

New Robust Statistics for Change Detection in Time Series of Multivariate SAR Images

Ammar MIAN^{1,2}, Guillaume GINOLHAC^{2,3}, Jean-Philippe OVARLEZ^{1,4}, Abdourrahmane M. ATTO^{2,3}

Abstract—This paper explores the problem of change detection in time series of heterogeneous multivariate synthetic aperture radar images. Classical change detection schemes have modelled the data as a realisation of Gaussian random vectors and have derived statistical tests under this assumption. However, when considering high-resolution images, the heterogeneous behaviour of the scatterers is not well described by a Gaussian model. In this paper, the data model is extended to Spherically Invariant Random Vectors where the heterogeneity of the images is accounted for through a deterministic texture parameter. Then three separate detection problems are considered and generalised likelihood ratio test technique is used to derive statistical tests for each problem. The constant false alarm rate property of the new statistics are studied both theoretically and through simulation. Finally, the performance of the new statistics are studied both in simulation and on real synthetic aperture radar data and compared to Gaussian-derived ones. The study yields promising results when the data are heterogeneous.

Index Terms—Image Time Series ; Change Detection ; Synthetic Aperture Radar; Robust detection ; Generalised Likelihood Ratio Test ; Spherically Invariant Random Vectors

I. INTRODUCTION

A. Motivations and relation to prior works

Recent years have seen an increase in the number of remotely sensed images of the earth. Synthetic Aperture Radar (SAR) images are more widely available thanks to space missions such as Sentinel-1. These radar systems are known for their all-weather sensing capabilities, which makes them a good source of information when studying the evolution of a large area over time. Change Detection (CD) in these image time series (ITS) is needed for a large variety of applications such as land-cover monitoring, disaster management or study of global warming.

CD in SAR images has been a popular subject of study in the recent years [1]. Since SAR systems are naturally subject to speckle noise, the statistical framework has been successful in the analysis of the acquired images. The Gaussian assumption has been widely used to model the pixels of the

¹ SONDRA, Centrale Supélec, Université Paris-Saclay, F-91190, Gif-sur-Yvette, France.

² Université Savoie Mont Blanc, LISTIC, France

³ Université Grenoble Alpes, France

⁴ DEMR, ONERA, Université Paris-Saclay, F-91123 Palaiseau, France.

The work was partially supported by PHOENIX ANR-15-CE23-0012 grant of the French *National Agency of Research*.

This paper has supplementary downloadable material available at <http://ieeexplore.ieee.org>, provided by the author. The material includes a detailed derivation of the statistics presented in the paper. This material is 308Ko in size. A demonstration (code) of the statistics is also available at <https://github.com/AmmarMian/DemonstrationANRPhoenix>.

images and has provided solid results for applications such as target detection [2]. When it comes to CD, several approaches have been explored. The Coherent Change Detection (CCD) is a well-known approach that considers local correlations of pixels between two dates. The pixels at both dates are concatenated into a unique vector which is then modelled by a given distribution. Under this formulation, the change has been parametrised through a scale factor between the covariance matrices of both dates in [3]. In [4], [5], binary hypothesis testing has been introduced and both Likelihood Ratio Test (LRT) and Generalised Likelihood Ratio Test (GLRT) have been derived. However, since these approaches use local correlations, they are sensitive to variation of phase between the two dates. If the conditions are not the same, many false alarms may arise due to the phase difference.

Other approaches have considered using statistical information theory to design a distance between the images [6], [7]. The pixels are modelled by a given distribution (typically Gaussian) and classic dissimilarity measures such as Kullback-Leibler (KL) divergence are used to obtain a comparative statistic. In this case, the methodology is less sensitive to the conditions since the local spatial distribution of the data is used to compute the change map. However, deciding on a threshold of detection is a rather difficult problem when using such distances.

Finally, covariance equality test has been introduced in [8], [9] for the case of two or three dates. Several works [10], [11] have considered variations of these tests aimed at specific applications. Recently, an extension to the general case of $T > 2$ images has been considered in [12] and a statistic has been derived using Generalised Likelihood Ratio Test (GLRT) methodology. In [13], Rao and Wald methodologies have been explored as well. Testing covariance equality is a classic problem within the statistical literature namely, various statistics have been suggested in [14], [15], [16]. When considering time series of multivariate vectors, the detection of change-point in a series of covariance matrices has been developed in [17], [18] for financial data analysis.

The works mentioned up to now use a Gaussian assumption which has been successful for standard SAR images. However, with the increase in resolution obtained in High Resolution (HR) SAR, the Normal distribution does not fit the observations well enough [19], [20], [21]. For those kind of images, an heterogeneity in the power of the pixels is observed locally. This behaviour has been modelled by introducing Spherically Invariant Random Vectors (SIRV) [22] which are a sub-family of the elliptical distributions [23]. This new model

has presented good results for radar applications [24], [25], [26].

Covariance equality test under non-Gaussian assumption has been explored in [27] where a Gaussian-derived statistic has been modified to work under a large variety of elliptical distributions. Although the proposed statistics have interesting asymptotic optimality properties, they use the Sample Covariance Matrix (SCM) estimator which lacks robustness in the presence of outliers. For SAR images specifically, [28] proposed to plug a robust Tyler estimate of covariance matrices into the Gaussian-derived statistic. Under SIRV model, [29] proposed a LRT statistic where the same robust estimates of covariances are used. However, these 2-step methodologies do not take into account a trace normalisation constraint which leads to a non Constant False Alarm (CFAR) property of the statistics. For the case $T = 2$, the present authors have proposed in [30] a GLRT statistic which conserves the CFAR property. The study shows interesting results of robustness under SIRV distributions.

In this paper, we consider an extension of [30]. The contributions of the present paper are summed-up as:

- We consider the problem of CD in HR SAR images for the general case $T > 2$. The heterogeneity on the spatial neighbourhood is taken into account by using a SIRV assumption. The texture is assumed to be deterministic and unknown, and is taken into account in three different ways leading to separate detection problems.
- We derive statistic of decision for each problem using the GLRT methodology. We also consider marginal statistics which are used for the change-point estimation strategy presented in [31].
- The derivation of the statistics leads to novel fixed-point estimates for the covariance matrices. The convergence properties of these estimates are considered. Then the CFAR property of the new statistics are studied.
- The new statistics are applied on two separate real datasets and have better performance than Gaussian-derived ones.

B. Paper Organisation

The paper is organised as follows: section II gives preliminary definitions and provides background on CD under Gaussian model. In section III an extension of CD under robust model is presented. Then GLRT for the different problems are derived in section IV. Section V considers the convergence of the novel covariance estimates. Then in section VI, statistical properties of the new statistics are explored. Simulations are done in section VII on both synthetic and real dataset. Finally, conclusions are presented in VIII. Proofs are given in Appendices.

In the scope of this paper, the following notations will be used: lower-case (resp. Upper-case) bold letters denotes vectors (resp. matrices). \mathbb{N}^p , \mathbb{R}^p and \mathbb{C}^p are the sets of integer, real and complex p -dimensional vectors. $\mathbb{S}_{\mathbb{H}}^p$ is the set of Hermitian semi-definite matrices of size $p \times p$. Given $(a, b) \in \mathbb{N}^2$, $b \geq a$, $\llbracket a, b \rrbracket$ denotes the set $\{a, \dots, b\}$. δ_{ik} is the

Kronecker symbol. Θ is an arbitrary parameter space. $\mathbf{0}_p$ is the p -dimensional null vector. \mathbf{I}_p is the identity matrix of size $p \times p$. For any given matrix, \bullet^T , \bullet^H represent respectively the transpose and transpose conjugate operators. $\Re(\bullet)$ and $\Im(\bullet)$ denote the real and imaginary parts. Notations $\text{Tr}(\bullet)$, $|\bullet|$ and $\|\bullet\|$ are the trace, determinant and euclidean norm operators. $\text{vec}(\bullet)$ is the vectorisation operator. Notation \bullet^{-1} is the inverse operation. The symbol \otimes denotes the Kronecker product. Given a scalar valued function f , $\frac{\partial f}{\partial \bullet}$ denotes the gradient of f w.r.t \bullet arranged in a column. $\mathbb{1}_K$ is the indicator function of set K . \mathbf{x} will always represent a random vector of size p . Any subscript or superscript serves to indicate a specific observation. Σ will always be an Hermitian matrix of size $p \times p$. The symbol \sim means "distributed as". H_0 and H_1 denote both possible hypothesis in a binary hypotheses test scheme.

II. BACKGROUND ON CHANGE DETECTION UNDER GAUSSIAN MODEL

In this section we give useful definitions that will be used in the paper. Then we give some background on CD under Gaussian model.

A. Preliminary definitions

In general, to detect changes, a small subset of the image is considered in the form of a sliding window. This window serves as a mask in order to select the observations corresponding to a local spatial neighbourhood. We define N_1 , N_2 the size of this window and $N = N_1 \times N_2$. We denote the observations on the window as $\mathbf{x}_k^{(t)}$. The subscript $k \in \llbracket 1, N \rrbracket$ serves to identify the pixel and $t \in \llbracket 1, T \rrbracket$, the date of observation. Let $(t_1, t_2) \in \llbracket 1, T \rrbracket^2$, we define $\mathcal{W}_{t_1, t_2} = \{\mathbf{x}_k^{(t)} | k \in \llbracket 1, N \rrbracket, k \in \llbracket t_1, t_2 \rrbracket\}$. Figure 1 gives an illustration of the local data selection.

To simplify the equations, we define the following quantities:

$$\begin{aligned} q(\Sigma, \mathbf{x}) &= \mathbf{x}^H \Sigma^{-1} \mathbf{x}, \\ \forall k, \forall t, \mathbf{S}_k^{(t)} &= \mathbf{x}_k^{(t)} \mathbf{x}_k^{(t)H}. \end{aligned} \quad (1)$$

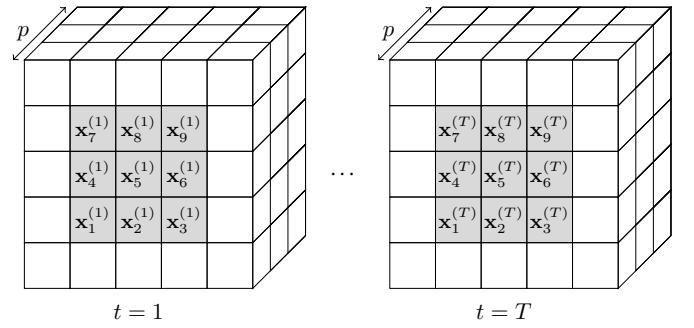


Figure 1. Illustration of local data selection ($N_1 = N_2 = p = 3$) for detection test. The gray area corresponds to $\mathcal{W}_{1,T}$ and the central pixel ($\mathbf{x}_5^{(t)}$) is the test pixel.

B. Data Model

Each pixel $\mathbf{x}_k^{(t)}$ of the SAR image is modelled as the realisation of a random multivariate random vector \mathbf{x} of size p with a given probability model denoted $p_{\mathbf{x}}(\mathbf{x}; \Omega(t))$, where $\Omega(t) = \{\boldsymbol{\theta}_1(t), \dots, \boldsymbol{\theta}_m(t)\} \in \Theta^m$ are the parameters of the Probability Density Function (PDF).

The Gaussian assumption is the most widely used in multivariate SAR image applications. Indeed, in those images, each pixel value is the sum of the contribution of many scatterers. Using the central limit theorem, the Gaussian assumption is the most natural one. This distribution is parametrised by a mean vector $\boldsymbol{\mu}$ and a covariance matrix Σ . The PDF is given by:

$$p_{\mathbf{x}}^{\mathcal{CN}}(\mathbf{x}; \boldsymbol{\mu}, \Sigma) = \frac{1}{\pi^p |\Sigma|} \exp(-q(\Sigma, \mathbf{x} - \boldsymbol{\mu})).$$

For SAR images, the mean is classically assumed to be zero due to the multiplicative nature of speckle noise: $\boldsymbol{\mu} = \mathbf{0}_p$ and will be omitted in the remainder of the paper. We will write $p_{\mathbf{x}}^{\mathcal{CN}}(\mathbf{x}; \Sigma) = p_{\mathbf{x}}^{\mathcal{CN}}(\mathbf{x}; \mathbf{0}_p, \Sigma)$.

The notation $\boldsymbol{\theta}(t) \stackrel{\Delta}{=} \boldsymbol{\theta}_t$ will be used henceforth.

