}_i^H \hat{\boldsymbol{\Sigma}}_\mathbf{y}^{-1} \mathbf{y}_i\right)^p} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda$$

where $\hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_0}} = \dfrac{p}{2N} \sum_{i=1}^N \left(\dfrac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H \hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_0}}^{-1} \mathbf{x}_i} + \dfrac{\mathbf{y}_i \mathbf{y}_i^H}{\mathbf{y}_i^H \hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_0}}^{-1} \mathbf{y}_i}\right)$

and $\hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_1}} = \dfrac{p}{N} \sum_{i=1}^N \dfrac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H \hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_1}}^{-1} \mathbf{x}_i}, \; \hat{\boldsymbol{\Sigma}}_\mathbf{y} = \dfrac{p}{N} \sum_{i=1}^N \dfrac{\mathbf{y}_i \mathbf{y}_i^H}{\mathbf{y}_i^H \hat{\boldsymbol{\Sigma}}_\mathbf{y}^{-1} \mathbf{y}_i}.$ (9)

Proof. See Appendix A. □

Discussion: The GLRT here is slightly more complex than the LRT. The estimation of $\hat{\boldsymbol{\Sigma}}_{\mathbf{x},H_0}$ differs from a classic Tyler's estimator. One can wonder about its convergence under H_1 regime, but this is out of the scope of the paper. The ratio can be seen as a product of two separate functions: the ratio of determinants which is similar to the GLRT under Gaussian assumption and a function that computes a ratio between the textures under H_1 and H_0 regime. Similarly to the LRT, the matrices are normalised by the trace.

3.3. Statistical properties

Proposition 3.4 (CFARness of the LRT and GLRT). $\hat{\Lambda}_{lrt}$ *is CFAR texture but not matrix.* $\hat{\Lambda}_{glrt}$ *is CFAR texture and matrix.*

Proof. See Appendix B. □

An interesting problem is the calculus of the detector distributions under H_0. Unfortunately, as the analytical expression is composed of statistically dependent quadratic forms, the calculus is intricate. The problem is let open as of now.

4. SIMULATION RESULTS

Table 1. Simulation-relevant parameters

α,β	ρ_x,ρ_y	p	N	SNR
shape and scale for Γ-distribution	coefficients for Toeplitz matrices	Size of vector	Number of observations	Signal to Noise Ratio

The CFARness behaviour of the two detectors has been studied in simulation. Table 1 summarizes the relevant parameters of the simulation. We generated observations through a K-distribution (the texture is distributed according to a Γ-distribution) with same covariance matrix. The covariance matrix $\boldsymbol{\Sigma_x} = \boldsymbol{\Sigma_y}$ (under H_0) was chosen to be a Toeplitz matrix of the form $(\boldsymbol{\Sigma_x})(m,n) = \rho_x^{|m-n|}$. Figure 2, corresponding to $P_{FA} - \lambda$ curves for several matrix and texture parameters validates CFAR behaviour of the GLRT. We observe that indeed, in practice $\hat{\Lambda}_{lrt}$ is not CFAR matrix due to the trace constraint. We can see that $\hat{\Lambda}_G$ detector does not respect the CFAR texture property which is expected as the textures are not taken into account in the design of the detector. It is thus not suited in our situation.
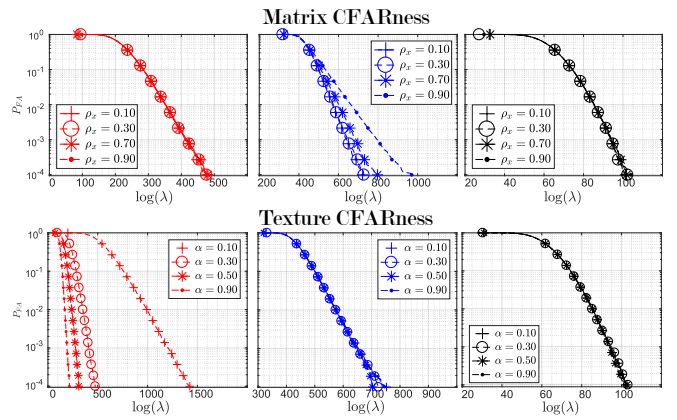

