
This version is available at https://strathprints.strath.ac.uk/59967/

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (https://strathprints.strath.ac.uk/) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator: strathprints@strath.ac.uk
A Multi-Family GLRT for Detection in Polarimetric SAR Images

L. Pallotta, Member, IEEE, C. Clemente, Member, IEEE, A. De Maio, Fellow, IEEE, and D. Orlando, Senior Member, IEEE

Abstract—This paper deals with detection from multi-polarization SAR images. The problem is cast in terms of a composite hypothesis test aimed at discriminating between the Polarimetric Covariance Matrix (PCM) equality (absence of target in the tested region) and the situation where the region under test exhibits a PCM with at least an ordered eigenvalue smaller than that of a reference covariance. This last setup reflects the physical condition where the backscattering associated with the target leads to a signal, in some eigen-directions, weaker than the one gathered from a reference area where it is a priori known the absence of targets. A Multi-family Generalized Likelihood Ratio Test (MGLRT) approach is pursued to come up with an adaptive detector ensuring the Constant False Alarm Rate (CFAR) property. At the analysis stage, the behaviour of the new architecture is investigated in comparison with a benchmark (but non-implementable) and some other adaptive sub-optimum detectors available in open literature. The study, conducted in the presence of both simulated and real data, confirms the practical effectiveness of the new approach.

Index Terms—CFAR, Covariance Matrix Equality, MGLRT.

I. INTRODUCTION

Polarimetric SAR images provide enhanced information on the imaged scene that can be exploited for improved target detection, recognition and scene classification [1]. Following the imaging stage, target detection can be applied and improved performance are achievable exploiting the multi-polarimetric nature of the data. Detectors exploiting the polarimetric information have been developed for specific applications including change [2], [3], [4], oil spill [5], [6] and ship detection [7].

In this paper, the problem of target detection is formulated in terms of a binary hypothesis test aimed at discriminating between the presence and the absence of variations in the Polarimetric Covariance Matrix (PCM) of the radar returns. The presence of targets such as oil spills and ship wakes modifies the backscattering of sea surface. The idea is to compare the region under test, which possibly contains targets, to a reference area where only echoes from the sea are present. Without loss of generality, in this paper we will focus on the specific case of oil spills, where it is reasonable to assume that the PCM of data containing oil slicks share eigenvalues smaller than or equal to the PCM of the sea returns. The decision problem is solved applying the GLRT and the devised architecture is referred to as Positive Definite Difference GLR T (PDD-GLRT). At the design stage, it is assumed that the rank difference between the two covariance matrices is known. However, this assumption might not be met in practical scenarios, since such a-priori information is not available at the receiver. In order to circumvent this drawback, the previous results are extended to come up with a decision rule capable of properly estimating the rank difference. This goal is achieved exploiting the Multi-family GLRT (MGLRT) [8] and the devised decision rule is referred to as Multi-family PDD-GLRT (MPDD-GLRT). Finally, numerical examples are provided to prove the effectiveness of the proposed approach also in comparison with existing strategies for oil spills detection.

The remainder of this paper is organized as follows. Section II is devoted to the problem formulation. The derivations of the PDD-GLRT and its multi-family modification are reported in Section III. Section IV analyzes the performance in terms of detection probability on simulated and real data. Some concluding remarks and future research tracks are given in Section V.

NOTATION

Vectors and matrices are denoted by boldface lower-case and upper-case letters, respectively. Symbols det(·) and tr(·) denote the determinant and the trace of a square matrix, respectively. Symbol $\mathbb{H}^N$ is used to represent the set of $N \times N$ hermitian matrices, while $I$ and $0$ represent the identity matrix and the null vector or matrix, respectively, both of proper dimensions. The curled inequality symbol $\succ$ is used to denote generalized matrix inequality: for any $A \in \mathbb{H}^N$, $A \succeq 0$ means that $A$ is a positive semi-definite matrix. Symbols $(\cdot)^T$ and $(\cdot)^\dagger$ stand for transpose and conjugate transpose, respectively. Finally, the acronym iid means independent and identically distributed while the symbol $E[\cdot]$ denotes statistical expectation.

