Invariant Rules for Multipolarization SAR Change Detection

Vincenzo Carotenuto, *Student Member, IEEE*, Antonio De Maio, *Fellow, IEEE*, Carmine Clemente, *Member, IEEE*, and John J. Soraghan, *Senior Member, IEEE*

***Abstract*—This paper deals with coherent (in the sense that both amplitudes and relative phases of the polarimetric returns are used to construct the decision statistic) multipolarization synthetic aperture radar (SAR) change detection assuming the availability of reference and test images collected from N multiple polarimetric channels. At the design stage, the change detection problem is formulated as a binary hypothesis testing problem, and the principle of invariance is used to come up with decision rules sharing the constant false alarm rate property. The maximal invariant statistic and the maximal invariant in the parameter space are obtained. Hence, the optimum invariant test is devised proving that a uniformly most powerful invariant detector does not exist. Based on this, the class of suboptimum invariant receivers, which also includes the generalized likelihood ratio test, is considered. At the analysis stage, the performance of some tests, belonging to the aforementioned class, is assessed and compared with the optimum clairvoyant invariant detector. Finally, detection maps on real high-resolution SAR data are computed showing the effectiveness of the considered invariant decision structures.**

***Index Terms*—Coherent change detection, invariant rules, maximal invariant, multipolarization.**

I. Introduction

A TECHNICAL challenge in synthetic aperture radar (SAR) signal processing is change detection, namely, the capability to identify temporal changes within a given scene [1]–[3] starting from a pair of coregistered images representing the area of interest, which is usually referred to as the reference and test pair. Two main approaches, known as incoherent and coherent, have been proposed in the open literature to process the image pair. The former attempts to detect changes in the mean power level of a given scene exploiting only the intensity information from the available images (thus neglecting phase information [4]–[6]): Differencing and ratioing are well-known techniques in this context [7]. The latter jointly use both amplitude and phase from the reference and test data to detect possible changes in the region of interest. In [4] and [5], a thorough comparison between incoherent and coherent change detection strategies, including the maximum-likelihood estimate of the SAR coherence parameter, is performed based on high-resolution [8] (0.3 m $\times$ 0.3 m) SAR images (the pixel sample spacing is 0.2 m $\times$ 0.2 m). In [7], several techniques for change detection have been presented and compared based on their probability of error and on results obtained using repeat-pass ERS-1 SAR data.

In [6], [9], and [10], the multipolarization signal model for the SAR change detection problem is laid down, the detection problem is formulated as a binary hypothesis test, and a decision rule based on the generalized likelihood ratio test (GLRT) is developed. Moreover, a performance analysis [10] of the GLRT is given in the form of receiver operating characteristics (ROCs), namely, detection probability (P_d) versus false alarm probability (P_{fa}), quantifying the benefits of multipolarization information in SAR change detection. Finally, a detection scheme based on canonical correlations analysis is applied to scalar EMISAR data [11], [12], whereas in [13], a mutual-information-based framework is developed to address coherent similarities between multichannel SAR images.

Starting from the multipolarization data model developed in [9] and [10], we propose a new and systematic framework for change detection based on the theory of invariance in hypothesis testing problems [14], [15]. This is a viable means to force some desired properties to a decision statistic at the design stage and has already been successfully applied in some different radar detection problems [16]–[18]. Otherwise stated, the principle of invariance allows for focus on decision rules that exhibit some natural symmetries implying important practical properties such as the constant false alarm rate (CFAR) behavior. Moreover, the use of invariance leads to a data reduction because all invariant tests can be expressed in terms of a statistic, called maximal invariant, which organizes the original data into equivalence classes. Furthermore, the parameter space is usually compressed after reduction by invariance, and the dependence on the original set of parameters becomes embodied into a maximal invariant in the parameter space (induced maximal invariant). With reference to the multipolarization SAR change detection problem, we first determine a maximal invariant statistic in terms of the eigenvalues of a data-dependent matrix constructed using both test and reference images. Then, we design the most powerful invariant (MPI)

Manuscript received March 12, 2014; revised June 23, 2014 and October 1, 2014; accepted November 12, 2014. This work was supported in part by the Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/K014307/1 and in part by the MOD University Defence Research Collaboration in Signal and the University of Naples "Federico II."

V. Carotenuto and A. De Maio are with the Dipartimento di Ingegneria Elettrica e delle Tecnologie dell'Informazione, Università degli Studi di Napoli "Federico II," 80125 Napoli, Italy (e-mail: vincenzo.carotenuto@unina.it; ademaio@unina.it).

C. Clemente and J. J. Soraghan are with the Centre for Excellence in Signal and Image Processing, University of Strathclyde, Glasgow G1 1XW, U.K. (e-mail: carmine.clemente@strath.ac.uk; j.soraghan@strath.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2014.2372900

receiver, as the Neyman–Pearson (NP) detector computed from the maximal invariant statistic, and show that no uniformly MPI (UMPI) test exists, namely, the optimum invariant detector is not implementable. Hence, we focus on suboptimum invariant receivers based on different functions of the maximal invariant such as arithmetic mean, harmonic mean, and maximum and minimum entries. Remarkably, the GLRT in [9] and [10] belongs to the proposed class of receivers. At the analysis stage, we assess the performance of the considered invariant decision rules in correspondence of a two-channel polarization diversity providing detection probability contours, for a given false alarm level, versus the induced maximal invariant, which turns out to be bidimensional. We also conduct performance comparisons with the benchmark MPI test showing that some analyzed receivers can outperform the GLRT and achieve a detection probability close to that ensured by the clairvoyant MPI structure. Moreover, we consider ROC curves and assess the impact of the polarimetric information on the detection performance.

Finally, to validate the behavior of the considered tests on real SAR images, we use a high-resolution change detection data set, available from the Air Force Research Laboratory (AFRL) [8] and collected from an X-band SAR. The conducted analysis highlights the detectors' capability to successfully operate in real environments; some of these new detectors have the interesting feature of being able to directly discriminate between car arrivals and departures without any additional radar cross section (RCS) comparison.

This paper is organized as follows. In Section II, we deal with the formulation of the multipolarization SAR change detection problem, whereas in Section III, we address maximal invariant and induced maximal invariant design. In Section IV, we devise the MPI test and show that a UMPI detector does not exist for the present problem. Thus, we consider the design of suboptimum invariant decision rules. In Sections V and VI, we assess the performance of the introduced invariant tests, respectively, on simulated and on real multipolarization SAR images. Finally, in Section VII, we draw conclusions and outline some possible future research tracks.

A. Notation

We adopt the notation of using boldface for vectors and matrices. The transpose and conjugate transpose operators are denoted, respectively, by the symbols $(\cdot)^T$ and $(\cdot)^\dagger$. $\mathrm{tr}(\cdot)$ and $\det(\cdot)$ are, respectively, the trace and the determinant of the square matrix argument. $\boldsymbol{I}$ and $\boldsymbol{0}$ denote, respectively, the identity matrix and the matrix with zero entries (their size is determined from the context). $\mathbf{Diag}(\boldsymbol{a})$ indicates the diagonal matrix whose ith diagonal element is the ith entry of $\boldsymbol{a}$. The curled inequality symbol $\succ$ is used to denote the generalized matrix inequality: For any Hermitian matrix $\boldsymbol{A}$, $\boldsymbol{A} \succ \boldsymbol{0}$ means that $\boldsymbol{A}$ is a positive definite matrix. The general linear group of degree N over the field of complex numbers, denoted by $\mathcal{GL}(N)$, is the set of $N \times N$ nonsingular matrices together with the operation of ordinary matrix multiplication. Finally, $\mathcal{H}_N^{++}$ denotes the set of $N \times N$ Hermitian positive definite matrices.