C. Detection Schemes

CD in an ITS is a large problem and can be posed in many different mathematical terms. When using a parametrised probability model for the pixels, the problem is seen as a comparison of parameters over the time. In this paper, we express the CD problematic as:

Consider a Time Series of random vectors $\mathbf{x}^{(t)} \sim p_{\mathbf{x}}(\mathbf{x}; \Omega(t))$; given parameters of interest denoted $\boldsymbol{\theta} \subset \Omega$, choose between the two following alternatives:

$$\begin{cases} H_0 : \boldsymbol{\theta}_1 = \dots = \boldsymbol{\theta}_T = \boldsymbol{\theta}_0, \\ H_1 : \exists (t, t') \in \llbracket 1, T \rrbracket^2, \boldsymbol{\theta}_t \neq \boldsymbol{\theta}_{t'} \end{cases}. \quad (2)$$

Under this general formulation, a subset $\boldsymbol{\theta}$ of the PDF's parameters is considered. If the value of these parameters of interest changes over time, it is considered as a change in the time series. $\Phi_t = \Omega_t \setminus \boldsymbol{\theta}$ are parameters of the PDF which are not considered to be significant for the change in the time series.

Another scheme of interest is to choose between the two following alternatives:

$$\begin{cases} H_0^{\text{marg}} : \boldsymbol{\theta}_1 = \dots = \boldsymbol{\theta}_{T-1} = \boldsymbol{\theta}_0 \text{ and } \boldsymbol{\theta}_T = \boldsymbol{\theta}_0, \\ H_1^{\text{marg}} : \boldsymbol{\theta}_1 = \dots = \boldsymbol{\theta}_{T-1} = \boldsymbol{\theta}_{01} \text{ and } \boldsymbol{\theta}_T \neq \boldsymbol{\theta}_{01} \end{cases}. \quad (3)$$

This scheme's intent is to test only the last image of the series while considering that there is no change before. It is useful in an on-line detection problem, where we want to integrate the knowledge that there was no change in order to obtain better performance than a bi-date scheme. It was considered in [31], for example, where an estimation strategy for the change-point is presented. We will consider the derivation of statistic for this problem as well. However, we will limit ourselves to the study of statistics derived for scheme (2).

D. Statistics of decision under Gaussian model

Under Gaussian model, it is clear that, under CD schemes presented previously, the sole possibility is $\theta_t = \Sigma_t$, $\Phi_t = \emptyset$. Many works have studied this problem [14], [15], [16], [27] and many statistics have been proposed. The case for $T = 2$ has been especially studied [8], [9], [10]. Recently, [13] did a comparative study and showed that many statistics are statistically equivalent and reduced the options available to:

- the GLRT statistic:

$$\hat{\Lambda}_G = \frac{\left| \hat{\Sigma}_0^{\text{SCM}} \right|^{TN}}{\prod_{t=1}^T \left| \hat{\Sigma}_t^{\text{SCM}} \right|^N} \stackrel{H_1}{\gtrless} \lambda, \quad (4)$$

where:

$$\forall t, \hat{\Sigma}_t^{\text{SCM}} = \frac{1}{N} \sum_{k=1}^N \mathbf{S}_k^{(t)} \text{ and } \hat{\Sigma}_0^{\text{SCM}} = \frac{1}{T} \sum_{t=1}^T \hat{\Sigma}_t^{\text{SCM}}. \quad (5)$$

- the t_1 statistic which is obtained from Terrell [32] or Rao [33] tests:

$$\hat{\Lambda}_{t_1} = \frac{1}{T} \sum_{t=1}^T \text{Tr} \left[\left(\hat{\Sigma}_0^{\text{SCM}} \right)^{-1} \hat{\Sigma}_t^{\text{SCM}} \right] \stackrel{H_1}{\gtrless} \lambda. \quad (6)$$

- the Wald statistic [34]:

$$\begin{aligned} \hat{\Lambda}_{\text{Wald}} &= N \sum_{t=2}^T \text{Tr} \left[\left(\mathbf{I}_p - \hat{\Sigma}_1^{\text{SCM}} (\hat{\Sigma}_t^{\text{SCM}})^{-1} \right)^2 \right] \\ &- q \left(N \sum_{t=1}^T (\hat{\Sigma}_t^{\text{SCM}})^{-T} \otimes (\hat{\Sigma}_t^{\text{SCM}})^{-1}, \text{vec} \left(\sum_{t=2}^T \boldsymbol{\Upsilon}_t \right) \right) \stackrel{H_1}{\gtrless} \lambda, \end{aligned} \quad (7)$$

where

$$\boldsymbol{\Upsilon}_t = N \left((\hat{\Sigma}_t^{\text{SCM}})^{-1} - (\hat{\Sigma}_t^{\text{SCM}})^{-1} \hat{\Sigma}_1^{\text{SCM}} (\hat{\Sigma}_t^{\text{SCM}})^{-1} \right). \quad (8)$$

For the GLRT statistic at eq. (4), the marginal statistic for scheme (3) has been derived in [31]:

$$\hat{\Lambda}_G^{\text{marg}} = \frac{\left| \sum_{t=1}^T \hat{\Sigma}_t^{\text{SCM}} \right|^{TN}}{\left| \hat{\Sigma}_T^{\text{SCM}} \right|^{(T-1)N}} \stackrel{H_1}{\gtrless} \lambda. \quad (9)$$

The statistics presented in this section are done using a Gaussian model which do not take into account the heterogeneity of the data. Indeed, as suggested in the introduction, HR SAR images may have heterogeneous behaviour since the number of scatterers in each pixel is reduced. In the next section, we will consider a robust model taking into account such behaviour.

III. EXTENSION TO NON-GAUSSIAN MODEL

A. Data Model

To take into account the heterogeneity of the data, the SIRV model is classically used. It is obtained by introducing a scale factor term, also known as the texture: $\mathbf{x} \sim \sqrt{\tau} \mathbf{z}$, where

$\mathbf{z} \sim \mathbb{C}\mathcal{N}(\mathbf{0}_p, \Sigma)$ and τ follows a given distribution. Since we want, in this paper, to derive statistics which are robust to many classes of distributions, no prior is given on τ . Instead, we will consider the texture terms as *deterministic* unknown parameters:

$$p_{\mathbf{x}_k^{(t)}} \left(\mathbf{x}_k^{(t)}; \tau_k^{(t)}, \Sigma_t \right) = p_{\mathbf{x}_k^{(t)}}^{\mathbb{C}\mathcal{N}} \left(\mathbf{x}_k^{(t)}; \tau_k^{(t)} \Sigma_t \right).$$

In this model, there exists an indetermination between the texture and covariance matrix. Indeed, we have $\forall \alpha \in \mathbb{R}^{+*}, p_{\mathbf{x}}(\mathbf{x}; \tau, \Sigma) = p_{\mathbf{x}}(\mathbf{x}; \alpha\tau, \Sigma/\alpha)$. Classically, without loss of generality, Σ is assumed to be normalised: $\text{Tr}(\Sigma) = p$.

Under the robust model, the estimation of Σ has been studied in [35], [36], [37], [38]. When the texture are considered deterministic, [20] proposed an approximation in the form of a fixed-point estimator also known as Tyler estimator:

$$\hat{\Sigma}_t^{\text{TE}} = \frac{p}{N} \sum_{k=1}^N \frac{\mathbf{S}_k^{(t)}}{q(\hat{\Sigma}_t^{\text{TE}}, \mathbf{x}_k^{(t)})}. \quad (10)$$

In contrast to the Gaussian case, the estimates of covariance matrices are normalised by the trace to ensure the uniqueness of the solution: since the fixed-point's equation has solutions to a given scale factor (due to the ambiguity problem), [39] proposed to impose $\text{Tr}(\hat{\Sigma}) = p$.

In [28], it was proposed for the case $T = 2$, to use this estimator in the statistic of eq. (4) in place of the SCM estimator in order to obtain a robust distance between covariance matrices. In fact, when considering this methodology, the statistic loses its CFAR matrix property. Indeed, this is caused by the normalisation constraint described previously: the ratio is not invariant when the estimates of covariance matrices are scaled. Since the normalisation is performed by scaling the estimates $\hat{\Sigma}_t$ by $p/\text{Tr}(\hat{\Sigma}_t)$, the statistic introduces a ratio of trace terms which are not CFAR.

Moreover, this methodology omit the textures parameters which may be useful for accounting changes. Indeed, since the matrices are normalised, the relative power between the images is *contained* in the textures parameters.

In this paper, we consider the design of statistics by GLRT technique using the robust model rather than plugging robust estimates in Gaussian-derived statistics.

B. Problems Statement

Since in the robust model, the PDF is characterised by two unknown parameters, several detection strategies are possible. We consider the following problems:

- Problem 1:

$$\begin{aligned} \theta_t &= \left\{ \tau_1^{(t)}, \dots, \tau_N^{(t)}, \Sigma_t \right\}, \\ \Phi_t &= \emptyset. \end{aligned} \quad (11)$$

In this detection problem, we want to detect a change corresponding jointly to a change in power and in the shape of covariance matrix. This differs from the classic

Gaussian detection test (where the power is implicitly tested through the covariance matrix) as the heterogeneity of the texture on the window of observations is taken into account in the model.

- Problem 2:

$$\begin{aligned} \theta_t &= \left\{ \Sigma_t \right\}, \\ \Phi_t &= \left\{ \tau_1^{(t)}, \dots, \tau_N^{(t)} \right\}. \end{aligned} \quad (12)$$

In this next detection problem, we want to detect changes in the local correlations between the pixels without taking into account their relative power. This scheme is intended for applications in which an alteration in the power is not a significant change (for example two images of a scene with different calibrations). In those situations, Problem 1 is not suited.

- Problem 3:

$$\begin{aligned} \theta_t &= \left\{ \tau_1^{(t)}, \dots, \tau_N^{(t)} \right\}, \\ \Phi_t &= \left\{ \Sigma_t \right\}. \end{aligned} \quad (13)$$

In the last detection scheme, the detection is done solely on the texture parameter. This leads to a statistical test where only the relative power between the images is taken into account for CD.

IV. DERIVATION OF GLRT FOR PROBLEMS 1,2 AND 3

In this section, we derive the GLRT for each problem of III-B for both omnibus scheme (2) and marginal scheme (3).

A. GLRT of Problem 1

Proposition IV.1. *The GLRT ratio under hypotheses of Problem 1 for omnibus scheme (2) is the following:*

$$\hat{\Lambda}_{\text{MT}} = \frac{\left| \hat{\Sigma}_0^{\text{MT}} \right|^{TN}}{\prod_{t=1}^T \left| \hat{\Sigma}_t^{\text{TE}} \right|^N} \prod_{k=1}^N \frac{\left(\sum_{t=1}^T q(\hat{\Sigma}_0^{\text{MT}}, \mathbf{x}_k^{(t)}) \right)^{Tp}}{T^{Tp} \prod_{t=1}^T \left(q(\hat{\Sigma}_t^{\text{TE}}, \mathbf{x}_k^{(t)}) \right)^p} \stackrel{H_1}{\gtrless} \lambda, \quad (14)$$

where

$$\hat{\Sigma}_0^{\text{MT}} = f_{N,T}^{\text{MT}} \left(\hat{\Sigma}_0^{\text{MT}} \right) = \frac{p}{N} \sum_{k=1}^N \frac{\sum_{t=1}^T \mathbf{S}_k^{(t)}}{\sum_{t=1}^T q(\hat{\Sigma}_0^{\text{MT}}, \mathbf{x}_k^{(t)})}. \quad (15)$$

Proof. See Appendix A. A step by step derivation is also provided in a supplementary material. \square

Discussion: The statistic obtained here is similar to the one obtained using Gaussian assumption. The term involving determinant is the same except that now the estimates are solution of a fixed-point equation. $\hat{\Sigma}_t^{\text{TE}}$ is the Tyler estimator of eq. (10). $\hat{\Sigma}_0^{\text{MT}}$ is similar but corresponds to a different fixed-point equation involving the observations for all the dates. The properties of this new estimate will be studied in the next section.

Due to the normalisation of covariance matrices, the term involving determinants is a test involving solely the structure of the covariance matrices and do not consider the relative power of the pixels between the dates. The ratio of the quadratic forms allows to test the change in power in the same way it is done for the correlations in the determinants term.