Fig. 2. $P_{FA} - \lambda$. Left = $\hat{\Lambda}_G$, Middle = $\hat{\Lambda}_{lrt}$, Right = $\hat{\Lambda}_{glrt}$. $p = 10$, $N = 25$. Top: $\alpha = 0.3$, $\beta = 0.1$. Bottom: $\rho_x = 0.3$, $\beta = 0.1$.
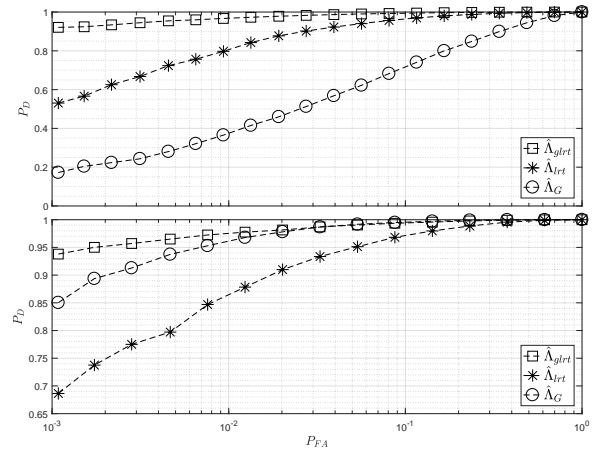

Fig. 3. $P_D - P_{FA}$ with $p = 10$, $N = 25$, $\rho_x = 0.1$, $\rho_y = 0.9$, $\alpha = 0.3$, $\beta = 0.1$. Top: SNR = 0 dB. Bottom: SNR= 20 dB.

The performances of the detectors have also been measured. We used two synthetic images where each pixel is generated through a K-distribution of same texture parameters and covariance matrices. On one image, a patch of 10×10 pixels, corresponding to the change, has been generated with a different covariance matrix. The SNR is applied through the scale parameter: as the product $\alpha\beta$ corresponds to the power

of the data, we increase the scale parameter on the patch to fit the given SNR. Figure 3 shows an example of a performance obtained through 1000 Monte-Carlo trials. As expected, the new detectors outperform the Gaussian one. The GLRT approach seems to fare better overall. We also observe that the performance of the new detectors do not vary much with the SNR comparatively to $\hat{\Lambda}_G$. This result was expected as the textures are systematically compensated in the ratio to test the structure of the covariances matrices alone.

5. CONCLUSION

In this paper, new detectors for Change Detection on highly heterogeneous multivariate images have been proposed. The data have been modeled as Compound Gaussian and by using Likelihood Ratio approaches, the detectors have been derived. The CFAR behaviour of the detectors have been studied both theoretically and in simulation. A ROC curve has also been computed to test the performances of the detectors and it shows that the new detectors, specifically $\hat{\Lambda}_{glrt}$, outperform the Gaussian one even at high SNR by a significant factor.

A. DERIVATION OF THE GLRT

The expression of the ratio in eq (8) leads us to an estimation of the parameters which maximize the likelihood under both H_0 and H_1 hypothesis. We will treat both separately.

Under H_0: Since the $\log$ function is monotone, we can focus on the following optimization problem:

$$\underset{\boldsymbol{\Sigma}_{\mathbf{x}},\boldsymbol{\tau}_{\boldsymbol{x}},\boldsymbol{\tau}_{\boldsymbol{y}}}{\text{argmax}} \log p_{W_{\mathbf{x}},W_{\mathbf{y}}}(W_{\mathbf{x}},W_{\mathbf{y}}/H_1,\boldsymbol{\Sigma}_{\mathbf{x}},\boldsymbol{\tau}_{\boldsymbol{x}},\boldsymbol{\tau}_{\boldsymbol{y}})$$

$$= -2N \log|\boldsymbol{\Sigma}_{\mathbf{x}}| - \sum_{i=1}^{N}\left(\frac{\mathbf{x}_i^H\boldsymbol{\Sigma}_{\mathbf{x}}^{-1}\mathbf{x}_i}{\tau_{\mathbf{x}_i}} - \frac{\mathbf{y}_i^H\boldsymbol{\Sigma}_{\mathbf{x}}^{-1}\mathbf{y}_i}{\tau_{\mathbf{y}_i}}\right).$$