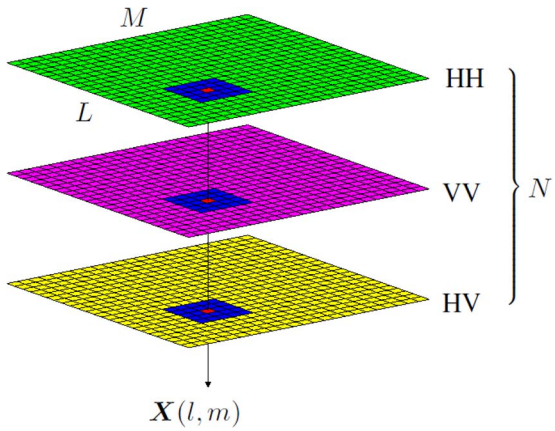


Fig. 1. Construction of the datacube.

II. Problem Formulation

A multipolarization SAR sensor measures for each pixel of the image under test $N \in \{2, 3\}$ complex returns, collected from different polarimetric channels (for instance, HH and VV for $N = 2$; HH, VV, and HV with reference to $N = 3$). The N returns from the same pixel are stacked to form the vector $\boldsymbol{X}(l, m)$, where $l = 1, \ldots, L$ and $m = 1, \ldots, M$ (L and M represent the vertical and horizontal size of the image, respectively). Therefore, the sensor provides a 3-D data stack $\boldsymbol{X}$ of size $M \times L \times N$, which is referred to in the following as a datacube and is illustrated in Fig. 1.

For SAR change detection applications, we suppose that two datacubes $\boldsymbol{X}$ (reference data) and $\boldsymbol{Y}$ (test data) of the same geographic area are available; they are collected from two different sensor passes and are accurately pixel aligned (coregistrated). We focus on the problem of detecting the presence of possible changes in a rectangular neighborhood $\mathcal{A}$, with size $K = W_1 \times W_2 \geq N$, of a given pixel. To this end, we denote by $\boldsymbol{R}_X$ ($\boldsymbol{R}_Y$) the matrix whose columns are the vectors of the polarimetric returns from the pixels of $\boldsymbol{X}$ ($\boldsymbol{Y}$), which fall in the region $\mathcal{A}$ and $\boldsymbol{S}_X = \boldsymbol{R}_X \boldsymbol{R}_X^\dagger$ ($\boldsymbol{S}_Y = \boldsymbol{R}_Y \boldsymbol{R}_Y^\dagger$).

The matrices $\boldsymbol{R}_X$ and $\boldsymbol{R}_Y$ are modeled as statistically independent random matrices. Moreover, the columns of $\boldsymbol{R}_X$ ($\boldsymbol{R}_Y$) are assumed statistically independent and identically distributed random vectors drawn from a complex circular zero-mean Gaussian distribution with a positive definite covariance matrix $\boldsymbol{\Sigma}_X$ ($\boldsymbol{\Sigma}_Y$). Under the aforementioned settings, the change detection problem in the region $\mathcal{A}$ can be formulated in terms of the following binary hypothesis test:

$$\begin{cases} H_0 : \boldsymbol{\Sigma}_X = \boldsymbol{\Sigma}_Y \\ H_1 : \boldsymbol{\Sigma}_X \neq \boldsymbol{\Sigma}_Y \end{cases} \tag{1}$$

where the null hypothesis H_0 of change absence is tested versus the alternative[1] H_1.

[1]This testing problem is also well known in statistical literature with reference to real observations [14, Ch. 8].

Exploiting the Gaussian assumption, we can write the joint probability density function (pdf) of $\boldsymbol{R}_X$ and $\boldsymbol{R}_Y$ as

$$f_{\boldsymbol{R}_X,\boldsymbol{R}_Y}(\boldsymbol{R}_X,\boldsymbol{R}_Y|H_1,\boldsymbol{\Sigma}_X,\boldsymbol{\Sigma}_Y) = \frac{1}{\pi^{2NK}\det(\boldsymbol{\Sigma}_X\boldsymbol{\Sigma}_Y)^K}\exp\left\{-\mathrm{tr}\left(\boldsymbol{\Sigma}_X^{-1}\boldsymbol{S}_X+\boldsymbol{\Sigma}_Y^{-1}\boldsymbol{S}_Y\right)\right\}. \tag{2}$$

Using the Fisher–Neyman factorization theorem [19], we can claim that a sufficient statistic for (1) is represented by the two sample Grammian matrices $\boldsymbol{S}_X$ and $\boldsymbol{S}_Y$, which are statistically independent and follow a complex Wishart distribution, i.e., [20]

$$f_{\boldsymbol{S}_X}(\boldsymbol{S}_X|H_1,\boldsymbol{\Sigma}_X) = \frac{c_W}{\det(\boldsymbol{\Sigma}_X)^K}\exp\left\{-\mathrm{tr}\left(\boldsymbol{\Sigma}_X^{-1}\boldsymbol{S}_X\right)\right\}\det(\boldsymbol{S}_X)^{K-N}, \quad \boldsymbol{S}_X\succ\boldsymbol{0} \tag{3}$$

$$f_{\boldsymbol{S}_Y}(\boldsymbol{S}_Y|H_1,\boldsymbol{\Sigma}_Y) = \frac{c_W}{\det(\boldsymbol{\Sigma}_Y)^K}\exp\left\{-\mathrm{tr}\left(\boldsymbol{\Sigma}_Y^{-1}\boldsymbol{S}_Y\right)\right\}\det(\boldsymbol{S}_Y)^{K-N}, \quad \boldsymbol{S}_Y\succ\boldsymbol{0} \tag{4}$$

with c_W as a normalization constant. From the sufficient statistic, we can evaluate the optimum NP detector as the likelihood ratio test (LRT), i.e.,

$$\frac{f_{\boldsymbol{S}_X,\boldsymbol{S}_Y}(\boldsymbol{S}_X,\boldsymbol{S}_Y|H_1,\boldsymbol{\Sigma}_X,\boldsymbol{\Sigma}_Y)}{f_{\boldsymbol{S}_X,\boldsymbol{S}_Y}(\boldsymbol{S}_X,\boldsymbol{S}_Y|H_0,\boldsymbol{\Sigma}_X)} = \frac{f_{\boldsymbol{S}_X}(\boldsymbol{S}_X|H_1,\boldsymbol{\Sigma}_X)f_{\boldsymbol{S}_Y}(\boldsymbol{S}_Y|H_1,\boldsymbol{\Sigma}_Y)}{f_{\boldsymbol{S}_X}(\boldsymbol{S}_X|H_0,\boldsymbol{\Sigma}_X)f_{\boldsymbol{S}_Y}(\boldsymbol{S}_Y|H_0,\boldsymbol{\Sigma}_X)} \underset{H_0}{\overset{H_1}{\gtrless}} T_{0,A}$$

where the first equality stems from the statistical independence between $\boldsymbol{S}_X$ and $\boldsymbol{S}_Y$; $T_{0,A}$ is the detecton threshold set to ensure a given P_{fa} value. Substituting the pdfs of $\boldsymbol{S}_X$ and $\boldsymbol{S}_Y$ in the LRT, after standard algebra and statistical equivalences (also absorbing irrelevant data-independent terms in the threshold), it can be recast as

$$\mathrm{tr}\left[\left(\boldsymbol{\Sigma}_X^{-1}-\boldsymbol{\Sigma}_Y^{-1}\right)\boldsymbol{S}_Y\right] \underset{H_0}{\overset{H_1}{\gtrless}} T_0 \tag{5}$$

where T_0 is the suitable modification of the original detection threshold. Evidently, test (5) is not uniformly most powerful, and consequently, it is not practically implementable because it requires knowledge (clairvoyant detector) of both $\boldsymbol{\Sigma}_X$ and $\boldsymbol{\Sigma}_Y$, which, in realistic applications, are usually unknown.