Proposition IV.2. *The GLRT ratio under hypotheses of problem 1 for marginal scheme (3) is the following:*

$$\hat{\Lambda}_{\text{MT}}^{\text{marg}} = \frac{\left| \hat{\Sigma}_0^{\text{MT}} \right|^{TN}}{\left| \hat{\Sigma}_{01}^{\text{MT}} \right|^{(T-1)N} \left| \hat{\Sigma}_T^{\text{TE}} \right|^N} \frac{((T-1)p)^{(T-1)Np} p^{Np}}{(Tp)^{TNp}} \times \prod_{k=1}^N \frac{\left(\sum_{t=1}^T q(\hat{\Sigma}_0^{\text{MT}}, \mathbf{x}_k^{(t)}) \right)^{Tp}}{\left(\sum_{t=1}^{T-1} q(\hat{\Sigma}_{01}^{\text{MT}}, \mathbf{x}_k^{(t)}) \right)^{(T-1)p} \left(q(\hat{\Sigma}_T^{\text{TE}}, \mathbf{x}_k^{(T)}) \right)^p} \stackrel{H_1}{\gtrless} \lambda, \quad (16)$$

where

$$\hat{\Sigma}_{01}^{\text{MT}} = f_{N,T-1}^{\text{MT}}(\hat{\Sigma}_{01}^{\text{MT}}). \quad (17)$$

Proof. See Appendix B. A step by step derivation is also provided in a supplementary material. \square

B. GLRT of Problem 2

Proposition IV.3. *The GLRT ratio under hypotheses of problem 2 for omnibus scheme (2) is the following:*

$$\hat{\Lambda}_{\text{Mat}} = \frac{\left| \hat{\Sigma}_0^{\text{Mat}} \right|^{TN}}{\prod_{t=1}^T \left| \hat{\Sigma}_t^{\text{TE}} \right|^N} \prod_{k=1}^{k=N} \frac{\left(q(\hat{\Sigma}_0^{\text{Mat}}, \mathbf{x}_k^{(t)}) \right)^p}{\left(q(\hat{\Sigma}_t^{\text{TE}}, \mathbf{x}_k^{(t)}) \right)^p} \stackrel{H_1}{\gtrless} \lambda, \quad (18)$$

where

$$\hat{\Sigma}_0^{\text{Mat}} = f_{N,T}^{\text{Mat}}(\hat{\Sigma}_0^{\text{Mat}}) = \frac{p}{TN} \sum_{k=1}^{k=T} \frac{\mathbf{s}_k^{(t)}}{q(\hat{\Sigma}_0^{\text{Mat}}, \mathbf{x}_k^{(t)})}.$$

Proof. The detail of this calculation for $T = 2$ can be found in [30]. A step by step derivation is also provided in a supplementary material. \square

Discussion: The statistic obtained here is different from the previous one since there is no consideration of the relative power between the dates. Indeed, this time, the estimates of the texture parameter are compensated and only serve to test the correlations between the dates for each pixel of the window.

In fact, the same statistic is obtained if we consider the complex elliptical symmetric model [40] on the self-normalised observations: $\mathbf{x}_k^{(t)} / \|\mathbf{x}_k^{(t)}\|$. In this model the relative power is lost in the normalisation and the correlation structure is considered.

$\hat{\Sigma}_0^{\text{Mat}}$ is also a new variant of Tyler's estimator and its properties will be studied in the next section.

Proposition IV.4. *The GLRT ratio under hypotheses of problem 2 for marginal scheme (3) is the following:*

$$\hat{\Lambda}_{\text{Mat}}^{\text{marg}} = \frac{\left| \hat{\Sigma}_0^{\text{Mat}} \right|^{TN}}{\left| \hat{\Sigma}_{01}^{\text{Mat}} \right|^{(T-1)N} \left| \hat{\Sigma}_T^{\text{TE}} \right|^N} \times \prod_{k=1}^N \frac{\left(\prod_{t=1}^T \left(q(\hat{\Sigma}_0^{\text{Mat}}, \mathbf{x}_k^{(t)}) \right)^p \right)^p}{\left(\prod_{t=1}^{T-1} \left(q(\hat{\Sigma}_{01}^{\text{Mat}}, \mathbf{x}_k^{(t)}) \right)^p \right) \left(q(\hat{\Sigma}_T^{\text{TE}}, \mathbf{x}_k^{(T)}) \right)^p} \stackrel{H_1}{\gtrless} \lambda, \quad (19)$$

where

$$\hat{\Sigma}_{01}^{\text{Mat}} = f_{N,T-1}^{\text{Mat}}(\hat{\Sigma}_{01}^{\text{Mat}}).$$

Proof. The calculation is very similar to the one done at Proposition IV.2. A step by step derivation is also provided in a supplementary material. \square

C. GLRT of Problem 3

Proposition IV.5. *The GLRT ratio under hypotheses of problem 3 for omnibus scheme (2) is the following:*

$$\hat{\Lambda}_{\text{Tex}} = \prod_{t=1}^T \frac{\left| \hat{\Sigma}_t^{\text{Tex}} \right|^N}{\left| \hat{\Sigma}_t^{\text{TE}} \right|^N} \prod_{k=1}^N \frac{\left(\sum_{t=1}^T q(\hat{\Sigma}_t^{\text{Tex}}, \mathbf{x}_k^{(t)}) \right)^{Tp}}{T^{Tp} \prod_{t=1}^T \left(q(\hat{\Sigma}_t^{\text{TE}}, \mathbf{x}_k^{(t)}) \right)^p} \stackrel{H_1}{\gtrless} \lambda, \quad (20)$$

where

$$\hat{\Sigma}_t^{\text{Tex}} = f_{N,T,t}^{\text{Tex}}(\hat{\Sigma}_1^{\text{Tex}}, \dots, \hat{\Sigma}_T^{\text{Tex}}), \quad (21)$$

$$= \frac{Tp}{N} \sum_{k=1}^N \frac{\mathbf{s}_k^{(t)}}{\sum_{t'=1}^T q(\hat{\Sigma}_{t'}^{\text{Tex}}, \mathbf{x}_k^{(t)})}. \quad (22)$$

Proof. Very similar to the one of Proposition IV.1 presented in Appendix A. A step by step derivation is also provided in a supplementary material. \square

Discussion: In this last statistic, the detection is done solely on the texture parameters. This leads to an interesting estimation: each $\hat{\Sigma}_t^{\text{Tex}}$ is solution of a fixed-point equation which involves all the estimates $\hat{\Sigma}_{t'}^{\text{Tex}}$. In practice, this can lead to convergence issues when considering the computation. This problematic will be treated in the next section and it can be shown that the estimates can be implemented simply.

The marginal statistic is omitted for this problem. As we will show hereafter, $\hat{\Lambda}_{\text{Tex}}$ does not have the CFAR matrix property and is thus not an interesting statistic for schemes whose objective is to ensure a given significance level.

V. CONVERGENCE CONSIDERATIONS

A. Theoretical study of convergence

We consider here the validity of the alternate maximisation done when deriving the new statistics and the convergence problems that arise. To this end, we consider geodesic convexity (g-convexity) on the manifold $\mathbb{S}_{\mathbb{H}}^p$ as presented in [41] which is defined as follows:

Definition 1. (Geodesic convexity) Let \mathcal{M} be an arbitrary manifold. For each pair $q_0, q_1 \in \mathcal{M}$, we define a geodesic $q_t^{q_0, q_1} \in \mathcal{M}$ for $t \in [0, 1]$. A real valued function f with domain \mathcal{M} if g-convex if $f(q_t^{q_0, q_1}) \leq tf(q_1) + (1-t)f(q_0)$ for any $q_0, q_1 \in \mathcal{M}$ and $t \in [0, 1]$.

The g-convexity, which extends the definition of the traditional Euclidean convexity to curved spaces, is useful for optimisations done on covariances matrices. Notably, we can use this property of the log-likelihood to show the following proposition:

Proposition V.1. $\hat{\Sigma}_0^{\text{MT}}$, $\hat{\Sigma}_0^{\text{Mat}}$ and $\hat{\Sigma}_t^{\text{Tex}}$ are the unique arguments of the global maxima of their respective log-likelihood cost functions over the observations

Proof. See Appendix C. \square

This proposition is necessary to justify the alternate maximisation done when deriving the expression of the statistics. However, when considering optimisation on manifolds, this in itself does not guarantee that the solution corresponding to the global maxima is part of the manifold. This point is important since we want a solution that is both computable and in the set $\mathbb{S}_{\mathbb{H}}^p$. The following proposition can be effectively shown:

Proposition V.2. $\hat{\Sigma}_0^{\text{MT}}$, $\hat{\Sigma}_0^{\text{Mat}}$ and $\hat{\Sigma}_t^{\text{Tex}}$ are the arguments to the global minima obtained inside $\mathbb{S}_{\mathbb{H}}^p$.

Proof. See Appendix D. \square

Now that we know that the solution of the fixed-point equations are the unique arguments to the global maxima of their log-likelihood and that they are obtained inside the manifold $\mathbb{S}_{\mathbb{H}}^p$, the convergence of the fixed-point algorithms can be considered. We have:

Theorem V.3. Let $\{\mathbf{x}_k^{(t)} | k \in \llbracket 1, N \rrbracket, t \in \llbracket 1, T \rrbracket\}$ be a set of observations. Let us define vectors $\mathbf{v}_i \in \mathbb{R}^p$ such that $\forall k, \forall t, \mathbf{v}_{(T-1)*N+k} = (\Re(\mathbf{x}_k^{(t)})^\top, \Im(\mathbf{x}_k^{(t)})^\top)^\top$ and $\mathbf{v}_{(2T-1)*N+k} = (-\Im(\mathbf{x}_k^{(t)})^\top, \Re(\mathbf{x}_k^{(t)})^\top)^\top$. Let $\mathbb{P}_{2TN}(\bullet)$ be the empirical distribution of samples $\{\mathbf{v}_i | i \in \llbracket 1, 2TN \rrbracket\}$. Then the fixed-point algorithms $(\Sigma_0^{\text{MT}})_{k+1} = f_{N,T}^{\text{MT}}((\Sigma_0^{\text{MT}})_{k+1})$ and $(\Sigma_0^{\text{Mat}})_{k+1} = f_{N,T}^{\text{Mat}}((\Sigma_0^{\text{Mat}})_{k+1})$ converge to unique solutions up to a scale factor if and only if the following condition is respected:

(C1) $\mathbb{P}_{2TN}(\{\mathbf{0}\}) = 0$ and for all linear subspaces $V \subset \mathbb{R}^{2p}$, we have $\mathbb{P}_{2TN}(V) < \dim(V)/2p$.

Proof. This result can be obtained using the complex to real equivalence provided in [23] and by plugging the expression of the new estimator at eq. (15) in the proof of theorem 3 of

[42]. Since most of the proof is equivalent, it will be provided in a supplementary material.

The main steps of the proof are the following:

- By using the complex to real transformation presented in the theorem, we show that the problem can be considered using real valued observations.
- Since the fixed point equation is preserved using the transformation $\mathbf{x} \rightarrow M\mathbf{x}$, for any non-singular matrix M , we can assume $\Sigma = I_{2p}$ without loss of generality.
- To prove the sufficient statement, we show using condition (C1), $\Sigma = I_{2p}$ and an appropriate bounding, that the largest and smallest eigenvalues of $\hat{\Sigma}$ both converge to one.
- The necessary statement is obtained by defining a projector Q on a proper subspace. Multiplying the fixed point equation when $\Sigma = I_{2p}$ by $(I_{2p} - Q)$ and using appropriate bounding allows to obtain the condition on the dimension.

\square

For practical purposes the condition (C1) can be achieved when there are at least $p+1$ linearly independent observations $\mathbf{x}_k^{(t)}$, which is ensured in the data model we considered in the paper. Again, the uniqueness is guaranteed by the trace normalisation which has to be imposed at each step of the algorithm. It is important to notice that the convergence of the algorithms is ensured for any set of observations $\{\mathbf{x}_k^{(t)} | t \in \llbracket 1, T \rrbracket, k \in \llbracket 1, N \rrbracket\}$ that respects condition (C1), even if the observations do not follow the same distributions (typically if the hypothesis H_1 is correct).