• Estimation of $\boldsymbol{\tau}_{\boldsymbol{x}}$, $\boldsymbol{\tau}_{\boldsymbol{y}}$: deriving $\log p_{W_{\mathbf{x}},W_{\mathbf{y}}}$ with respect to $\tau_{\boldsymbol{x}}$ gives, for each element of $\boldsymbol{\tau}_{\boldsymbol{x}}$: $\forall i, \hat{\tau}_{\mathbf{x}_i} = \mathbf{x}_i^H\boldsymbol{\Sigma}_{\mathbf{x}}^{-1}\mathbf{x}_i/p$. Similarly, $\forall i, \hat{\tau}_{\mathbf{y}_i} = \mathbf{y}_i^H\boldsymbol{\Sigma}_{\mathbf{x}}^{-1}\mathbf{y}_i/p$.
• Estimation of $\boldsymbol{\Sigma}_{\mathbf{x}}$: deriving $\log p_{W_{\mathbf{x}},W_{\mathbf{y}}}$ with respect to $\boldsymbol{\Sigma}_{\mathbf{x}}$ lead to $\hat{\boldsymbol{\Sigma}}_{\mathbf{x}} = \frac{1}{2N}\sum_{i=1}^{N}\left(\frac{\mathbf{x}_i\mathbf{x}_i^H}{\tau_{\mathbf{x}_i}} + \frac{\mathbf{y}_i\mathbf{y}_i^H}{\tau_{\mathbf{y}_i}}\right)$ and replacing the texture by their estimates gives the expression of eq (9). For clarity, this estimator will be noted as $\hat{\boldsymbol{\Sigma}}_{\mathbf{x},H_0}$.

Under H_1: We have the following optimization problem:
$$\underset{\boldsymbol{\Sigma}_{\mathbf{x}},\boldsymbol{\Sigma}_{\mathbf{y}},\boldsymbol{\tau}_{\boldsymbol{x}},\boldsymbol{\tau}_{\boldsymbol{y}}}{\text{argmax}} \log p_{W_{\mathbf{x}},W_{\mathbf{y}}}(W_{\mathbf{x}},W_{\mathbf{y}}/H_1,\boldsymbol{\Sigma}_{\mathbf{x}},\boldsymbol{\Sigma}_{\mathbf{y}},\boldsymbol{\tau}_{\boldsymbol{x}},\boldsymbol{\tau}_{\boldsymbol{y}}).$$
• Estimation of $\boldsymbol{\tau}_{\boldsymbol{x}}$, $\boldsymbol{\tau}_{\boldsymbol{y}}$: By the same procedure as H_0, we have $\forall i, \hat{\tau}_{\mathbf{x}_i} = \mathbf{x}_i^H\boldsymbol{\Sigma}_{\mathbf{x}}^{-1}\mathbf{x}_i/p$, $\hat{\tau}_{\mathbf{y}_i} = \mathbf{y_i}^H\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{y}_i/p$.
• Estimation of $\boldsymbol{\Sigma}_{\mathbf{x}}$, $\boldsymbol{\Sigma}_{\mathbf{y}}$: A simple optimization on the covariance matrices using the differential identities given previously results in Tyler's fixed point estimators for both. For clarity, we will note $\hat{\boldsymbol{\Sigma}}_{\mathbf{x}}$ as $\hat{\boldsymbol{\Sigma}}_{\mathbf{x},H_1}$.