III. Data Reduction and Invariance Issues

Both hypotheses under test are composite or, in other words, H_0 and H_1 are equivalent to a partition of the parameter space Θ into the two disjoint sets

$$\Theta_0 = \left\{\boldsymbol{\Sigma}_X=\boldsymbol{\Sigma}_Y,\ (\boldsymbol{\Sigma}_X,\boldsymbol{\Sigma}_Y)\in\mathcal{H}_N^{++}\times\mathcal{H}_N^{++}\right\}$$
$$\Theta_1 = \left\{\boldsymbol{\Sigma}_X\neq\boldsymbol{\Sigma}_Y,\ (\boldsymbol{\Sigma}_X,\boldsymbol{\Sigma}_Y)\in\mathcal{H}_N^{++}\times\mathcal{H}_N^{++}\right\}. \tag{6}$$

This formulation emphasizes that the individual values of the nuisance parameters are irrelevant: One must only discern to which hypothesis they pertain, namely, whether the covariances are equal or not. This suggests that we can cluster the data considering transformations that leave the following unaltered:

a) the two composite hypotheses, namely, the partition of the parameter space;
b) the families of distributions under the two hypotheses.

This goal can be achieved through the *principle of invariance* [15]. According to such principle, we look for transformations that preserve the formal structure of the hypothesis testing problem and then for decision rules invariant to them. Such a principle also acts as a data reduction technique leading to a reduced observation space of significantly lower dimensionality than the original one.

It is not difficult to prove that our testing problem is invariant under the group of transformations G acting on the sufficient statistic as

$$G=\left\{g:\boldsymbol{S}_X\to\boldsymbol{B}\boldsymbol{S}_X\boldsymbol{B}^\dagger,\ \boldsymbol{S}_Y\to\boldsymbol{B}\boldsymbol{S}_Y\boldsymbol{B}^\dagger,\ B\in\mathcal{GL}(N)\right\}. \tag{7}$$

In fact, the families of distributions are preserved because if $\boldsymbol{S}_X$ ($\boldsymbol{S}_Y$) is Wishart distributed, $\boldsymbol{B}\boldsymbol{S}_X\boldsymbol{B}^\dagger$ ($\boldsymbol{B}\boldsymbol{S}_Y\boldsymbol{B}^\dagger$) is also Wishart with the same scalar parameters and matrix parameter $\boldsymbol{B}\boldsymbol{\Sigma}_X\boldsymbol{B}^\dagger$ ($\boldsymbol{B}\boldsymbol{\Sigma}_Y\boldsymbol{B}^\dagger$), where $\boldsymbol{B}\in\mathcal{GL}(N)$. Moreover, the original partition of the parameter space is left unaltered since if $\boldsymbol{\Sigma}_X\neq\boldsymbol{\Sigma}_Y$, then $\boldsymbol{B}\boldsymbol{\Sigma}_X\boldsymbol{B}^\dagger\neq\boldsymbol{B}\boldsymbol{\Sigma}_Y\boldsymbol{B}^\dagger$, and if $\boldsymbol{\Sigma}_X=\boldsymbol{\Sigma}_Y$, then $\boldsymbol{B}\boldsymbol{\Sigma}_X\boldsymbol{B}^\dagger=\boldsymbol{B}\boldsymbol{\Sigma}_Y\boldsymbol{B}^\dagger$.

A. Maximal Invariant Design

The invariance property induces a partition of the data space into orbits (or equivalence classes) where, over each orbit, every point is related to every other through a transformation that is a member of the group G. Any statistic that identifies different orbits in a one-to-one way significantly reduces the total amount of data necessary for solving the hypothesis testing problem and constitutes the compressed data set to be used in the design of any invariant detector. This kind of statistics is called maximal invariants since they are constant over each orbit (invariance) while assuming different values on different orbits (maximality).

Formally, a statistic $\mathbf{T}(\boldsymbol{S}_X,\boldsymbol{S}_Y)$ is said to be a maximal invariant with respect to the group of transformations G if and only if

- Invariance: $\mathbf{T}(\boldsymbol{S}_X,\boldsymbol{S}_Y)=\mathbf{T}[g(\boldsymbol{S}_X,\boldsymbol{S}_Y)],\forall\, g\in G$.
- Maximality: $\mathbf{T}(\boldsymbol{S}_{X_1},\boldsymbol{S}_{Y_1})=\mathbf{T}(\boldsymbol{S}_{X_2},\boldsymbol{S}_{Y_2})$ implies that $\exists\, g\in G$ such that $(\boldsymbol{S}_{X_2},\boldsymbol{S}_{Y_2})=g(\boldsymbol{S}_{X_1},\boldsymbol{S}_{Y_1})$.

Notice that there are many maximal invariant statistics, but they are equivalent in that they yield statistically equivalent detectors. Moreover, all invariant tests can be expressed as a function of the maximal invariant statistic [14], [21], which, for the problem of interest, is provided by the following.

Proposition 1: A maximal invariant statistic for problem (1) with respect to the group of transformations (7) is the N-dimensional vector of the eigenvalues $\lambda_1, \ldots, \lambda_N$ of

$$\boldsymbol{S}_X \boldsymbol{S}_Y^{-1}. \tag{8}$$

Proof: The invariance of $\lambda_1, \ldots, \lambda_N$ follows because the eigenvalues of $\boldsymbol{B}\boldsymbol{S}_X\boldsymbol{B}^\dagger(\boldsymbol{B}\boldsymbol{S}_Y\boldsymbol{B}^\dagger)^{-1} = \boldsymbol{B}\boldsymbol{S}_X\boldsymbol{S}_Y^{-1}\boldsymbol{B}^{-1}$ are the same as those of $\boldsymbol{S}_X\boldsymbol{S}_Y^{-1}$. As to the maximality, we have to prove that if $\boldsymbol{S}_{X,1}\boldsymbol{S}_{Y,1}^{-1}$ and $\boldsymbol{S}_{X,2}\boldsymbol{S}_{Y,2}^{-1}$ share the same eigenvalues, then there exists a nonsingular matrix $\boldsymbol{B}$ such that $\boldsymbol{S}_{X,1} = \boldsymbol{B}\boldsymbol{S}_{X,2}\boldsymbol{B}^\dagger$ and $\boldsymbol{S}_{Y,1} = \boldsymbol{B}\boldsymbol{S}_{Y,2}\boldsymbol{B}^\dagger$. To this end, denoting by $\bar{\boldsymbol{\Lambda}} = \mathbf{Diag}([\lambda_1, \ldots, \lambda_N]^T)$ and exploiting Theorem [22, Corollary 4.6.12, p. 250], we can claim that there exist two nonsingular matrices $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ such that

$$\boldsymbol{B}_1\boldsymbol{S}_{X,1}\boldsymbol{B}_1^\dagger = \bar{\boldsymbol{\Lambda}}\,, \qquad \boldsymbol{B}_1\boldsymbol{S}_{Y,1}\boldsymbol{B}_1^\dagger = \boldsymbol{I}$$
$$\boldsymbol{B}_2\boldsymbol{S}_{X,2}\boldsymbol{B}_2^\dagger = \bar{\boldsymbol{\Lambda}}\,, \qquad \boldsymbol{B}_2\boldsymbol{S}_{Y,2}\boldsymbol{B}_2^\dagger = \boldsymbol{I}\,.$$

Hence

$$\boldsymbol{S}_{X,1} = \boldsymbol{B}_1^{-1}\boldsymbol{B}_2\boldsymbol{S}_{X,2}\boldsymbol{B}_2^\dagger(\boldsymbol{B}_1^{-1})^\dagger = \boldsymbol{B}\boldsymbol{S}_{X,2}\boldsymbol{B}^\dagger$$
$$\boldsymbol{S}_{Y,1} = \boldsymbol{B}_1^{-1}\boldsymbol{B}_2\boldsymbol{S}_{Y,2}\boldsymbol{B}_2^\dagger(\boldsymbol{B}_1^{-1})^\dagger = \boldsymbol{B}\boldsymbol{S}_{Y,2}\boldsymbol{B}^\dagger$$

with $\boldsymbol{B} = \boldsymbol{B}_1^{-1}\boldsymbol{B}_2$, and the proof is thus completed.