The case of $\hat{\Sigma}_t^{\text{Tex}}$ is in this regard trickier. Indeed, since each step requires the knowledge of the others estimates, we propose the cyclic algorithm 1 that will iterate each matrix alternatively. While it is easy to show that if only one of the matrices is unknown¹, the fixed-point algorithm will converge, it is difficult to conclude on a theoretical standpoint about the convergence of the alternate estimation algorithm. Nonetheless, when doing extensive simulations, as will be shown shortly afterwards, on both theoretical and real-data, there has been no case when the algorithm do not converge, except when the condition (C1) is not respected.

B. Experimental study of convergence

In order to test the convergence property of matrix estimates, we consider realisations of random variables $\mathbf{x} = \sqrt{\tau} \mathbf{z}$ where τ follows a Γ -distribution with shape parameter α and scale parameter β . \mathbf{z} is generated through a Gaussian realisation with covariance matrix chosen to be Toeplitz of the form:

$$\Sigma = (\sigma_{m,n})_{\substack{1 \leq m \leq p, \\ 1 \leq n \leq p}},$$

where : $\sigma_{m,n} = \rho^{|m-n|}$.

We consider two settings: first, we generate a time series $\{\mathbf{x}_k^{(t)} | k \in \llbracket 1, N \rrbracket, t \in \llbracket 1, T \rrbracket\}$ where each $\mathbf{x}_k^{(t)}$ is distributed

¹using the same considerations as in the previous theorem.

Algorithm 1 Computation of $\hat{\Sigma}_t^{\text{Tex}}$

- 1: Initialize $\forall t \in [1, T]$, $\hat{\Sigma}_t^{\text{Tex}(0)} = \mathbf{I}_p$
- 2: **while** $\delta > \epsilon$ **do**
- 3: **for** $t \in [1, T]$ **do**
- 4: Compute:

$$\hat{\Sigma}_t^{\text{Tex}(n+1)} = f_{N,T,t}^{\text{Tex}} \left(\hat{\Sigma}_1^{\text{Tex}(n)}, \dots, \hat{\Sigma}_T^{\text{Tex}(n)} \right).$$
- 5: Impose Trace normalisation by:

$$\hat{\Sigma}_t^{\text{Tex}(n+1)} = \frac{p \hat{\Sigma}_t^{\text{Tex}(n+1)}}{\text{Tr}(\hat{\Sigma}_t^{\text{Tex}(n+1)})}.$$
- 6: **end for**
- 7: Compute criterion
- 8:
- 9: $\delta = \max \left\{ \frac{\| \hat{\Sigma}_t^{\text{Tex}(n+1)} - \hat{\Sigma}_t^{\text{Tex}(n)} \|}{\| \hat{\Sigma}_t^{\text{Tex}(n)} \|} / t \in [1, T] \right\}.$
- 9: **end while**

with the same covariance matrix Σ_0 . Then, we generate a time series where each $\mathbf{x}_k^{(t)}$ is distributed with a covariance matrix Σ_t different for each date. Figure 2 presents a Monte-Carlo (MC) simulation where the criterion δ of convergence is plotted against the number of iterations n of the fixed-point algorithm. The plot shows that for whatever the setting, all estimates converge since the criterion attains the working precision of the machine. We observe that $\hat{\Sigma}_t^{\text{Tex}}$ needs more iterations to converge. This was expected, since in this case three different matrices were estimated while for the others a single matrix was computed. These results comfort the theoretical considerations of V-A.

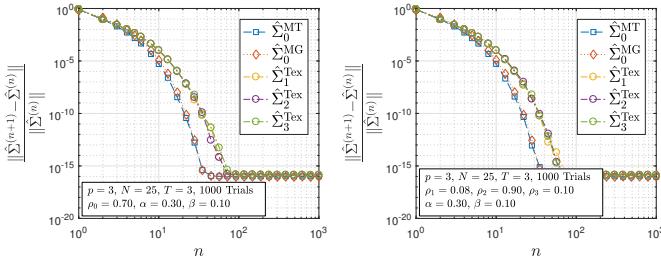


Figure 2. Convergence property of estimates. Left: Same matrices at each date. Right: Different matrices at each date. The textures are different at each date for both settings.

VI. STUDY OF CFAR PROPERTY

A. Theoretical study of the CFAR property

Here, we study the properties of the statistics derived in Section IV. We consider the CFAR property which is primordial if we want to apply the statistic in a decision scheme where the significance level is important.

We have the following propositions:

Proposition VI.1. $\hat{\Lambda}_{\text{MT}}$ (*resp.* $\hat{\Lambda}_{\text{MT}}^{\text{marg}}$) is CFAR texture and matrix for Problem 1 (*resp.* marginal Problem 1).

Proof. See Appendix E. \square

Proposition VI.2. $\hat{\Lambda}_{\text{Mat}}$ (*resp.* $\hat{\Lambda}_{\text{Mat}}^{\text{marg}}$) is CFAR texture and matrix for Problem 2 (*resp.* marginal Problem 2).

Proof. The same arguments as used in Proposition VI.1 are applied here. \square

Proposition VI.3. $\hat{\Lambda}_{\text{Tex}}$ is CFAR texture but is not CFAR matrix for Problem 3.

Proof. See Appendix F. \square

B. Experimental study of the CFAR property

The CFAR texture and matrix behaviour of the new statistics have been tested in simulation. To this end, a time series has been generated under the H_0 regime of Problem 1 which also corresponds to H_0 for the other problems. The statistics have been computed in MC trials to generate the plots shown at Figure 3. The Gaussian statistics of II-D have also been computed. The plots show that these Gaussian statistics vary when the texture changes and thus, have not the texture CFAR property. The new statistics however, do not vary for any texture parameter tested, which is an improvement. In this regards, $\hat{\Lambda}_{\text{Mat}}$ is the most robust one since the statistic does not vary even if the texture equality between the dates is not respected.

Next, the matrix CFAR behaviour is tested using $\forall k, \forall t, \tau_k^{(t)} = 1$. Figure 4 shows plots of MC trials where the coefficients for the covariance matrix vary. The plots show that the Gaussian statistics are CFAR which was demonstrated in [13]. It shows that $\hat{\Lambda}_{\text{MT}}$, $\hat{\Lambda}_{\text{Mat}}$ have the CFAR matrix behaviour while $\hat{\Lambda}_{\text{Tex}}$ has not. This result is coherent with the theoretical analysis.

VII. PERFORMANCE STUDY OF NEW STATISTICS

In this section, we present results of simulation performed on synthetic and real dataset. For synthetic data, the same methodology as the previous section is used for generating data. Table I reminds the parameters relevant for the simulation.

Table I
SIMULATION-RELEVANT PARAMETERS

α, β	ρ_t	p	N	T
Shape and Scale for Γ -distribution	Coefficients for Toeplitz matrices	Size of vector	Number of observations	Number of Images

A. Test of statistics on synthetic dataset

We consider analysing the theoretical performance of the new statistics. To this end, we consider a time series with $T = 10$, $p = 3$, $N = 7$ with a change at $t = 5$ and plot Receiver Operating Characteristic (ROC) curves for each problem.

- Problem 1: Before change, the covariance matrix is associated with $\rho = 0.1$. The textures are generated with

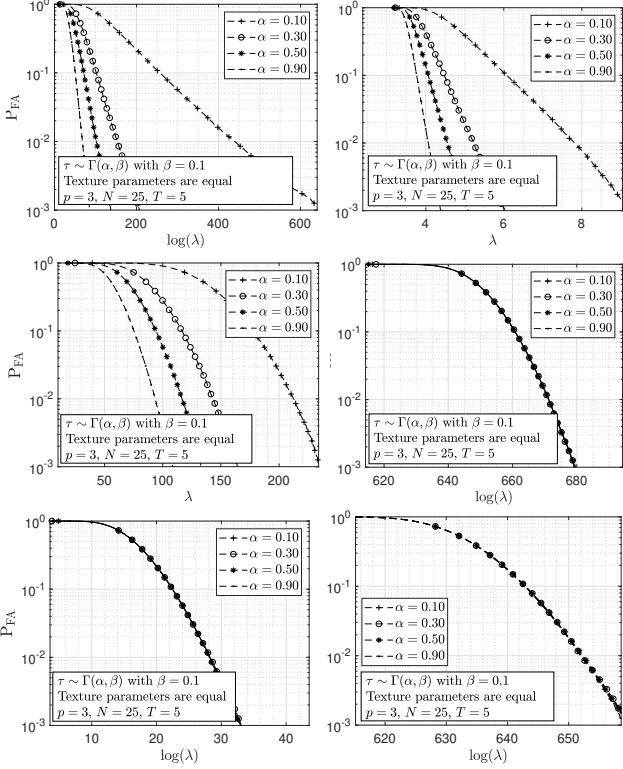


Figure 3. Texture CFAR behaviour. Top-left: $\hat{\Lambda}_G$. Top-right: $\hat{\Lambda}_{t_1}$. Middle-left: $\hat{\Lambda}_{Wald}$. Middle-right: $\hat{\Lambda}_{MT}$. Bottom-left: $\hat{\Lambda}_{Mat}$. Bottom-Right: $\hat{\Lambda}_{Tex}$. $\rho = 0.3$ at each date for all the curves.

$\alpha = 0.3, \beta = 0.1$ and are equal for each date. After the change, we have $\rho = 0.8$ and $\alpha = 0.3, \beta = 0.3$ and the textures are equal for each date.

- Problem 2: Before change, the covariance matrix is associated with $\rho = 0.1$. After the change, we have $\rho = 0.8$. For any date, the textures are generated using $\alpha = 0.3, \beta = 0.1$ and are different for each date.
- Problem 3: Before and after the change, the covariance matrix is obtained using a random value for ρ and the textures are generated using $\alpha = 0.3, \beta = 0.3$. Before change, the textures are equal and after the change, they are different.
- Gaussian problem: the textures are all fixed to one. Before the change, the covariance matrix is associated with $\rho = 0.1$. After the change, the covariance matrix is associated with $\rho = 0.8$.

Figure 5 gives the results obtained by MC trials. The thresholds for a given P_{Fa} are computed numerically using the H_0 regime of the problem considered. Although not realistic on real images, this allows, on synthetic data, to have an experimental threshold that matches the objective P_{Fa} even if the test is not CFAR for the problem considered.

For each problem, the statistic derived yields the best expected result. The Gaussian statistics have poorer performance than $\hat{\Lambda}_{MT}$ and $\hat{\Lambda}_{Tex}$ for testing a change in the texture. $\hat{\Lambda}_{Mat}$ performs the best when there is only a change in the covariance matrix shape. For the third problem, since there is no change in the matrices, it is not surprising that the detection rate is low.

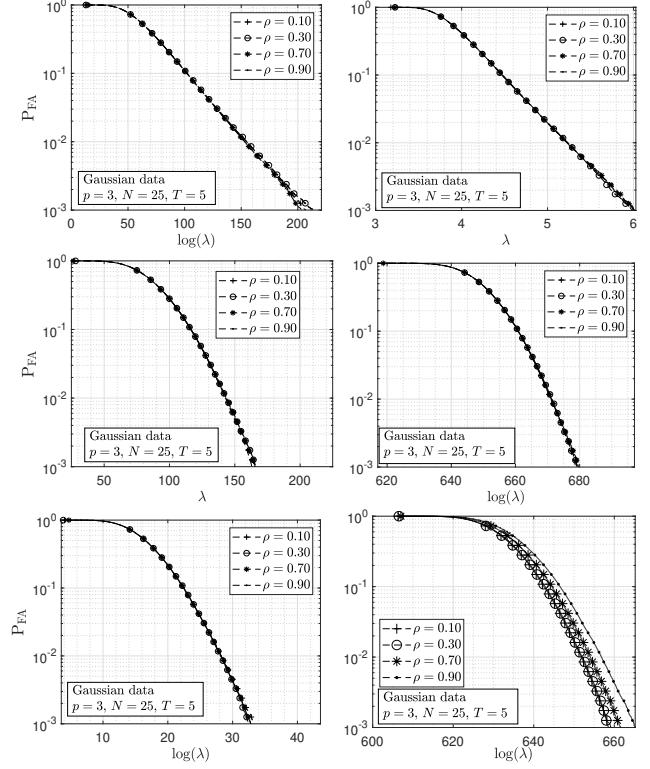


Figure 4. Matrix CFAR behaviour. Top-left: $\hat{\Lambda}_G$. Top-right: $\hat{\Lambda}_{t_1}$. Middle-left: $\hat{\Lambda}_{Wald}$. Middle-right: $\hat{\Lambda}_{MT}$. Bottom-left: $\hat{\Lambda}_{Mat}$. Bottom-Right: $\hat{\Lambda}_{Tex}$. $\rho = 0.3$ at each date for all the curves.