Ratio: Finally, we replace the unknown parameters by their estimate in the likelihood ratio which allows to maximise the likelihoods. The calculus leads to (9).

B. CFARNESS OF THE 2-STEP LRT AND GLRT

CFAR texture: It suffices to write $\forall i$, $\mathbf{x}_i = \sqrt{\tau_{\mathbf{x}_i}}\ddot{\mathbf{x}}_i$ and $\mathbf{y}_i = \sqrt{\tau_{\mathbf{y}_i}}\ddot{\mathbf{y}}_i$ with $\ddot{\mathbf{x}}_i \sim \mathcal{CN}(\mathbf{0},\boldsymbol{\Sigma}_{\mathbf{x}})$, $\ddot{\mathbf{y}}_i \sim \mathcal{CN}(\mathbf{0},\boldsymbol{\Sigma}_{\mathbf{y}})$. The LRT is rewritten as:

$$\hat{\Lambda}_{lrt} = \frac{|\hat{\boldsymbol{\Sigma}}_{\mathbf{x}}|^N}{|\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}|^N}\exp\left(p\sum_{i=1}^{N}\frac{\ddot{\mathbf{y}}_i^H\hat{\boldsymbol{\Sigma}}_{\mathbf{x}}^{-1}\ddot{\mathbf{y}}_i}{\ddot{\mathbf{y}}_i^H\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}^{-1}\ddot{\mathbf{y}}_i}\right) \text{ with}$$

$$\hat{\boldsymbol{\Sigma}}_{\mathbf{x}} = \frac{p}{N}\sum_{i=1}^{N}\frac{\ddot{\mathbf{x}}_i\ddot{\mathbf{x}}_i^H}{\ddot{\mathbf{x}}_i^H\hat{\boldsymbol{\Sigma}}_{\mathbf{x}}^{-1}\ddot{\mathbf{x}}_i}, \hat{\boldsymbol{\Sigma}}_{\mathbf{y}} = \frac{p}{N}\sum_{i=1}^{N}\frac{\ddot{\mathbf{y}}_i\ddot{\mathbf{y}}_i^H}{\ddot{\mathbf{y}}_i^H\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}^{-1}\ddot{\mathbf{y}}_i}.$$

Here, the textures simplify, which means that the ratio does not change for any given set of textures. Thus, the LRT is CFAR texture. The proof is similar for the GLRT.

CFAR Matrix: First of all we verify that the trace normalization constraint does not affect the ratio. One way to see the normalization is by replacing the estimated matrix $\hat{\boldsymbol{\Sigma}}_\bullet$ by $p\frac{\hat{\boldsymbol{\Sigma}}_\bullet}{\text{Tr}(\hat{\boldsymbol{\Sigma}}_\bullet)}$ in the detectors' expression. For the 2-step LRT, we write:

$$\hat{\Lambda}_{lrt} = \frac{|\hat{\boldsymbol{\Sigma}}_{\mathbf{x}}|^N\text{Tr}(\hat{\boldsymbol{\Sigma}}_{\mathbf{y}})^{pN}}{|\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}|^N\text{Tr}(\hat{\boldsymbol{\Sigma}}_{\mathbf{x}})^{pN}}\exp\left(p\frac{\text{Tr}(\hat{\boldsymbol{\Sigma}}_{\mathbf{y}})}{\text{Tr}(\hat{\boldsymbol{\Sigma}}_{\mathbf{x}})}\sum_{i=1}^{N}\frac{\mathbf{y}_i^H\hat{\boldsymbol{\Sigma}}_{\mathbf{x}}^{-1}\mathbf{y}_i}{\mathbf{y}_i^H\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}^{-1}\mathbf{y}_i}\right).$$

The traces are dependent on the covariances matrices of the data, which means that even if we whiten the data, the ratio will be dependent on the covariance matrix of the data. Thus, in practice the 2-step LRT is not CFAR matrix.