Interestingly, the principle of invariance realizes a significant data reduction: The maximal invariant statistic is a real N-dimensional vector, whereas the original sufficient statistic is composed of the two $N \times N$ Grammian matrices $\boldsymbol{S}_X$ and $\boldsymbol{S}_Y$. For instance, with reference to three polarimetric channels, the original sufficient statistic is composed of 18 real entries, whereas in the compressed domain (i.e., after reduction by invariance), the maximal invariant is just 3-D.

B. Induced Maximal Invariant Design

Data transformation induces parameter transformation, which leaves unaltered the two composite hypotheses. In other words, by the principle of invariance, the parameter space is also partitioned into orbits and, usually, involves a reduced set of parameters. The relevant parameters are embodied into any *induced maximal invariant*, namely, any function of the parameters that is constant over each orbit of the parameter space (invariance) but assumes different values over different orbits (maximality).

For the case at hand, an induced maximal invariant is composed of the eigenvalues $\boldsymbol{\delta} = [\delta_1, \ldots, \delta_N]^T$ of the matrix

$$\boldsymbol{\Sigma}_X\boldsymbol{\Sigma}_Y^{-1}\,. \tag{9}$$

The proof of this claim can be done following the same steps as in the proof of Proposition 1 and is omitted for the sake of compactness.

The previous claim highlights that the principle of invariance yields a significant reduction of the number of parameters: In fact, the induced maximal invariant is an N-dimensional vector while the original parameter space was composed of the two covariance matrices $\boldsymbol{\Sigma}_X$ and $\boldsymbol{\Sigma}_Y$.

We explicitly observe that in the reduced parameter space, the partition corresponding to the two composite hypotheses of test (1) is $\Xi_0 = \{\mathbf{1}_N\}$, relative to $\boldsymbol{\Sigma}_X = \boldsymbol{\Sigma}_Y$, and $\Xi_1 = \overline{\{\mathbf{1}_N\}}$, relative to $\boldsymbol{\Sigma}_X \neq \boldsymbol{\Sigma}_Y$, where $\overline{\{\mathbf{1}_N\}}$ is the set of N-dimensional column vectors with at least one entry different from 1. The structure of Ξ_0, which now corresponds to a simple H_0 hypothesis, clearly shows that all invariant receivers that process a maximal invariant statistic through a transformation independent of $\delta_1, \ldots, \delta_N$, achieve the CFAR property.

IV. Design of the Optimum and Suboptimum Invariant Detectors

This section is devoted to the design of the optimum and suboptimum invariant detectors. Section IV-A deals with the design of the MPI detector for the problem of interest and the existence of a UMPI test. Since no UMPI detector exists, Section IV-B introduces some suboptimum invariant architectures.

A. Design of the MPI Detector

Based on the NP criterion, the MPI test can be obtained as the LRT based on any maximal invariant [15], [23], which, as already pointed out, provides sufficient information for constructing any invariant decision rule. It is thus necessary, for further developments, to introduce the pdf of the devised maximal invariant under both the H_0 and the H_1 hypotheses [20]

$$f_{\boldsymbol{\Lambda}}(\boldsymbol{\Lambda}|H_1, \boldsymbol{\Delta}) = c_{N,K}\det(\boldsymbol{\Delta})^{-K}\,{}_1\tilde{F}_0(2K; -\boldsymbol{\Delta}^{-1}, \boldsymbol{\Lambda}) \times \det(\boldsymbol{\Lambda})^{K-N}\prod_{j=1}^{N}\prod_{i<j}(\lambda_i - \lambda_j)^2 \tag{10}$$

$$f_{\boldsymbol{\Lambda}}(\boldsymbol{\Lambda}|H_0) = f_{\boldsymbol{\Lambda}}(\boldsymbol{\Lambda}|H_1, \boldsymbol{I}) = c_{N,K} \times \det(\boldsymbol{I}+\boldsymbol{\Lambda})^{-2K} \times \det(\boldsymbol{\Lambda})^{K-N}\prod_{j=1}^{N}\prod_{i<j}(\lambda_i-\lambda_j)^2 \tag{11}$$

where[2] $\boldsymbol{\Lambda} = \mathbf{Diag}([\lambda_1, \lambda_2, \ldots, \lambda_N]^T)$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N > 0$, $\boldsymbol{\Delta} = \mathbf{Diag}(\boldsymbol{\delta})$, ${}_1\tilde{F}_0(\cdot;\cdot,\cdot)$ denotes the hypergeometric function of matrix argument [20, eq. (88)], and $c_{N,K}$ is a normalization constant.

Hence, the LRT can be written as

$$\frac{f_{\boldsymbol{\Lambda}}(\boldsymbol{\Lambda}|H_1, \boldsymbol{\Delta})}{f_{\boldsymbol{\Lambda}}(\boldsymbol{\Lambda}|H_0)} = \frac{{}_1\tilde{F}_0(2K; -\boldsymbol{\Delta}^{-1}, \boldsymbol{\Lambda})}{\det(\boldsymbol{I}+\boldsymbol{\Lambda})^{-2K}} \underset{H_0}{\overset{H_1}{\gtrless}} T \tag{12}$$

where T denotes the detection threshold, set to ensure a given P_{fa} level.

[2]With a slight notation abuse, we continue to indicate with $\lambda_i, i = 1, \ldots, N$, the ordered (in descending order) eigenvalues.

In order to proceed further, we focus on the case $N = 2$ and exploit expression [25, eq. (51)] to recast ${}_1\tilde{F}_0(2K; -\boldsymbol{\Delta}^{-1}, \boldsymbol{\Lambda})$ as a rational function, i.e.,

$$
\begin{aligned}
&{}_1\tilde{F}_0(2K; -\boldsymbol{\Delta}^{-1}, \boldsymbol{\Lambda}) \\
&\quad = \frac{c_2 \det \begin{bmatrix} \left(1+\frac{\lambda_1}{\delta_1}\right)^{1-2K} & \left(1+\frac{\lambda_1}{\delta_2}\right)^{1-2K} \\ \left(1+\frac{\lambda_2}{\delta_1}\right)^{1-2K} & \left(1+\frac{\lambda_2}{\delta_2}\right)^{1-2K} \end{bmatrix}}{(\lambda_1-\lambda_2)\left(-\frac{1}{\delta_1}+\frac{1}{\delta_2}\right)}, \quad \delta_i \neq \delta_j
\end{aligned} \tag{13}
$$

with c_2 as a constant with respect to λ_i and δ_i.

Moreover, for $\delta_1 = \delta_2 = \delta$, the splitting formula [20, eq. (92)] yields

$$
{}_1\tilde{F}_0(2K; -\boldsymbol{\Delta}^{-1}, \boldsymbol{\Lambda}) = \det\left(\boldsymbol{I} + \frac{\boldsymbol{\Lambda}}{\delta}\right)^{-2K}, \qquad \delta_i = \delta_j = \delta\,. \tag{14}
$$

Otherwise stated, we have (15), shown at the bottom of the page.