$\hat{\Lambda}_{MT}$ appears to be the best option for testing changes on the textures since for both problems 1 and 3, the performance are good. This is explained by the fact that the distribution under H_0 of the statistic is less sensitive to a violation of the matrix equality assumption than it would be from a texture one. Since in Problem 3 the textures are equal before the change, the threshold to guarantee the P_{Fa} is still low enough to guarantee good performance. Finally, when the data is strictly Gaussian, $\hat{\Lambda}_{t_1}$ and $\hat{\Lambda}_G$ have better results than the robust statistics. This result is expected, since there is a trade-off between robustness and performance when considering robust methods. Among the new statistics, $\hat{\Lambda}_{Tex}$ does not allow do detect a change in the shape so its results are expected to be lowest.

B. Test of statistics on real dataset

1) *Data description:* The proposed statistics have been tested on real images coming from two different datasets: SDMS (Courtesy AFRL/RYA) [43] and UAVSAR (Courtesy NASA/JPL-Caltech). From SDMS, three images of the same scene, presented at Figure 6, are used. The ground truth is obtained from [44] for the two dates and [13] for the three dates. From UAVSAR, two scenes with two images each are used. They are presented in Figure 7. The ground truth is collected from [45], [46]. Table II gives an overall perspective of the scenes used in the study.

2) *First analysis:* We first try the various Gaussian and new statistics on the three dates of SDMS and on the Scene 1 of UAVSAR dataset. Figure 8 gives the results relative to the statistics for SDMS data and Figure 9 for UAVSAR. The

Table II
DESCRIPTION OF SAR DATA USED

Dataset	Url	Resolution	Scene	p	T	Size	Coordinates (top-left px)
SDMS CCD Challenge	https://www.sdms.afrl.af.mil/	Rg: 0.2m Az: 0.2m	FP0121-FP0124	3	2	1000 × 1000 px	[Rg, Az] = [2055, 2055]
			FP0120-FP0121-FP0124	3	3	1000 × 1000 px	[Rg, Az] = [2055, 2055]
UAVSAR SanAnd_26524_03 Segment 4 April 23, 2009 - May 11, 2015	https://uavstar.jpl.nasa.gov	Rg: 1.67m Az: 0.6m	Scene 1	3	2	2360 × 600 px	[Rg, Az] = [2891, 28891]
			Scene 2	3	2	2300 × 600 px	[Rg, Az] = [3236, 25601]

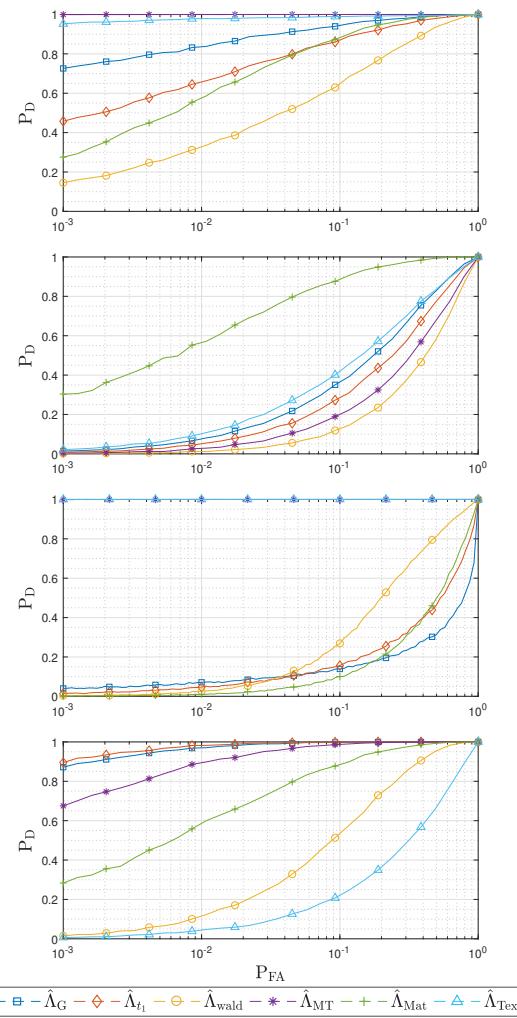


Figure 5. ROC curves obtained on synthetic data. Top: Problem 1. Middle-Up: Problem 2. Middle-down: Problem 3. Bottom: Gaussian setting.

values for $\hat{\Lambda}_{t_1}$ and $\hat{\Lambda}_{Wald}$ are omitted since they have similar behaviour than $\hat{\Lambda}_G$. Qualitatively, each statistic is high at the location of the changes given by the ground truth. For both dataset, $\hat{\Lambda}_{Mat}$ seems to have poorer performance since the values of the statistic are not much higher on the changes compared to the background. For UAVSAR data, a linear pattern appears in the bottom-right corner and responds highly

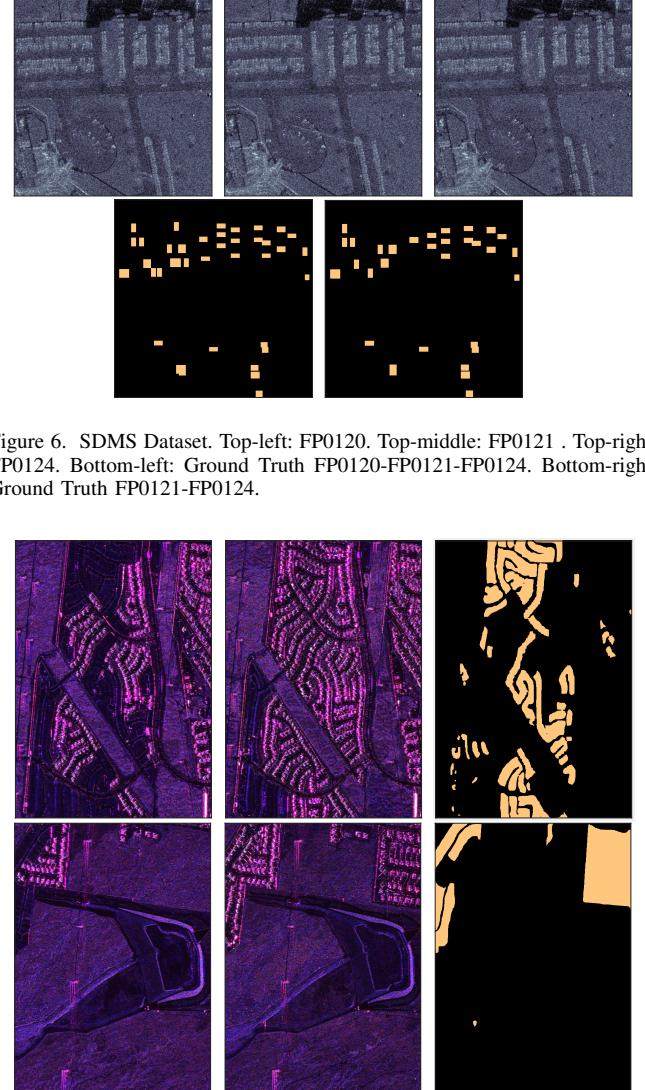


Figure 6. SDMS Dataset. Top-left: FP0120. Top-middle: FP0121 . Top-right: FP0124. Bottom-left: Ground Truth FP0120-FP0121-FP0124. Bottom-right: Ground Truth FP0121-FP0124.

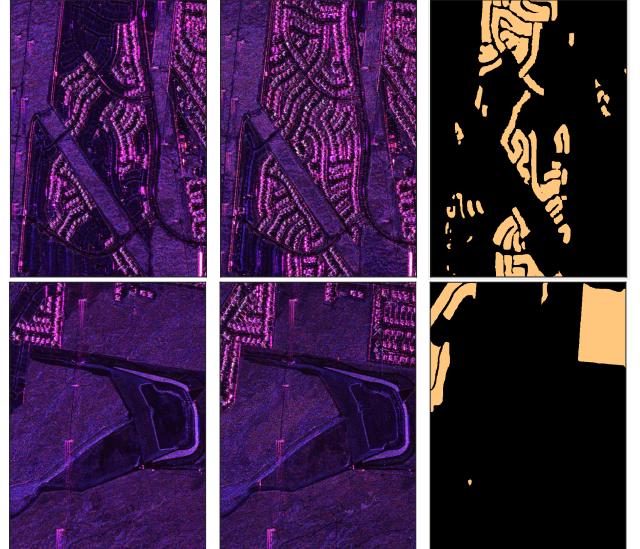


Figure 7. UAVSAR Dataset in Pauli representation. Left: April 23, 2009. Middle: May 15, 2011. Right: Ground Truth. Top: Scene 1. Bottom: Scene 2.

for all detectors except for $\hat{\Lambda}_{MT}$ and $\hat{\Lambda}_{Tex}$. However, it is difficult to conclude solely on those qualitative terms.

To quantify the performance of the statistics, experimental ROC curves are plotted using the Ground truth, denoted

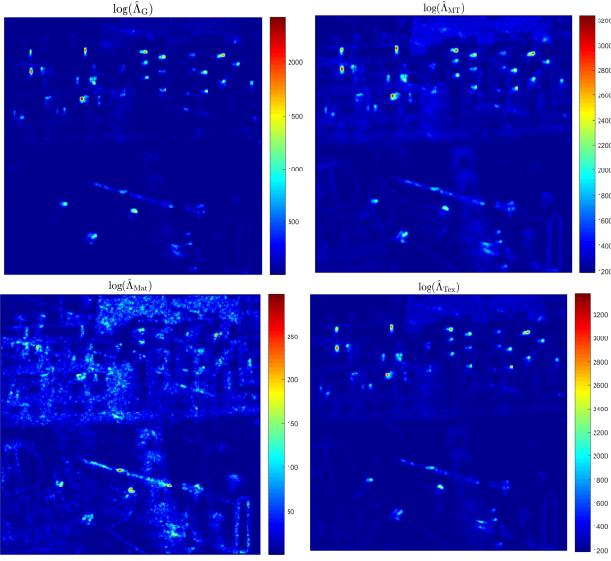


Figure 8. Value of the different statistics for SDMS FP0120-FP0121-FP0124.
 $p = 3$, $N_1 = N_2 = 11$.

$\mathbf{I}_{\text{GT}}(x, y)$ associated with spatial coordinates (x, y) , by computing the following:

- Probability of false alarm:

$$P_{FA} = N_{FD}/N_{NC},$$

$$\text{where: } N_{\text{FD}} = \sum_{x,y} \left(\hat{\Lambda}(x,y) \geq \lambda \right) \times (1 - \mathbb{1}_{\text{Gt}}(x,y)) ,$$

$$N_{\text{NC}} = \sum_{x,y} (1 - \mathbb{1}_{\text{Gt}}(x,y)) .$$

- Probability of detection:

$$P_D = N_{GD}/N_C,$$

$$\text{where: } N_{\text{GD}} = \sum_{x,y} \left(\hat{\Lambda}(x,y) \geq \lambda \right) \times \mathbb{1}_{\text{Gt}}(x,y),$$

$$N_{\text{C}} = \sum_{x,y} \mathbb{1}_{\text{Gt}}(x,y).$$

Figure 10 shows the results for each dataset and a size of analysis window of $N_1 = N_2 = 11$. It appears that $\hat{\Lambda}_{MT}$ has the overall best performance: it has similar results to $\hat{\Lambda}_G$ on SDMS dataset but performs better on UAVSAR dataset. Each Gaussian-derived statistic has similar performance but $\hat{\Lambda}_G$ appears to have better results than $\hat{\Lambda}_{t_1}$ and $\hat{\Lambda}_{Wald}$.