The GLRT, on the other hand, is written as:

$$\hat{\Lambda}_{glrt} = \frac{|\hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_0}}|^{2N}\text{Tr}(\hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_1}})^{pN}\text{Tr}(\hat{\boldsymbol{\Sigma}}_{\mathbf{y}})^{pN}}{|\hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_1}}|^N|\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}|^N\text{Tr}(\hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_0}})^{2pN}} \times$$

$$\prod_{i=1}^{N}\frac{\text{Tr}(\hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_0}})^{2p}}{\text{Tr}(\hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_1}})^p\text{Tr}(\hat{\boldsymbol{\Sigma}}_{\mathbf{y}})^p}\frac{(\mathbf{x}_i^H\hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_0}}^{-1}\mathbf{x}_i)^p(\mathbf{y}_i^H\hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_0}}^{-1}\mathbf{y}_i)^p}{(\mathbf{x}_i^H\hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_1}}^{-1}\mathbf{x}_i)^p(\mathbf{y}_i^H\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}^{-1}\mathbf{y}_i)^p}.$$

It is easy to see that the traces terms simplify in the expression. We note $\tilde{\boldsymbol{\Sigma}}_\bullet = \boldsymbol{\Sigma}_{\mathbf{x}}^{-\frac{1}{2}}\hat{\boldsymbol{\Sigma}}_\bullet\boldsymbol{\Sigma}_{\mathbf{x}}^{-\frac{1}{2}}$ and $\tilde{\mathbf{x}}_i = \boldsymbol{\Sigma}_{\mathbf{x}}^{-\frac{1}{2}}\ddot{\mathbf{x}}_i$, $\tilde{\mathbf{y}}_i = \boldsymbol{\Sigma}_{\mathbf{x}}^{-\frac{1}{2}}\ddot{\mathbf{y}}_i$. And we can rewrite the GLRT as:

$$\hat{\Lambda}_{glrt} = \frac{|\tilde{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_0}}|^{2N}}{|\tilde{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_1}}|^N|\tilde{\boldsymbol{\Sigma}}_{\mathbf{y}}|^N}\prod_{i=1}^{N}\frac{(\tilde{\mathbf{x}}_i^H\tilde{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_0}}^{-1}\tilde{\mathbf{x}}_i)^p(\tilde{\mathbf{y}}_i^H\tilde{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_0}}^{-1}\tilde{\mathbf{y}}_i)^p}{(\tilde{\mathbf{x}}_i^H\tilde{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_1}}^{-1}\tilde{\mathbf{x}}_i)^p(\tilde{\mathbf{y}}_i^H\tilde{\boldsymbol{\Sigma}}_{\mathbf{y}}^{-1}\tilde{\mathbf{y}}_i)^p}.$$

We have $\tilde{\mathbf{x}}_i \sim \mathcal{CN}(\mathbf{0},\mathbf{I})$ and $\tilde{\mathbf{y}}_i \sim \mathcal{CN}(\mathbf{0},\mathbf{I})$, which are not dependent on $\boldsymbol{\Sigma}_{\mathbf{x}}$. Since $\hat{\boldsymbol{\Sigma}}_{\mathbf{x}}$ and $\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}$ are classic Tyler fixed point estimators, $\tilde{\boldsymbol{\Sigma}}_{\mathbf{x}}$ and $\tilde{\boldsymbol{\Sigma}}_{\mathbf{y}}$ are estimators of identity matrix $\mathbf{I}$ as per the results of [18]. For $\tilde{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_0}}$, we name $\mathbf{x}_{N+1},...,\mathbf{x}_{2N} = \mathbf{y}_1,...,\mathbf{y}_N$ so that $\hat{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_0}}$ is also a Tyler's estimator with $2N$ samples, and thus $\tilde{\boldsymbol{\Sigma}}_{\mathbf{x},\mathbf{H_0}}$ is also an estimator of $\mathbf{I}$. In the end, no member of the ratio is dependent on $\boldsymbol{\Sigma}_{\mathbf{x}}$ under H_0 which ends the proof.