As a consequence, after standard statistical equivalences, the MPI test can be recast as

$$
\begin{aligned}
&\left[\frac{(1+\lambda_1)}{(\delta_1+\lambda_1)}\frac{(1+\lambda_2)}{(\delta_2+\lambda_2)}\right]^{2K}\left[\frac{(\delta_1+\lambda_1)(\delta_2+\lambda_2)}{\lambda_1-\lambda_2}\right] + \\
&\quad -\left[\frac{(1+\lambda_1)}{(\delta_2+\lambda_1)}\frac{(1+\lambda_2)}{(\delta_1+\lambda_2)}\right]^{2K}\left[\frac{(\delta_2+\lambda_1)(\delta_1+\lambda_2)}{\lambda_1-\lambda_2}\right] \\
&\quad \times \operatorname{sign}(\delta_1-\delta_2) \underset{H_0}{\overset{H_1}{\gtrless}} T_A\,, \qquad \delta_1 \neq \delta_2
\end{aligned} \tag{16}
$$

$$
\frac{(1+\lambda_1)(1+\lambda_2)}{(\delta+\lambda_1)(\delta+\lambda_2)} \underset{H_0}{\overset{H_1}{\gtrless}} T_B\,, \qquad \delta_1 = \delta_2 = \delta \tag{17}
$$

with T_A and T_B suitable modifications of the original detection threshold in (12).

The derived expression of the MPI test clearly highlights that it cannot, in general, be implemented unless the induced maximal invariant is known (clairvoyant detector): Unfortunately, such hypothesis is not realistic, particularly for the SAR change detection problem. In other words, the MPI detector, which provides the best performance for a given $\boldsymbol{\Delta}^\star$, is not, in general, optimum for different values of $\boldsymbol{\Delta}$, namely, no UMPI detector exists. Nevertheless, the MPI detector is still noteworthy since its performance upper bounds those of any other invariant receiver, operating under the same signal and disturbance models and, hence, can be used to assess the loss of any implementable, although suboptimal, invariant test. Before concluding, we highlight that the case $N = 3$ can be dealt with a similar technique (the derivations are given in Appendix A).

B. Maximal-Invariant-Based Detectors

The lack of a UMPI test suggests an investigatigation of invariant decision rules based upon different strategies. However, there is no criterion for choosing *a priori* a receiver instead of another. An intuitive rule to select invariant tests for our problem could be based on the following asymptotic observation. For large values of K, the eigenvalues of $\boldsymbol{S}_X^{-1}\boldsymbol{S}_Y$ tend to δ_i, $i = 1, \ldots, N$; hence, decision rules

$$
h(\lambda_1, \ldots, \lambda_N) \underset{H_0}{\overset{H_1}{\gtrless}} T \tag{18}
$$

1) are very effective in discriminating deviations $\delta_i \gg 1$, when $h(\cdot, \ldots, \cdot)$ is an increasing function of the arguments. However, they perform poor when δ_i are smaller than 1.
2) are very effective in discriminating deviations $\delta_i \ll 1$, when $h(\cdot, \ldots, \cdot)$ is a decreasing function of the arguments. Nevertheless, they perform poor when δ_i is greater than 1.
3) in principle could achieve good detection levels for both $\delta_i \gg 1$ and $\delta_i \ll 1$, when $h(\cdot, \ldots, \cdot)$ complies with $h((1/\lambda_1), \ldots, (1/\lambda_N)) = h(\lambda_1, \ldots, \lambda_N)$.

On the other hand, one cannot analyze all possible reasonable detectors; hence, in the following, we focus on six decision rules, which, based on extensive numerical analysis, are seen to achieve satisfactory detection performance.

1) GLRT: This approach is equivalent to replacing the unknown parameters in the likelihood ratio with their maximum-likelihood estimates, under each hypothesis [26]. Interestingly, under very mild technical assumptions, the GLRT is invariant [27]. For the present problem, it has been proposed in [9] and [10][3] and is given by

$$
\frac{\max\limits_{\boldsymbol{\Sigma}_X}\max\limits_{\boldsymbol{\Sigma}_Y} f_{\boldsymbol{S}_X,\boldsymbol{S}_Y}(\boldsymbol{S}_X, \boldsymbol{S}_Y | H_1, \boldsymbol{\Sigma}_X, \boldsymbol{\Sigma}_Y)}{\max\limits_{\boldsymbol{\Sigma}_X} f_{\boldsymbol{S}_X,\boldsymbol{S}_Y}(\boldsymbol{S}_X, \boldsymbol{S}_Y | H_0, \boldsymbol{\Sigma}_X)} \underset{H_0}{\overset{H_1}{\gtrless}} T_1 \tag{19}
$$

with T_1 as the detection threshold. After optimizations and monotonic transformations, (19) can be shown to be statistically equivalent to

$$
\prod_{i=1}^{N} \frac{(1+\lambda_i)^2}{\lambda_i} \underset{H_0}{\overset{H_1}{\gtrless}} T_1 \tag{20}
$$

[3]With reference to real observations, it is derived in [14, Ch. 8].

$$
{}_1\tilde{F}_0(2K; -\boldsymbol{\Delta}^{-1}, \boldsymbol{\Lambda}) = \begin{cases} \dfrac{c_2\left[\left(1+\frac{\lambda_1}{\delta_1}\right)^{1-2K}\left(1+\frac{\lambda_2}{\delta_2}\right)^{1-2K} - \left(1+\frac{\lambda_1}{\delta_2}\right)^{1-2K}\left(1+\frac{\lambda_2}{\delta_1}\right)^{1-2K}\right]}{(\lambda_1-\lambda_2)\left(-\frac{1}{\delta_1}+\frac{1}{\delta_2}\right)} & \delta_i \neq \delta_j \\ \left(1+\frac{\lambda_1}{\delta}\right)^{-2K}\left(1+\frac{\lambda_2}{\delta}\right)^{-2K} & \delta_1 = \delta_2 = \delta \end{cases} \tag{15}
$$

where the same symbol T_1 has been used to denote the modified detection threshold.

Interestingly, the GLRT complies with condition 3 that was given at the beginning of the section, namely, we expect the GLRT to be capable of achieving good detection levels both when $\delta_i \gg 1$ and $\delta_i \ll 1$.

2) Arithmetic- and Harmonic-Mean-Based Detectors: These decision rules are, respectively, given by

$$\sum_{i=1}^{N} \lambda_i \underset{H_0}{\overset{H_1}{\gtrless}} T_2 \tag{21}$$

$$\sum_{i=1}^{N} \frac{1}{\lambda_i} \underset{H_0}{\overset{H_1}{\gtrless}} T_3 \tag{22}$$

where T_2 and T_3 are the detection thresholds.

The former complies with condition 1, whereas the latter complies with condition 2. As a consequence, (21)[4] is suitable for detecting deviations $\delta_i \gg 1$ while (22) for $\delta_i \ll 1$.

From (21) and (22), it is also possible to construct another decision rule satisfying condition 3, merely summing up the decision statistics, i.e.,

$$\sum_{i=1}^{N} \left(\frac{1}{\lambda_i} + \lambda_i \right) \underset{H_0}{\overset{H_1}{\gtrless}} T_4 \tag{23}$$

with T_4 as the detection threshold. Before concluding this paragraph, we highlight that (21) can be also obtained substituting in the optimum LRT detector (5) the sample covariance matrices $(1/K)\boldsymbol{S}_X$ and $(1/K)\boldsymbol{S}_Y$ in place of $\boldsymbol{\Sigma}_X$ and $\boldsymbol{\Sigma}_Y$, respectively. Otherwise stated, the arithmetic-mean-based detector can be interpreted as an adaptive implementation of the LRT.