3) *ROC plots*: These results can be interpreted as follows: on the SDMS dataset, while the resolution is high, the images are globally homogeneous. In fact, much of the details are not visible and the objects appear to be blurry. This means that in practice, on a small local neighbourhood, the Gaussian model is accurate and thus that the Gaussian-derived statistics perform well. Nonetheless, the new statistics, except $\hat{\Lambda}_{\text{Mat}}$, do not have lower performance and can still have better performance when the size of the neighbourhood chosen is high as will be shown afterwards. On the other hand, the objects are better resolved on the UAVSAR. The transitions

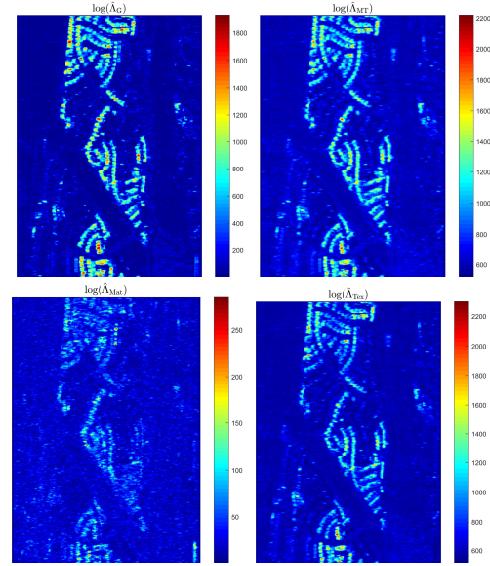


Figure 9. Value of the different statistics for UAVSAR Scene 1. $p = 3$, $N_1 = N_2 = 11$.

are sharper which means that an heterogeneous model is more accurate and thus that the new statistics will perform better. The difference of performance for UAVSAR scene 2 can be explained by the fact that dynamic between the darker zones and the bright ones is much higher than in the scene 1.

For the datasets used in this paper, $\hat{\Lambda}_{\text{Mat}}$ does not perform well. This is due to the fact that the detection omits the texture parameters which are responsible for the power. In these datasets, the ground truth corresponds to the arrival or disappearance of strong scatterers and thus, the power has an important role. As explained before, $\hat{\Lambda}_{\text{Mat}}$ allows to detect changes which are focused on the correlation structures and is not appropriate for those kinds of change.

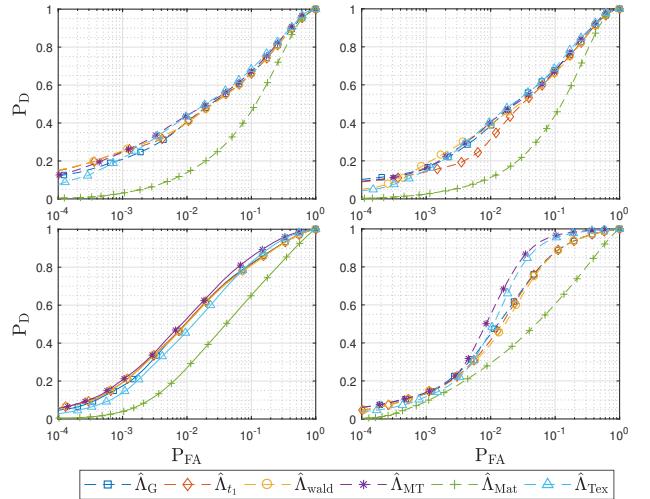


Figure 10. P_D versus P_{FA} on real data. Top-left: SDMS FP0121-FP0124. Top-right: FP0120-FP0121-FP0124. Bottom-left: UAVSAR Scene1. Bottom-right: UAVSAR Scene 2. For all images, $p = 3$, $N_1 = N_2 = 11$.

4) *Increasing the size of window:* In order to test the impact of the size of the analysis window, we fix an experimental $P_{FA} = 10^{-2}$ and plot the P_D against the size of the window. Figure 11 gives the results for all datasets.

By increasing the size of window, the detection rate improves. It can be explained by the fact that the estimation step has been performed on more data and is thus more precise. The drawback is that the detection is obtained with a lower spatial resolution.

When increasing the size of the window, $\hat{\Lambda}_{MT}$ and $\hat{\Lambda}_{Tex}$ perform better than the Gaussian statistics, especially on UAVSAR Scene 2. This is expected, since increasing the size of the window means that the data are spread over a large spatial leading to an increase of the heterogeneity due to the presence of many scatterers in the scene.

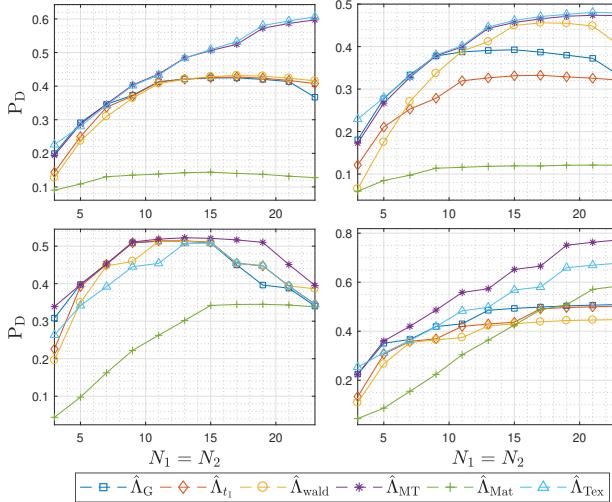


Figure 11. P_D as a function of window size at $P_{FA} = 10^{-2}$. Top-left: SDMS FP0121-FP0124. Top-right: FP0120-FP0121-FP0124. Bottom-left: UAVSAR Scene1. Bottom-right: UAVSAR Scene2. For all images, $p = 3$.

5) *Increasing the dimension of pixels:* Finally we consider the performance if the size of vector p increases. To this end, we exploit the wavelet decomposition method presented in [24], [47] which allows to decompose a monovariate SAR image into canals corresponding to a physical behaviour of the scatterers. Using this decomposition on all polarimetric canals of SDMS dataset allows to have an image with $p = 27$. The decomposition is not performed on UAVSAR dataset, since it does not exhibit a physical diversity using the wavelet decomposition. Figure 12 gives the result of CD for all the statistics. When compared to the performance using solely polarimetric information, it appears that using this method, the performance are lower when the P_{FA} is very low, while they are improved for $P_{FA} > 10^{-1}$. The case of $\hat{\Lambda}_{Tex}$, which has significantly better performance, highlights again that the texture parameter plays a main role in CD applications.

VIII. CONCLUSION

In this paper, we considered the problem of CD in an ITS of heterogeneous SAR images. by taking into account the heterogeneity through a SIRV model, we proposed three

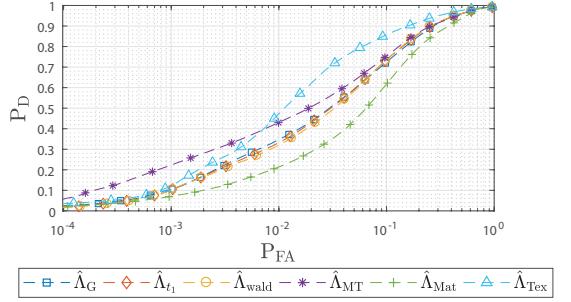


Figure 12. P_D versus P_{FA} on SDMS FP0121-FP0124 with wavelet decomposition.

detection schemes and derived statistics using GLRT techniques. The convergence and consistency properties of the estimates have been considered and the CFAR properties of these statistics have been studied on a theoretical standpoint and in simulation. $\hat{\Lambda}_{Tex}$ does not have the matrix CFAR property which is essential in many applications.

The statistics have been applied in simulation, where each statistic has good results under the conditions of its detection problem. $\hat{\Lambda}_{Mat}$ has proven to have the most robust behaviour. Finally, the statistics have been tested on real SAR data and $\hat{\Lambda}_{MT}$ has obtained the overall best performance in terms of detection. This highlights that for many changes, the texture information has to be taken into account in the detection problem.

APPENDIX A PROOF OF PROPOSITION IV.1

Proof. The GLRT assumes computing the two following likelihoods:

$$\begin{aligned}\mathcal{L}_0 &= p_{\mathcal{W}_{1,T}}(\mathcal{W}_{1,T}; \boldsymbol{\theta}_0, \Phi_1, \dots, \Phi_T), \\ \mathcal{L}_1 &= p_{\mathcal{W}_{1,T}}(\mathcal{W}_{1,T}; \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_T, \Phi_1, \dots, \Phi_T),\end{aligned}$$

where

$$\begin{aligned}\boldsymbol{\theta}_0 &= \{\tau_1, \dots, \tau_N, \Sigma_0\}, \\ \forall t, \boldsymbol{\theta}_t &= \left\{ \tau_1^{(t)}, \dots, \tau_N^{(t)}, \Sigma_t \right\}, \\ \forall t, \Phi_t &= \emptyset,\end{aligned}$$

and then computing

$$\max_{\{\boldsymbol{\theta}_0, \Phi_1, \dots, \Phi_T\}} \mathcal{L}_0 \text{ and } \max_{\{\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_T, \Phi_1, \dots, \Phi_T\}} \mathcal{L}_1.$$

- Let us consider \mathcal{L}_0 first. Since the observations are assumed to be independent, we have

$$\mathcal{L}_0 = \prod_{k=1}^{t=T} p_{\mathbf{x}_k^{(t)}}^{\mathbb{CN}} \left(\mathbf{x}_k^{(t)}; \tau_k \Sigma_0 \right).$$

Then in order to maximise \mathcal{L}_0 , we consider the optimisation:

$$\hat{\boldsymbol{\theta}}_0 = \underset{\boldsymbol{\theta}_0}{\operatorname{argmax}} \log \mathcal{L}_0(\boldsymbol{\theta}_0).$$

This is done by optimising separately for each separate parameter, assuming the others being constant, and then plugging back the estimates when needed. The validity

of this methodology is tackled in section V.

We have:

$$\begin{aligned} \log \mathcal{L}_0 = & -\pi^{TNp} - TN \log |\Sigma_0| - T p \sum_{k=1}^N \log(\tau_k) \\ & - \sum_{\substack{k=N \\ t=1 \\ k=1}}^{t=T} \frac{q(\Sigma_0, \mathbf{x}_k^{(t)})}{\tau_k}. \end{aligned} \quad (23)$$

Let $k \in \llbracket 1, N \rrbracket$, we solve:

$$\frac{\partial \log \mathcal{L}_0}{\partial \tau_k} = -Tp \sum_{k=1}^N \frac{1}{\tau_k} + \sum_{t=1}^T \frac{q(\Sigma_0, \mathbf{x}_k^{(t)})}{\tau_k^2} = 0,$$

which yields:

$$\hat{\tau}_k = \frac{1}{Tp} \sum_{t=1}^T q(\Sigma_0, \mathbf{x}_k^{(t)}). \quad (24)$$

Recall complex differentiation results [20]:

$$\begin{aligned} \frac{\partial \log |\Sigma|}{\partial \Sigma} &= \Sigma^{-1}, \\ \frac{\partial q(\Sigma, \mathbf{x}_k^{(t)})}{\partial \Sigma} &= -\mathbf{S}_k^{(t)} \Sigma^{-2}. \end{aligned} \quad (25)$$

We solve:

$$\frac{\partial \log \mathcal{L}_0}{\partial \Sigma_0} = -TN \Sigma_0^{-1} + \sum_{\substack{k=N \\ t=1}}^{k=N} \frac{\mathbf{S}_k^{(t)}}{\tau_k} \Sigma_0^{-2} = \mathbf{0}_{p^2},$$

which yields:

$$\hat{\Sigma}_0 = \frac{1}{TN} \sum_{\substack{k=1 \\ k=1}}^{k=N} \frac{\mathbf{S}_k^{(t)}}{\tau_k}. \quad (26)$$

Then by plugging back the estimates of textures at eq. (24), we obtain the expression given at eq. (17) that we denote $\hat{\Sigma}_0^{\text{MT}}$. We have:

$$\hat{\theta}_0 = \left\{ \hat{\tau}_1, \dots, \hat{\tau}_N, \hat{\Sigma}_0^{\text{MT}} \right\}. \quad (27)$$

- Now for \mathcal{L}_1 , we consider the same procedure:

$$\left\{ \hat{\theta}_1, \dots, \hat{\theta}_T \right\} = \underset{\{\theta_1, \dots, \theta_T\}}{\text{argmax}} \log \mathcal{L}_1(\theta_1, \dots, \theta_T).$$

We have:

$$\begin{aligned} \log \mathcal{L}_1 = & -\pi^{TNp} - N \sum_{t=1}^T \log |\Sigma_t| - p \sum_{k=1}^{k=N} \log \left(\tau_k^{(t)} \right) \\ & - \sum_{\substack{k=N \\ t=1 \\ k=1}}^{t=T} \frac{q(\Sigma_t, \mathbf{x}_k^{(t)})}{\tau_k^{(t)}}. \end{aligned} \quad (28)$$