II. PROBLEM FORMULATION

A multi-polarization SAR sensor, for each pixel of the image under test, measures $N = 3$ complex returns, which are collected from three different polarimetric channels (namely HH, HV, and VV). The $N$ returns associated with the same
pixel are organized in the specific order HH, HV, and VV to form the vector \( \mathbf{X}(l_1, l_2) \), where \( l_1 = 1, \ldots, L_1 \) and \( l_2 = 1, \ldots, L_2 \) (\( L_1 \) and \( L_2 \) represent the vertical and horizontal size of the image, respectively). Therefore, the sensor provides a 3-D data stack \( \mathbf{X} \) of size \( L_1 \times L_2 \times N \) which is referred to as datacube, whose pictorial representation is given in Fig. 1.

Starting from the datacube \( \mathbf{X} \) of the scene illuminated by the radar, for each pixel under test, we extract a rectangular neighborhood \( \mathcal{A} \) of size \( K = W_1 \times W_2 \geq N \). We denote by \( \mathbf{Z} = [z_1 \ldots z_K] \in \mathbb{C}^{N \times K} \) the matrix whose columns are the vectors of the polarimetric returns from the pixels of \( \mathbf{X} \) which fall in the region \( \mathcal{A} \). The matrix \( \mathbf{Z} \) is modeled as a random matrix, whose columns are assumed iid random vectors drawn from a complex circular zero-mean Gaussian distribution with positive definite covariance matrix \( \mathbf{R} \).

Our goal is to identify those regions of \( \mathbf{X} \) that exhibit variations of the covariance matrix \( \mathbf{R} \) with respect to a preassigned reference region. Specifically, the presence of specific objects in the observed scene yields signal echoes with spectral properties different from those associated with the background [5], [6]. For instance, consider an oil spill on the sea surface, whose reflectivity coefficient reduces the variations of the covariance matrix \( \mathbf{R} \) and

\[
\begin{align*}
\Lambda_{p}(\mathbf{Z}_K, \mathbf{Y}_M) = \begin{cases} 
1, & \text{if } p^* < p, \\
\prod_{i=1}^{p^*} (1 + \delta_i)^{K+M} / \delta_i^M, & \text{otherwise}, 
\end{cases}
\end{align*}
\]

where

\[
\delta_i, i = 1, \ldots, p^*, \text{ are } p^* \text{ eigenvalues of } \mathbf{G}^{-1} \mathbf{H} \text{ with } p^* \text{ being the minimum between the number of eigenvalues of } \mathbf{G}^{-1} \mathbf{H} \text{ greater than } M/K \text{ and } p.
\]

The above architecture will be referred to in the following as PDD-GLRT. In the next subsection, we apply the Exponential Embedded Family (EEF) framework [8], [9] to devise an architecture based upon the PDD-GLRT that does not use any a priori information on the rank of \( \mathbf{R}_2 \).

\[1\] For the sake of simplicity, in the following, we neglect the possible low correlation level between adjacent pixels.

\[2\] The proof is omitted for the sake of brevity.
A. Multi-Family PDD-GLRT

In the present subsection, we introduce a variant to the PDD-GLRT which does not need the above a-priori information. Specifically, the binary hypotheses considered before become multiple nested instances each tied up to a rank value of $R_k$, that is denoted by $r(i) = i, i = 1, \ldots, N$. In this scenario, the classical GLRT cannot be used and, hence, we resort to the MGLRT [8] [9] that allows the PDD-GLRT to be utilized also in the case where the rank is not known. Specifically, the MGLRT can be written in terms of the EEF computed for a given model order $i, i = 1, \ldots, N$, [9] namely

$$\max_{i \in \{1, \ldots, N\}} \text{EEF}(i) = \frac{H_1}{H_0} > \eta,$$  \hspace{1cm} (8)

where

$$\text{EEF}(i) = \left( \zeta_{G_i}(Z_K, Y_M) \right) - \left( \frac{r(i)}{\log \left( \frac{\zeta_{G_1}(Z_K, Y_M)}{r(i)} \right) + 1} \right) \times$$  \hspace{1cm} (9)

$$\times \left( \frac{\zeta_{G_i}(Z_K, Y_M)}{r(i) - 1} \right), \quad i = 1, \ldots, N.$$