3) Maximum and Minimum Eigenvalue-Based Detectors: These tests are, respectively, based on the following comparisons:

$$\left(\lambda_1 + \frac{1}{\lambda_N} \right) \underset{H_0}{\overset{H_1}{\gtrless}} T_5 \tag{24}$$

$$\max \left(\lambda_1 \, , \, \frac{1}{\lambda_N} \right) \underset{H_0}{\overset{H_1}{\gtrless}} T_6 \tag{25}$$

with T_5 and T_6 as the decision thresholds. An intuitive explanation to the decision rules is based on the following arguments: The former term, i.e., λ_1, dominates for large deviations $\delta_i \gg 1$, whereas the latter term, i.e., $1/\lambda_N$, if $\delta_i \ll 1$. Hence, (24) and (25) are supposed to perform well both for $\delta_i \gg 1$ and $\delta_i \ll 1$.

V. Performance Analysis on Simulated Data

This section presents the performance analysis via computer simulated data of the detectors introduced in Section IV. It is organized in two subsections. The first analyzes P_d, for a given P_{fa}, versus the induced maximal invariant, which represents the set of relevant parameters ruling the P_d's behavior (otherwise stated, different pairs of covariance matrices $(\boldsymbol{\Sigma}_X, \boldsymbol{\Sigma}_Y)$ sharing the same induced maximal invariant $\boldsymbol{\delta}$ lead to the same detection performance). The second shows standard ROCs for a given value of $\boldsymbol{\delta}$. Finally, the impact of the number of polarimetric channels used to perform change detection is studied.

A. Detection Probability Versus the Induced Maximal Invariant

The analysis is conducted in terms of P_d for a given P_{fa} level, assuming zero-mean complex circular multivariate Gaussian observations with equal (distinct) covariance matrices under H_0 (H_1). For comparison purposes, the MPI test is used as a benchmark to the performance of any feasible invariant receiver. Finally, a square inspection window is considered, namely, $W_1 = W_2 = W$.

The case of two polarizations is chosen because for $N = 2$, the induced maximal invariant, which rules the performance of any invariant detector, is bidimensional. This means that contour plots of P_d versus the components of the induced maximal invariant δ_1 and δ_2 completely characterize the detection performance of receivers (20)–(25). Specifically, the curves can be interpreted as follows. Given a pair of true covariance matrices $(\boldsymbol{\Sigma}_X, \boldsymbol{\Sigma}_Y)$, compute the bidimensional induced maximal invariant $\boldsymbol{\delta}$ and, from the countours, read the value of P_d corresponding to a given P_{fa} level.

In the present analysis, Monte Carlo simulation is used to set the detection thresholds assuming $100/P_{fa}$ independent runs and $P_{fa} = 10^{-4}$. Moreover, 5000 independent trials are used to estimate P_d. Fig. 2 shows the contours corresponding to $W = 3$. It highlights that receivers complying with condition 3 in Section IV-B are capable of detecting deviations δ_i either $\gg 1$ or $\delta_i \ll 1$. Moreover, (23), (24), and (25) achieve a comparable detection performance level, namely, $P_d \geq 0.9$ for $\delta_i \notin [10^{-1.3}\,, 10^{1.3}]$, $i = 1, 2$, whereas the same level of performance is achieved by (20) for $\delta_i \notin [10^{-1.36}\,, 10^{1.36}]$, $i = 1, 2$. A different behavior is shown by tests (21) and (22); the former operates with a satisfactory detection performance when $\delta_i \notin [0\,, 10^{1.3}]$, $i = 1, 2$, whereas the latter when $(\delta_1, \delta_2) \in \,]0\,, 10^{-1.3}]^2$. This is a useful feature as, shown on real data, tests (21) and (22) can be used for postdetection classification, namely, to discriminate between $\delta_i \gg 1$ and $\delta_i \ll 1$.

In Fig. 3, the performance of the considered invariant tests is compared with the benchmark MPI receiver. For this test, $W = 3$ and four different cuts of the detection probability surface (P_d versus δ_1 and δ_2) are considered: $\delta_2 = \delta_1$, $\delta_2 = 0.75\delta_1$, $\delta_2 = 0.50\delta_1$, and $\delta_2 = 0.25\delta_1$. The curves highlight that tests (21) and (22), respectively, optimized to the specific situation $\delta \gg 1$ (21) and $\delta \ll 1$ (22) slightly outperform the other detection rules; additionally, their performance is near the optimum curve in the region $\delta_1 > 10$ (21) and $\delta_1 < 0.1$ (22). However, their P_d is close to zero when $\delta_1 < 10$ (21) and $\delta_1 > 0.1$ (22). All the remaining receivers follow the same performance behavior as the benchmark curve.

[4]This detector is also known as the Hotelling–Lawley trace and has been discussed for change detection purposes in [28].