Let $k \in \llbracket 1, N \rrbracket, t \in \llbracket 1, T \rrbracket$, solving

$$\frac{\partial \log \mathcal{L}_1}{\partial \tau_k^{(t)}} = 0,$$

yields:

$$\hat{\tau}_k^{(t)} = \frac{1}{p} q(\Sigma_t, \mathbf{x}_k^{(t)}). \quad (29)$$

Let $t \in \llbracket 1, T \rrbracket$, by solving

$$\frac{\partial \log \mathcal{L}_1}{\partial \Sigma_t} = \mathbf{0}_{p^2},$$

and by plugging estimates of eq. (29), we obtain the expression of $\hat{\Sigma}_t$ given at eq. (17) that we denote $\hat{\Sigma}_t^{\text{MT}}$. We have:

$$\hat{\theta}_t = \left\{ \hat{\tau}_1^{(t)}, \dots, \hat{\tau}_N^{(t)}, \hat{\Sigma}_t^{\text{MT}} \right\}. \quad (30)$$

Finally, the closed form of the statistic is obtained by:

$$\hat{\Lambda}_{\text{MT}} = \frac{\mathcal{L}_1(\hat{\theta}_1, \dots, \hat{\theta}_T)}{\mathcal{L}_0(\hat{\theta}_0)} \underset{\text{H}_0}{\gtrless} \lambda.$$

The cumbersome calculation is omitted and yields the expression given at eq. (14). \square

APPENDIX B PROOF OF PROPOSITION IV.2

Proof. The GLRT assumes computing the two following likelihoods:

$$\begin{aligned} \mathcal{L}_0 &= p_{\mathcal{W}_{1,T}}(\mathcal{W}_{1,T}; \theta_0, \Phi_1, \dots, \Phi_T), \\ \mathcal{L}_1 &= p_{\mathcal{W}_{1,T}}(\mathcal{W}_{1,T}; \theta_{01}, \theta_T, \Phi_1, \dots, \Phi_T), \end{aligned}$$

where

$$\begin{aligned} \theta_0 &= \{\tau_1, \dots, \tau_N, \Sigma_0\}, \\ \theta_{01} &= \{\tau_1^{(01)}, \dots, \tau_N^{(01)}, \Sigma_{01}\}, \\ \theta_T &= \{\tau_1^{(T)}, \dots, \tau_N^{(T)}, \Sigma_T\}, \\ \forall t, \Phi_t &= \emptyset, \end{aligned}$$

and then computing

$$\max_{\{\theta_0, \Phi_1, \dots, \Phi_T\}} \mathcal{L}_0 \text{ and } \max_{\{\theta_{01}, \theta_T, \Phi_1, \dots, \Phi_T\}} \mathcal{L}_1.$$

Using the optimisation methodology of Appendix A, $\hat{\theta}_0$ is the same as eq. (27), $\hat{\theta}_T$ is obtained by eq. (30) and $\hat{\theta}_{01}$ is obtained from eq. (27) where T is replaced by $T-1$ in all estimates.

Then the statistic is given by:

$$\hat{\Lambda}_{\text{MT}}^{\text{marg}} = \frac{\mathcal{L}_1(\hat{\theta}_{01}, \hat{\theta}_T)}{\mathcal{L}_0(\hat{\theta}_0)} \underset{\text{H}_0}{\gtrless} \lambda.$$

The cumbersome calculation is omitted and yields the expression at eq. (16). \square

APPENDIX C
PROOF OF PROPOSITION V.1

We will consider here only the proof for $\hat{\Sigma}_0^{\text{MT}}$, since the same procedure can be applied to show the property for the others. In other words, we will show that $\hat{\tau}_k$ at eq. (24) and $\hat{\Sigma}_0^{\text{MT}}$ at eq. (26) are the global maxima of the following log-likelihood function:

$$\begin{aligned} \log \mathcal{L}(\tau_1, \dots, \tau_N, \Sigma_0) = & -\pi^{TNp} - TN \log |\Sigma_0| - \\ & Tp \sum_{k=1}^N \log(\tau_k) - \sum_{\substack{k=1 \\ t=1}}^{t=T} \frac{q(\Sigma_0, \mathbf{x}_k^{(t)})}{\tau_k}. \end{aligned} \quad (31)$$

We first recall two useful results:

Lemma 1. Any local minimum of a g-convex function over \mathcal{M} is a global minimum.

Lemma 2. Consider the manifold $\mathbb{S}_{\mathbb{H}}^p$ and the following geodesic:

$$\Sigma_t^{\Sigma_0, \Sigma_1} = \Sigma_0^{\frac{1}{2}} \left(\Sigma_0^{-\frac{1}{2}} \Sigma_1 \Sigma_0^{-\frac{1}{2}} \right)^t \Sigma_0^{\frac{1}{2}}, \quad t \in [0, 1],$$

between two points Σ_0, Σ_1 .

Let $\mathbf{h}_i \in \mathbb{C}^p$, $a \in \pm 1$, $a' \in \pm 1$ for $i = 1, \dots, m$ and $\mathbf{H}_i \in \mathbb{C}^{q,p}$ for $i = 1, \dots, n$. The function

$$\mathcal{L}(\Sigma) = \log \left| \sum_{i=1}^n \mathbf{H}_i \Sigma^a \mathbf{H}_i^H \right| + \sum_{i'=1}^m \mathbf{h}_{i'}^H \Sigma^{a'} \mathbf{h}_{i'}, \quad (32)$$

is strictly g-convex in $\Sigma \in \mathbb{S}_{\mathbb{H}}^p$.

When looking at the negative of function $\log \mathcal{L}$ in eq. (31), straightforward application of Lemma 2 allows to conclude that it is jointly g-convex in Σ_0 and for all τ_k :

- the g-convexity for each τ_k is obtained by rewriting the negative of eq. (31) in the form of eq. (32), if we take²:

$$\begin{aligned} \Sigma &= \tau_k, \\ a &= 1, \\ a' &= -1, \\ \{\mathbf{h}_i | i \in \llbracket 1, m \rrbracket\} &= \left\{ \mathbf{x}_k^{(t)} \Sigma_0^{-\frac{1}{2}} | k \in \llbracket 1, N \rrbracket, t \in \llbracket 1, T \rrbracket \right\}, \\ \mathbf{H}_i &= \tau_i \delta_{ik}. \end{aligned}$$

- the g-convexity in Σ_0 is obtained by rewriting the negative of eq. (31) in the form of eq. (32), if we take³:

$$\begin{aligned} \Sigma &= \Sigma_0, \\ a &= 1, \\ a' &= -1, \\ \{\mathbf{h}_i | i \in \llbracket 1, m \rrbracket\} &= \left\{ \frac{\mathbf{x}_k^{(t)}}{\sqrt{\tau_k}} | k \in \llbracket 1, N \rrbracket, t \in \llbracket 1, T \rrbracket \right\}, \\ \mathbf{H}_i &= \delta_{i1}. \end{aligned}$$

So we have the strict g-convexity, application of Lemma 1 allows us to conclude that the estimates correspond to unique global maxima.

²Considering solely the terms involving the considered texture parameter.

³Considering solely the terms involving the covariance matrix.

APPENDIX D
PROOF OF PROPOSITION V.2

Again, we will only consider the case of $\hat{\Sigma}_0^{\text{MT}}$, since the same considerations lead to the result for the others. Up to now we have only shown that the negative log-likelihood $-\log \mathcal{L}$ is g-convex. To show that it has a unique minimum in $\mathbb{S}_{\mathbb{H}}^p$, and thus that the fixed-point equation to Σ_0^{MT} admits a unique solution within the manifold, it suffices to show that the minimum of $\log \mathcal{L}$ occurs in the interior of $\mathbb{S}_{\mathbb{H}}^p$. To this end we have to show that $\log \mathcal{L}(\Sigma) \rightarrow \infty$ as $\Sigma \rightarrow \text{Bound}(\mathbb{S}_{\mathbb{H}}^p)$, the boundary of $\mathbb{S}_{\mathbb{H}}^p$.

Let $\lambda_1(\Sigma), \dots, \lambda_p(\Sigma)$, be the ordered eigenvalues of Σ . We can rewrite $-\log \mathcal{L}$ as⁴:

$$TN \sum_{j=1}^p \log \lambda_j(\Sigma) + \sum_{\substack{k=1 \\ t=1}}^{t=T} \frac{q(\Sigma, \mathbf{x}_k^{(t)})}{\tau_k}.$$

Now, decomposing Σ as $\Sigma \underset{\text{EVD}}{=} \mathbf{P}^H \mathbf{D} \mathbf{P}$ we can write $q(\Sigma, \mathbf{x}_k^{(t)})$ as:

$$\sum_{j=1}^p \left| [\mathbf{y}_k^{(t)}]_j \right|^2 / \lambda_j(\Sigma),$$

where $[\mathbf{y}_k^{(t)}]_j$ is the j -th element of $\mathbf{y}_k^{(t)} = \mathbf{P}^H \mathbf{x}_k^{(t)}$. Then we have:

$$-\log \mathcal{L}(\Sigma) \geq \sum_{\substack{k=1 \\ j=p}}^{k=N} \frac{\left| [\mathbf{y}_k^{(t)}]_j \right|^2}{\lambda_j(\Sigma) \tau_k} + TN \sum_{j=1}^p \log \lambda_j(\Sigma),$$

Finally, $\Sigma \rightarrow \text{Bound}(\mathbb{S}_{\mathbb{H}}^p)$ if and only if $\lambda_1(\Sigma) \rightarrow \infty$ and/or $\lambda_p(\Sigma) \rightarrow 0$. Under both regimes the right-hand side of the previous equation goes to ∞ , which concludes the proof.

APPENDIX E
PROOF OF PROPOSITION VI.1

Proof. We consider separately the texture and matrix properties:

- *Texture CFAR:* First, $\hat{\Sigma}_0^{\text{MT}}$ and $\hat{\Sigma}_t^{\text{MT}}$ are invariant by the substitution $\mathbf{x}_k^{(t)} \rightarrow \mathbf{x}_k^{(t)} / \tau_k^{(0)}$. Then, the different

$$\frac{\left(\sum_{t=1}^T q(\hat{\Sigma}_0^{\text{MT}}, \mathbf{x}_k^{(t)}) \right)^{Tp}}{T^{Tp} \prod_{t=1}^T \left(q(\hat{\Sigma}_t^{\text{MT}}, \mathbf{x}_k^{(t)}) \right)^p}$$

terms are also invariant by the same substitution. This means that the values of $\{\tau_k^{(0)} | k \in \llbracket 1, N \rrbracket\}$ do not affect the statistic of $\hat{\Lambda}_{\text{MT}}$, which is the definition of texture CFAR property in this problem.

- *Matrix CFAR:* As said in the discussion of IV-A, the estimates of matrices are subject to an indetermination

⁴Omitting the constants with regards to Σ_0

which is resolved by an appropriate normalisation. For any estimate $\hat{\Sigma} \in \{\hat{\Sigma}_0^{\text{MT}}, \hat{\Sigma}_1^{\text{MT}}, \dots, \hat{\Sigma}_T^{\text{MT}}\}$, when replacing $\hat{\Sigma}$ by $p\hat{\Sigma}/\text{Tr}(\hat{\Sigma})$ in eq. (14), the trace terms simplify in the expression of $\hat{\Lambda}_{\text{MT}}$. Thus, the statistic is homogeneous by the normalisation constraint.

Then, the statistic is invariant for the group of transformation:

$$\mathcal{G} = \left\{ \mathbf{G} \mathbf{x}_k^{(t)} \mid t \in \llbracket 1, T \rrbracket, k \in \llbracket 1, N \rrbracket, \mathbf{G} \in \mathbb{S}_{\mathbb{H}}^p \right\}.$$

Indeed, we can write:

$$\hat{\Lambda}_{\text{MT}} = \frac{\left| \mathbf{G} \hat{\Sigma}_0^{\text{MT}} \mathbf{G} \right|^{TN}}{\prod_{t=1}^T \left| \mathbf{G} \hat{\Sigma}_t^{\text{TE}} \mathbf{G} \right|^N} \prod_{k=1}^N \frac{\left(\sum_{t=1}^T q(\hat{\Sigma}_0^{\text{MT}}, \mathbf{G} \mathbf{x}_k^{(t)}) \right)^{Tp}}{T^{Tp} \prod_{t=1}^T \left(q(\hat{\Sigma}_t^{\text{TE}}, \mathbf{G} \mathbf{x}_k^{(t)}) \right)^p},$$

where all terms $\mathbf{G} \hat{\Sigma}_0^{\text{MT}} \mathbf{G}$, $\mathbf{G} \hat{\Sigma}_t^{\text{TE}} \mathbf{G}$, $q(\hat{\Sigma}_0^{\text{MT}}, \mathbf{G} \mathbf{x}_k^{(t)})$, $q(\hat{\Sigma}_t^{\text{TE}}, \mathbf{G} \mathbf{x}_k^{(t)})$ can be written as functions of $\{\mathbf{G} \mathbf{x}_k^{(t)} \mid t \in \llbracket 1, T \rrbracket, k \in \llbracket 1, N \rrbracket\}$.