In the last equation, $\zeta_{G_i}(Z_K, Y_M)$ is two times the logarithm of the PDD-GLRT decision statistic with $p = i$, namely

$$\zeta_{G_i}(Z_K, Y_M) = 2 \log \Lambda_i(Z_K, Y_M),$$  \hspace{1cm} (10)

where $u(\cdot)$ is the Heaviside step function. Thus, when there are multiple nested alternative hypotheses, the MGLRT first computes the PDD-GLRT for each hypothesis, $\zeta_{G_i}(\cdot)$ say, applies the transformation (9) to construct the EEF(i), and chooses the hypothesis attaining the maximum. For the sake of completeness, we provide below the explicit expression of $\zeta_{G_i}(Z_K, Y_M)$, which contains the constants which have been incorporated into the threshold of the PDD-GLRT, i.e.,

$$\zeta_{G_i}(Z_K, Y_M) = -2(K + M)r(i) \log(K + M) + 2Kr(i) \log(K) + 2Mr(i) \log(M)$$

$$- 2M \sum_{n=1}^{r(i)} \log \delta_n + 2(K + M) \sum_{n=1}^{r(i)} \log(1 + \delta_n).$$  \hspace{1cm} (11)

IV. PERFORMANCE ANALYSES

In this section, we investigate the performances of the proposed detectors in terms of Probability of Detection ($P_D$). For comparison purposes, we also plot the performance curves of other approaches available in the open literature. More precisely, we consider the GLRT devised in [3], [4], [6], whose expression is

$$\Lambda_{\text{GLRT}} = \frac{[\det(G + H)]^{(K+M)}}{[\det(G)]^{K}[\det(H)]^{M}} \frac{H_1}{H_0} > \eta,$$  \hspace{1cm} (12)

1The model order is represented by the rank of $R_2$.

the Maximum Likelihood Detector (MLD) proposed in [6] given by

$$\Lambda_{\text{MLD}} = \frac{\det(H)}{\det(G)} \frac{H_1}{H_0} \frac{H_1}{H_0} > \eta,$$  \hspace{1cm} (13)

and the Single Likelihood Detector (SLD) proposed in [6] whose decision statistic is

$$\Lambda_{\text{SLD}} = \text{tr} \left( G^{-1} H \right).$$  \hspace{1cm} (14)

In order to better analyze the behavior of all these detectors, the comparisons also include the clairvoyant counterparts, namely the Likelihood Ratio Test (LRT) and the clairvoyant SLD (C-SLD). More precisely, the former is statistically equivalent to

$$\Lambda_{\text{LRT}} = \text{tr} \left[ R_1^{-1} (G + H) \right]$$

$$- R_1^{-1} G - (R_1 + R_2)^{-1} H \frac{H_1}{H_0} \frac{H_1}{H_0} > \eta,$$  \hspace{1cm} (15)

while the latter has the following expression

$$\Lambda_{\text{C-SLD}} = \text{tr} \left( R_1^{-1} H \right).$$  \hspace{1cm} (16)

The analysis is conducted on both simulated and real SAR data. In the latter case, we prove the effectiveness of the newly proposed detectors in terms of the capability to detect oil spills over the sea surface.

A. Detection Capability on Simulated Data

This subsection is devoted to the performance analysis using computer simulated data for a preassigned value of Probability of False Alarm ($P_{FA}$). Specifically, the numerical examples are obtained by means of standard Monte Carlo counting techniques. The detection thresholds and the $P_D$ values are evaluated resorting to $10^3/P_{FA}$ and $10^3$ independent trials, respectively. The nominal $P_{FA}$ is set to $10^{-4}$. As to the vector size, it is chosen equal to 3 to account for the fact that SAR sensors collect data using three different polarizations, whereas as to the parameter $p$ the value $p = 2$ is considered.

The simulated data, $z_k, k = 1, \ldots, K,$ and $y_m, m = 1, \ldots, M,$ are modeled as $N$-dimensional zero-mean complex circular Gaussian vectors with covariance matrices $R_1 = I$ and $R_1 + R_2$, respectively. Matrix $R_2$ is rank deficient and is defined as $R_2 = [\alpha]^2 p_1 p_1^* + [\beta]^2 p_2 p_2^*$, where the $N$-dimensional steering vectors, $p_1$ and $p_2$, have been chosen as

$$p_1 = [1, 0, \ldots, 0]^T,$$

$$p_2 = [0, 1, 0, \ldots, 0]^T,$$

and $|\alpha|^2 = |\beta|^2$, so as $R_2$ is a rank-2 matrix.

The SNR is defined as

$$\text{SNR} = |\alpha|^2 |p_1|^2 R_1^{-1} p_1 + |\beta|^2 |p_2|^2 R_1^{-1} p_2 = 2|\alpha|^2.$$  \hspace{1cm} (18)

In Fig. 2, we plot $P_D$ versus SNR assuming $K = M = 9$. It turns out that the PDD-GLRT and the MLD perform better than all the other adaptive architectures. In addition, the M-PDD-GLRT exhibits some losses with respect to the PDD-
GLRT even though both detectors outperform the GLRT and the SLD with a gain of about 1 dB at $P_D = 0.9$. As last remark, note that the PDD-GLRT suffers a detection loss of about 5 dB with respect to the LRT (or the C-SLD).