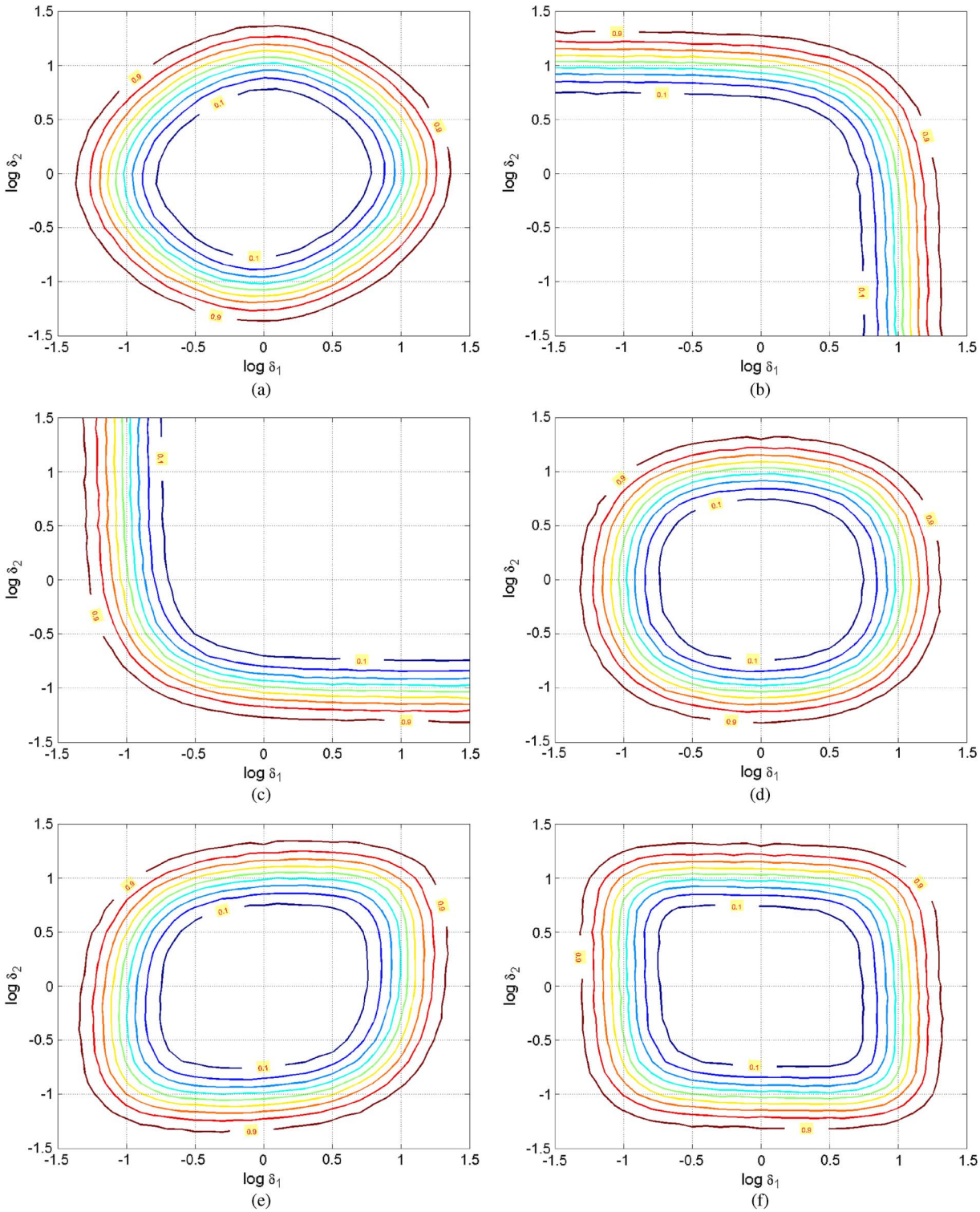


Fig. 2. P_d contours versus $\log \delta_1$ and $\log \delta_2$ for $N = 2$ and $W = 3$. (a) Detector (20): $\prod_{i=1}^{N}((1+\lambda_i)^2/\lambda_i)$. (b) Detector (21): $\sum_{i=1}^{N}\lambda_i$. (c) Detector (22): $\sum_{i=1}^{N}(1/\lambda_i)$. (d) Detector (23): $\sum_{i=1}^{N}((1/\lambda_i)+\lambda_i)$. (e) Detector (24): $(\lambda_1 + (1/\lambda_N))$. (f) Detector (25): $\max(\lambda_1\,,\,(1/\lambda_N))$.

The case $W = 5$ is not reported for conciseness; however, it highlights that increasing W leads to better detection performances. Moreover, the performance gap between the considered invariant detectors and the MPI test is smaller than that observed for $W = 3$. This was expected and can be explained observing that a greater number of homogeneous data vectors have been used to construct the Grammians $\boldsymbol{S}_X$ and $\boldsymbol{S}_Y$ whose scaled versions ($1/K$ scale factor) are unbiased and consistent estimates of the covariance matrices $\boldsymbol{\Sigma}_X$ and $\boldsymbol{\Sigma}_Y$.

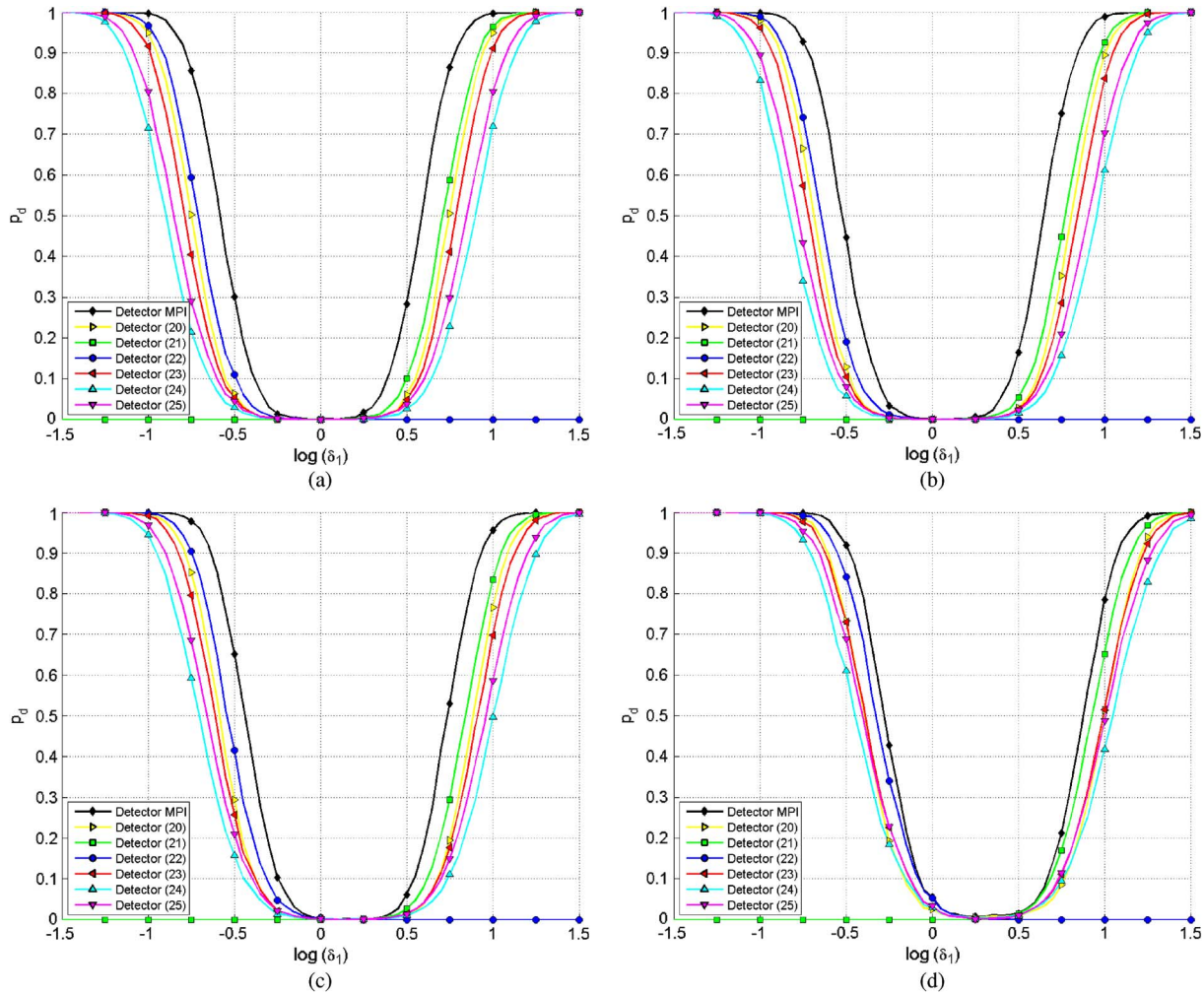


Fig. 3. P_d versus $\log \delta_1$ for $N = 2$ and $W = 3$. Detector (20): $\prod_{i=1}^{N}((1+\lambda_i)^2/\lambda_i)$; detector (21): $\sum_{i=1}^{N}\lambda_i$; detector (22): $\sum_{i=1}^{N}(1/\lambda_i)$; detector (23): $\sum_{i=1}^{N}((1/\lambda_i)+\lambda_i)$; detector (24): $(\lambda_1 + (1/\lambda_N))$; detector (25): $\max(\lambda_1\ ,\ (1/\lambda_N))$. (a) $\delta_1 = \delta_2$. (b) $\delta_1 = 0.75\delta_2$. (c) $\delta_1 = 0.5\delta_2$. (d) $\delta_1 = 0.25\delta_2$.

B. Standard ROC Analysis

The analysis is conducted in terms of ROC curves (namely, P_d versus P_{fa}) for a fixed value of the induced maximal invariant $\boldsymbol{\delta} = [1/2, 1/2, 1/2]^T$, corresponding to the pair of covariance matrices exploited in [10]. Of course, any other covariance pair with the same value of $\boldsymbol{\delta}$ leads to the same ROCs, whereas a different value of $\boldsymbol{\delta}$ leads to different ROCs. For comparison purposes, the benchmark ROC of the clairvoyant MPI detector and that of the detector proposed in [10] are also plotted. The detector proposed in [10] is an adaptive implementation of the LRT, whose expression in terms of the maximal invariant statistic is

$$\sum_{i=1}^{N}\left(\frac{1}{\lambda_i} - \ln\frac{1}{\lambda_i}\right) \underset{H_0}{\overset{H_1}{\gtrless}} T_H\,. \tag{26}$$

In order to set the detection threshold, Monte Carlo simulations are used assuming $100/P_{fa}$ independent runs. In the simulations, 5000 independent trials were used to estimate P_d. Fig. 4 shows the ROCs of the considered receivers for both two[5] and three polarimetric channels, i.e., $W = 3$ and $W = 5$.