Finally by taking $\mathbf{G} = \Sigma_0^{-1/2}$, the statistic is a function of $\{\Sigma_0^{-1/2} \mathbf{x}_k^{(t)} \mid t \in \llbracket 1, T \rrbracket, k \in \llbracket 1, N \rrbracket\}$ where $\Sigma_0^{-1/2} \mathbf{x}_k^{(t)} \sim \mathbb{C}\mathcal{N}(\mathbf{0}_p, \mathbf{I}_p)$. It follows that the statistic is independent of Σ_0 that ends the proof.

The same arguments of invariance are used for $\hat{\Lambda}_{\text{MT}}^{\text{marg}}$. \square

APPENDIX F PROOF OF PROPOSITION VI.3

Proof. The texture CFAR property is done using the same procedure as propositions VI.1 and VI.2.

The matrix CFAR property cannot be ensured due to the trace normalisation. For any estimate $\hat{\Sigma} \in \{\hat{\Sigma}_1^{\text{TE}}, \dots, \hat{\Sigma}_T^{\text{TE}}, \hat{\Sigma}_1^{\text{Tex}}, \dots, \hat{\Sigma}_T^{\text{Tex}}\}$, when replacing $\hat{\Sigma}$ by $p\hat{\Sigma}/\text{Tr}(\hat{\Sigma})$ in eq. (20), we have:

$$\hat{\Lambda}_{\text{Tex}} = \prod_{t=1}^T \frac{\left| \hat{\Sigma}_t^{\text{Tex}} \right|}{\left| \hat{\Sigma}_t^{\text{TE}} \right| \text{Tr}(\hat{\Sigma}_t^{\text{Tex}})} \times \prod_{k=1}^N \frac{\left(\sum_{t=1}^T \text{Tr}(\hat{\Sigma}_t^{\text{Tex}}) q(\hat{\Sigma}_t^{\text{Tex}}, \mathbf{x}_k^{(t)}) \right)^{Tp}}{T^{Tp} \prod_{t=1}^T \left(q(\hat{\Sigma}_t^{\text{TE}}, \mathbf{x}_k^{(t)}) \right)^p}.$$

In this expression, the trace terms do not simplify. \square

REFERENCES

- [1] M. Hussain, D. Chen, A. Cheng, H. Wei, and D. Stanley, "Change detection from remotely sensed images: From pixel-based to object-based approaches," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 80, pp. 91 – 106, 2013. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0924271613000804>
- [2] J. Li and E. G. Zelnio, "Target detection with synthetic aperture radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, no. 2, pp. 613–627, April 1996.
- [3] L. M. Novak, "Coherent change detection for multi-polarization SAR," in *Conference Record of the Thirty-Ninth Asilomar Conference on Signals, Systems and Computers*, 2005., Oct 2005, pp. 568–573.
- [4] M. Preiss and N. J. S. Stacy, "Polarimetric SAR coherent change detection," in *7th European Conference on Synthetic Aperture Radar*, June 2008, pp. 1–4.
- [5] J. Barber, "A generalized likelihood ratio test for coherent change detection in polarimetric SAR," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 9, pp. 1873–1877, Sept 2015.
- [6] A. M. Atto, E. Trouve, Y. Berthoumieu, and G. Mercier, "Multidate divergence matrices for the analysis of SAR image time series," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 4, pp. 1922–1938, April 2013.
- [7] C. Lesniewska-Choquet, A. M. Atto, G. Mauris, and G. Mercier, "Image change detection by possibility distribution dissemblance," in *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, July 2017, pp. 1–6.
- [8] K. Conradsen, A. A. 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)*, vol. 6, 2001, pp. 2628–2630 vol.6.
- [9] L. M. Novak, "Change detection for multi-polarization multi-pass SAR," in *Algorithms for Synthetic Aperture Radar Imagery XII*, vol. 5808. International Society for Optics and Photonics, 2005, pp. 234–247.
- [10] V. Carotenuto, A. D. 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.
- [11] A. D. Maio, D. Orlando, L. Pallotta, and C. Clemente, "A multifamily GLRT for oil spill detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 1, pp. 63–79, Jan 2017.
- [12] 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.
- [13] D. Ciuonzo, V. Carotenuto, and A. D. 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.
- [14] H. Nagao, "On some test criteria for covariance matrix," *Ann. Statist.*, vol. 1, no. 4, pp. 700–709, 07 1973. [Online]. Available: <https://doi.org/10.1214/aos/1176342464>
- [15] J. R. Schott, "Some tests for the equality of covariance matrices," *Journal of Statistical Planning and Inference*, vol. 94, no. 1, pp. 25 – 36, 2001. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378375800002093>
- [16] T. Anderson, *An Introduction to Multivariate Statistical Analysis*, ser. Wiley Series in Probability and Statistics. Wiley, 2003. [Online]. Available: <https://books.google.es/books?id=Cmm9QgAACAAJ>
- [17] P. Galeano and D. Pea, "Covariance changes detection in multivariate time series," *Journal of Statistical Planning and Inference*, vol. 137, no. 1, pp. 194 – 211, 2007. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378375805002673>
- [18] A. Aue, S. Hrmann, L. Horvth, and M. Reimherr, "Break detection in the covariance structure of multivariate time series models," *Ann. Statist.*, vol. 37, no. 6B, pp. 4046–4087, 12 2009. [Online]. Available: <https://doi.org/10.1214/09-AOS707>
- [19] 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.
- [20] F. Gini and M. Greco, "Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter," *Signal Process.*, vol. 82, no. 12, pp. 1847–1859, Dec. 2002. [Online]. Available: [http://dx.doi.org/10.1016/S0165-1684\(02\)00315-8](http://dx.doi.org/10.1016/S0165-1684(02)00315-8)
- [21] E. Conte, A. D. Maio, and C. Galdi, "Statistical analysis of real clutter at different range resolutions," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 40, no. 3, pp. 903–918, July 2004.
- [22] 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.
- [23] E. Ollila, D. E. Tyler, V. Koivunen, and H. V. Poor, "Complex elliptically symmetric distributions: Survey, new results and applications," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5597–5625, Nov 2012.

- [24] J.-P. Ovarlez, G. Ginolhac, and A. M. Atto, "Multivariate linear time-frequency modeling and adaptive robust target detection in highly textured monovariate SAR image," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2017, pp. 4029–4033.
- [25] F. Pascal, P. Forster, J.-P. Ovarlez, and P. Larzabal, "Theoretical analysis of an improved covariance matrix estimator in non-gaussian noise," in *Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, vol. 4, March 2005, pp. iv/69–iv/72 Vol. 4.
- [26] A. D. Maio, Ed., *Modern Radar Detection Theory*, ser. Radar, Sonar & Navigation. Institution of Engineering and Technology, 2015. [Online]. Available: <http://digital-library.theiet.org/content/books/ra/sbra509e>
- [27] M. Hallin and D. Paindaveine, "Optimal tests for homogeneity of covariance, scale, and shape," *Journal of Multivariate Analysis*, vol. 100, no. 3, pp. 422 – 444, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0047259X08001474>
- [28] P. Formont, F. Pascal, G. Vasile, J.-P. Ovarlez, and L. Ferro-Famil, "Statistical classification for heterogeneous polarimetric SAR images," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 3, pp. 567–576, June 2011.
- [29] M. Liu, H. Zhang, and C. Wang, "Change detection in urban areas of high-resolution polarization SAR images using heterogeneous clutter models," in *2011 3rd International Asia-Pacific Conference on Synthetic Aperture Radar (APSAR)*, Sept 2011, pp. 1–4.
- [30] A. Mian, J.-P. Ovarlez, G. Ginolhac, and A. M. Atto, "A robust change detector for highly heterogeneous images." 2018 IEEE International Conference on Acoustics, Speech and Signal Processing, 2018, pp. 3429–3433.
- [31] K. Conradsen, A. A. Nielsen, and H. Skriver, "Determining the points of change in time series of polarimetric SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 5, pp. 3007–3024, May 2016.
- [32] G. R. Terrell, "The gradient statistic," *Computing Science and Statistics*, vol. 34, no. 34, pp. 206–215, 2002.
- [33] C. Radhakrishna Rao, "Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 44, no. 1, p. 5057, 1948.
- [34] A. Wald, "Tests of statistical hypotheses concerning several parameters when the number of observations is large," *Transactions of the American Mathematical Society*, vol. 54, no. 3, pp. 426–482, 1943. [Online]. Available: <http://www.jstor.org/stable/1990256>
- [35] D. E. Tyler, "A distribution-free m -estimator of multivariate scatter," *Ann. Statist.*, vol. 15, no. 1, pp. 234–251, 03 1987. [Online]. Available: <https://doi.org/10.1214/aos/1176350263>
- [36] F. Pascal, Y. Chitour, J.-P. Ovarlez, P. Forster, and P. Larzabal, "Covariance structure maximum-likelihood estimates in compound gaussian noise: Existence and algorithm analysis," *IEEE Transactions on Signal Processing*, vol. 56, no. 1, pp. 34–48, Jan 2008.
- [37] I. Soloveychik and A. Wiesel, "Performance analysis of Tyler's covariance estimator," *IEEE Transactions on Signal Processing*, vol. 63, no. 2, pp. 418–426, Jan 2015.
- [38] G. Draković and F. Pascal, "New properties for Tyler's covariance matrix estimator," in *2016 50th Asilomar Conference on Signals, Systems and Computers*, Nov 2016, pp. 820–824.
- [39] F. Pascal, P. Forster, J.-P. Ovarlez, and P. Larzabal, "Performance analysis of covariance matrix estimates in impulsive noise," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2206–2217, June 2008.
- [40] E. Ollila, J. Eriksson, and V. Koivunen, "Complex elliptically symmetric random variables - generation, characterization, and circularity tests," *IEEE Transactions on Signal Processing*, vol. 59, no. 1, pp. 58–69, Jan 2011.
- [41] A. Wiesel, "Geodesic convexity and covariance estimation," *IEEE Transactions on Signal Processing*, vol. 60, no. 12, pp. 6182–6189, Dec 2012.
- [42] J. T. Kent and D. E. Tyler, "Maximum likelihood estimation for the wrapped cauchy distribution," *Journal of Applied Statistics*, vol. 15, no. 2, pp. 247–254, 1988. [Online]. Available: <https://doi.org/10.1080/02664768800000029>
- [43] S. M. Scarborough, L. Gorham, M. J. Minardi, U. K. Majumder, M. G. Judge, L. Moore, L. Novak, S. Jaroszewksi, L. Spoldi, and A. Pieramico, "A challenge problem for SAR change detection and data compression," pp. 7699 – 7699 – 5, 2010. [Online]. Available: <http://dx.doi.org/10.1117/12.855378>
- [44] V. Carotenuto, A. D. Maio, C. Clemente, J. J. Soraghan, and G. Alfano, "Forcing scale invariance in multipolarization SAR change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 1, pp. 36–50, Jan 2016.
- [45] D. Ratha, S. De, T. Celik, and A. Bhattacharya, "Change detection in polarimetric SAR images using a geodesic distance between scattering mechanisms," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 7, pp. 1066–1070, July 2017.
- [46] A. D. C. Nascimento, A. C. Frery, and R. J. Cintra, "Detecting Changes in Fully Polarimetric SAR Imagery with Statistical Information Theory," *ArXiv e-prints*, Jan. 2018.
- [47] A. Mian, J.-P. Ovarlez, G. Ginolhac, and A. Atto, "Multivariate change detection on high resolution monovariate SAR image using linear time-frequency analysis," in *2017 25th European Signal Processing Conference (EUSIPCO)*, Aug 2017, pp. 1942–1946.