In Fig. 3, we focus on a smaller number of test data, i.e., $M = 4$, leaving the same value of $K$ as in Fig. 2. From the figure, it can be observed that there is a general worsening in the receivers’ performance. The figure highlights also that the PDD-GLRT and the M-PDD-GLRT achieve the best performances with the former performing better than the latter. Finally, the GLRT is the worst with a loss higher than 3 dB if compared to the PDD-GLRT and its multi-family version.

Summarizing, the PDD-GLRT seems a reasonable choice since it exhibits a more robust behavior than the counterparts being the one less sensitive to parameters variations and giving in all the analyzed situations acceptable performances. In other words, the PDD-GLRT represents a kind of “all seasons” architecture.

B. Test on Real Radar Data

As final part of the performance analysis, in this subsection, we evaluate the effectiveness of the new detectors on real SAR data. More precisely, we use a data set obtained from the GOMoil_07601_10052_101_100622_L090_CX_02 SAR image which is composed of sea data containing an oil spill on it. The image has been acquired on 22nd of June 2010, during the British Petroleum oil spill incident in the Gulf of Mexico (known also as the Deepwater Horizon Oil Spill). This image contains a scene acquired with a sensor exploiting all the polarizations and the corresponding polarimetric overlay is reported in Fig. 4.

The data set is a sub-image of $3000 \times 2000$ pixels, whose span (i.e., $|HH| + |VV| + 2|HV|$) [1] expressed in dB values is also displayed in Fig. 5.

Fig. 4. Three polarization color overlay of the SAR image GOMoil_07601_10052_101_100622_L090_CX_02.

Fig. 5. Real L-band data SAR image GOMoil_07601_10052_101_100622_L090_CX_02. Span (in dB) of the reference image of size $3000 \times 2000$ pixels.

Fig. 6 shows the detection results applying detector (8) over the described SAR image, where the test is performed assuming the following parameter setting (notice that $N = 3$...
since we have three polarization images). A sea data pixel is chosen as reference to compute the sample covariance $H$ from a $3 \times 3$ (i.e., $M = 9$) window centered in that pixel. Then, a window of size $3 \times 3$ (i.e., $K = 9$) is slid over the SAR image to test all the pixels and to compute the sample matrix $G$. Finally, the threshold is chosen to ensure a nominal $P_{FA}$ of $10^{-3}$, extracting a cluster of only sea data composed of $10^5$ pixels and choosing the 100-th value in the decreasing ordered statistics. The results is a detection map where the white pixels are those associated to oil spills detections, while the black pixels are representative of the sea data (no-detections). The results show that the multi-family PDD is able to ensure reliable detection of oil spills over sea. An important notice is that the number of looks, in a real environment, has to be chosen to manage the trade off between the accuracy of covariance estimation and the amount of heterogeneity.

V. CONCLUSIONS
Multi-polarization SAR detection has been considered in this paper to test the equality of two polarimetric sample covariance matrices constructed from a reference area (where it is known the absence of targets) and a test region (where it is necessary to establish the spot presence), respectively. The alternative hypothesis is represented by the instance where the PCM within the reference area exhibits at least an ordered eigenvalue greater than the corresponding one extracted from the PCM of the area under test. First of all, assuming the exact knowledge about the number of different eigenvalues between the reference and tested PCMs, the PDD-GLRT is devised. Then, to come up with a fully adaptive detector without any a priori assumption about the aforementioned number, the M-PDD-GLRT is introduced.

At the analysis stage, the oil spill detection case has been analysed and examples have been provided to prove the effectiveness of the proposed approach also in comparison with existing strategies. To this end, both simulated and real multi-polarization SAR data have been used. Possible future research track might concern the extension of the approach to the case of a joint multi-frequency and multi-polarization processing as well as to consider the presence of a non-Gaussian backscattering due to sea.

ACKNOWLEDGMENT
The authors would like to acknowledge NASA/JPL for providing the SAR data. The work of C. Clemente was sponsored by the Engineering and Physical Sciences Research Council (EPSRC) Grant number EP/K014307/1 and the MOD University Defence Research Collaboration in Signal Processing.

REFERENCES

![Image](image_url)