C. REFERENCES

[1] Paul Thompson, Daniel E. Wahl, Paul H. Eichel, Dennis C. Ghiglia, and Charles V. Jakowatz, *Spotlight-Mode Synthetic Aperture Radar: A Signal Processing Approach*, Kluwer Academic Publishers, Norwell, MA, USA, 1996.

[2] Y.Z Hsu, H.-H Nagel, and G Rekers, "New likelihood test methods for change detection in image sequences," *Computer Vision, Graphics, and Image Processing*, vol. 26, no. 1, pp. 73 – 106, 1984.

[3] K. Conradsen, A. Aasbjerg Nielsen, J. Schou, and H. Skriver, "Change detection in polarimetric SAR data and the complex wishart distribution," in *IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No.01CH37217)*, 2001, vol. 6, pp. 2628–2630 vol.6.

[4] L. M. Novak, "Coherent change detection for multipolarization SAR," in *Conference Record of the Thirty-Ninth Asilomar Conference onSignals, Systems and Computers, 2005.*, Oct 2005, pp. 568–573.

[5] L. Pallotta, C. Clemente, A. De Maio, and D. Orlando, "A multi-family GLRT for detection in polarimetric SAR images," in *2016 Sensor Signal Processing for Defence (SSPD)*, Sept 2016, pp. 1–5.

[6] V. Carotenuto, A. De Maio, C. Clemente, and J. Soraghan, "Unstructured versus structured glrt for multipolarization sar change detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 8, pp. 1665–1669, Aug 2015.

[7] M. S. Greco and F. Gini, "Statistical analysis of high-resolution SAR ground clutter data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 3, pp. 566–575, March 2007.

[8] E. Ollila, D. E. Tyler, V. Koivunen, and H. V. Poor, "Compound-gaussian clutter modeling with an inverse gaussian texture distribution," *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 876–879, Dec 2012.

[9] K. Yao, "A Representation Theorem and its Applications to Spherically Invariant Random Processes," *Information Theory, IEEE Transactions on*, vol. 19, pp. 600–608, September 1973.

[10] G. Vasile, J. P. Ovarlez, F. Pascal, and C. Tison, "Coherency matrix estimation of heterogeneous clutter in high-resolution polarimetric SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 4, pp. 1809–1826, April 2010.

[11] Lionel Bombrun, Gabriel Vasile, and Michel Gay, "Segmentation of Polarimetric SAR images based on the SIRV model and the Kummeru PDF," in *XXIIème colloque GRETSI (GRETSI 2009)*, Dijon, France, Sept. 2009, p. 4.

[12] M. Liu, H. Zhang, C. Wang, and F. Wu, "Change detection of multilook polarimetric SAR images using heterogeneous clutter models," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 12, pp. 7483–7494, Dec 2014.

[13] A. A. Nielsen, K. Conradsen, and H. Skriver, "Omnibus test for change detection in a time sequence of polarimetric SAR data," in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, July 2016, pp. 3398–3401.

[14] D. Ciuonzo, V. Carotenuto, and A. De Maio, "On multiple covariance equality testing with application to SAR change detection," *IEEE Transactions on Signal Processing*, vol. 65, no. 19, pp. 5078–5091, Oct 2017.

[15] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 231, pp. 289–337, 1933.

[16] Fulvio Gini and Maria Greco, "Covariance matrix estimation for cfar detection in correlated heavy tailed clutter," *Signal Process.*, vol. 82, no. 12, pp. 1847–1859, Dec. 2002.

[17] David E. Tyler, "A distribution-free m-estimator of multivariate scatter," *Ann. Statist.*, vol. 15, no. 1, pp. 234–251, 03 1987.

[18] F. Pascal, J.-P. Ovarlez, P. Forster, and P. Larzabal, "On a SIRV-CFAR Detector with Radar Experimentations in Impulsive Noise," *Proc. EUSIPCO*, September 2006.