In all the analyzed cases, the ROC highlights that the harmonic-mean-based detector (22) outperforms the counterparts, and the loss with respect to the benchmark curve is acceptable for almost all the parameter values chosen in the simulation. The GLRT (20) and the detector (23) achieves almost the same performance levels and come third and are ranked just after detector (26). Considering the arithmetic-mean-based detector, its P_d for the given value of $\boldsymbol{\delta}$ is completely unsatisfactory, which is not surprising and is actually anticipated by the theory, since the decision statistic belongs to the class 1) of Section IV-B. Summarizing, the simulation curves evidently show that the GLRT (20) is often outperformed by the counterparts, and as predicted by the theory, there is not a detector that uniformly outperforms the others.

[5]For the covariance pairs in [10], the HH, VV, and HV polarimetric channels are perfectly equivalent because the induced maximal invariant exhibits equal components.

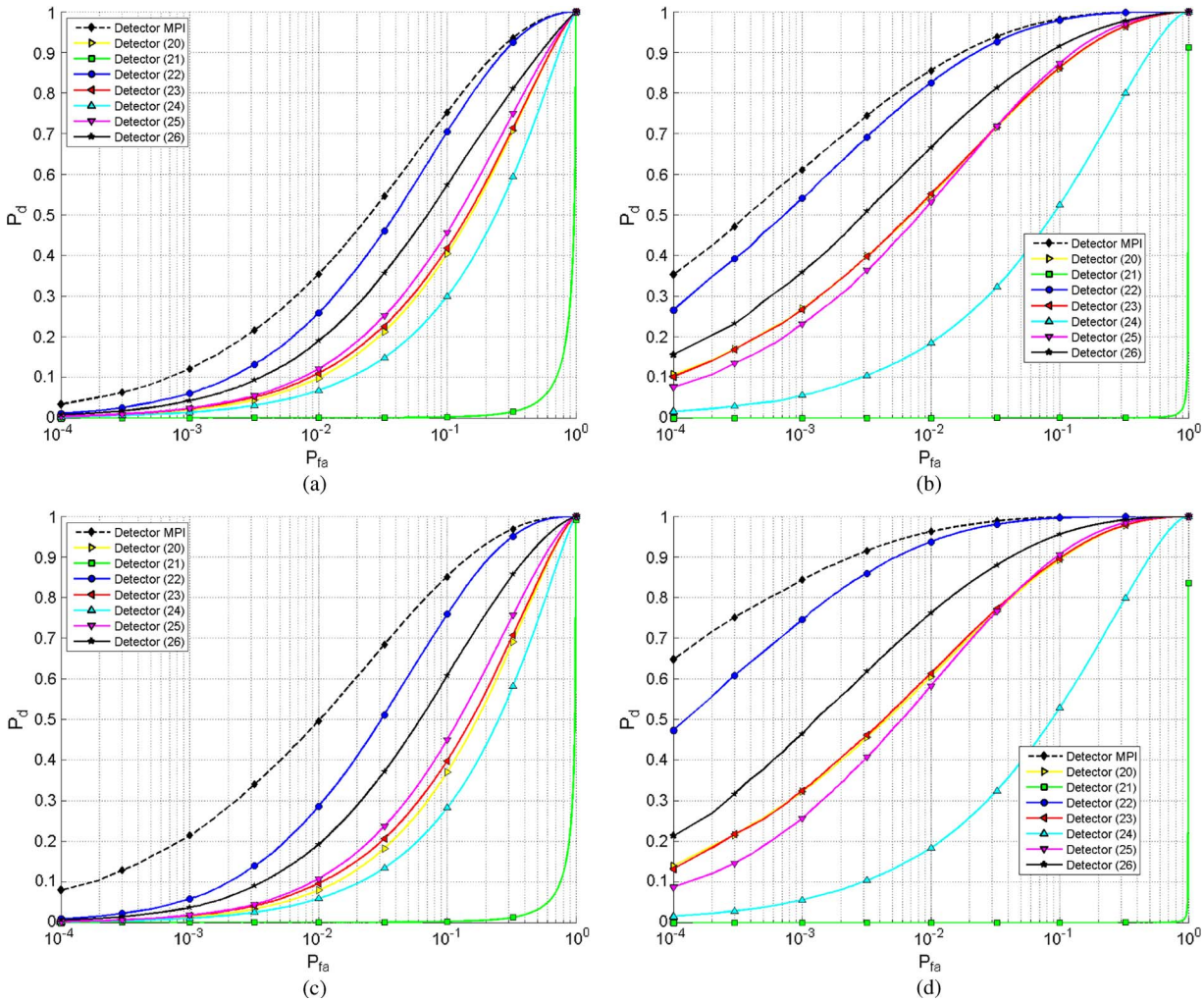


Fig. 4. P_d versus P_{fa}. Detector (20): $\prod_{i=1}^{N}((1+\lambda_i)^2/\lambda_i)$; detector (21): $\sum_{i=1}^{N}\lambda_i$; detector (22): $\sum_{i=1}^{N}(1/\lambda_i)$; detector (23): $\sum_{i=1}^{N}((1/\lambda_i)+\lambda_i)$; detector (24): $(\lambda_1+(1/\lambda_N))$; detector (25): $\max(\lambda_1,(1/\lambda_N))$; detector (26): $\sum_{i=1}^{N}((1/\lambda_i)-\ln(1/\lambda_i))$. (a) $N=2$, $W=3$. (b) $N=2$, $W=5$. (c) $N=3$, $W=3$. (d) $N=3$, $W=5$.

The effects on the performance of the number of processed polarimetric channels are studied in Fig. 5. For conciseness, the analysis is limited to the ROC curves of detectors (20), (22), (25), and (26), which are reported for $W=5$, $N=3$, $N=2$, and $N=1$. Based on footnote 4, the curves for $N=1$ can be either interpreted as the ROC of the HH, VV, or HV channel. It is also worth mentioning that for $N=1$, no UMPI detector exists. In fact, for $N=1$, the MPI receiver

$$\frac{\boldsymbol{S}_X}{\boldsymbol{S}_Y}\text{sign}(\delta-1)\underset{H_0}{\overset{H_1}{\gtrless}}T_{A,1}$$

requires knowledge of the induced maximal invariant $\delta=\boldsymbol{\Sigma}_X/\boldsymbol{\Sigma}_Y$.

The analysis of the curves in Fig. 5 shows that the detection performance for a given P_{fa} level improves as N increases. For example, with reference to detector (20) and a P_{fa} of 10^{-3}, a P_d of 0.18 is achieved for $N=1$, whereas for $N=2$ and $N=3$, the P_d results are 0.27 and 0.32, respectively. This is in accordance with the intuitive observation that the higher the polarimetric information about the scene, the better the performance.

So far, all the numerical analyses performed in this section were based on the induced maximal invariant value $\boldsymbol{\delta}=[1/2,1/2,1/2]^T$. Thus, in order to provide an average behavior of the ROC curves for the considered detectors, the following simulation is conducted.

1) 500 random values[6] of the induced maximal invariant are selected [modeling the entries of $\boldsymbol{\delta}$ as independent and identically distributed uniform random variables within the intervals (0, 1), (1, 2), and (0, 2)]. The corresponding three simulation setups will be referred to in the following as case 1, case 2, and case 3.
2) The ROC curve of a given detector is evaluated in correspondence of each induced maximal invariant realization.
3) For each detector, the previously obtained 500 ROC curves are averaged for a specific simulation setup (i.e., case 1, 2, or 3).

[6]The seed of the rand MATLAB function is set to the initial state so as to guarantee the reproducibility of the curves.

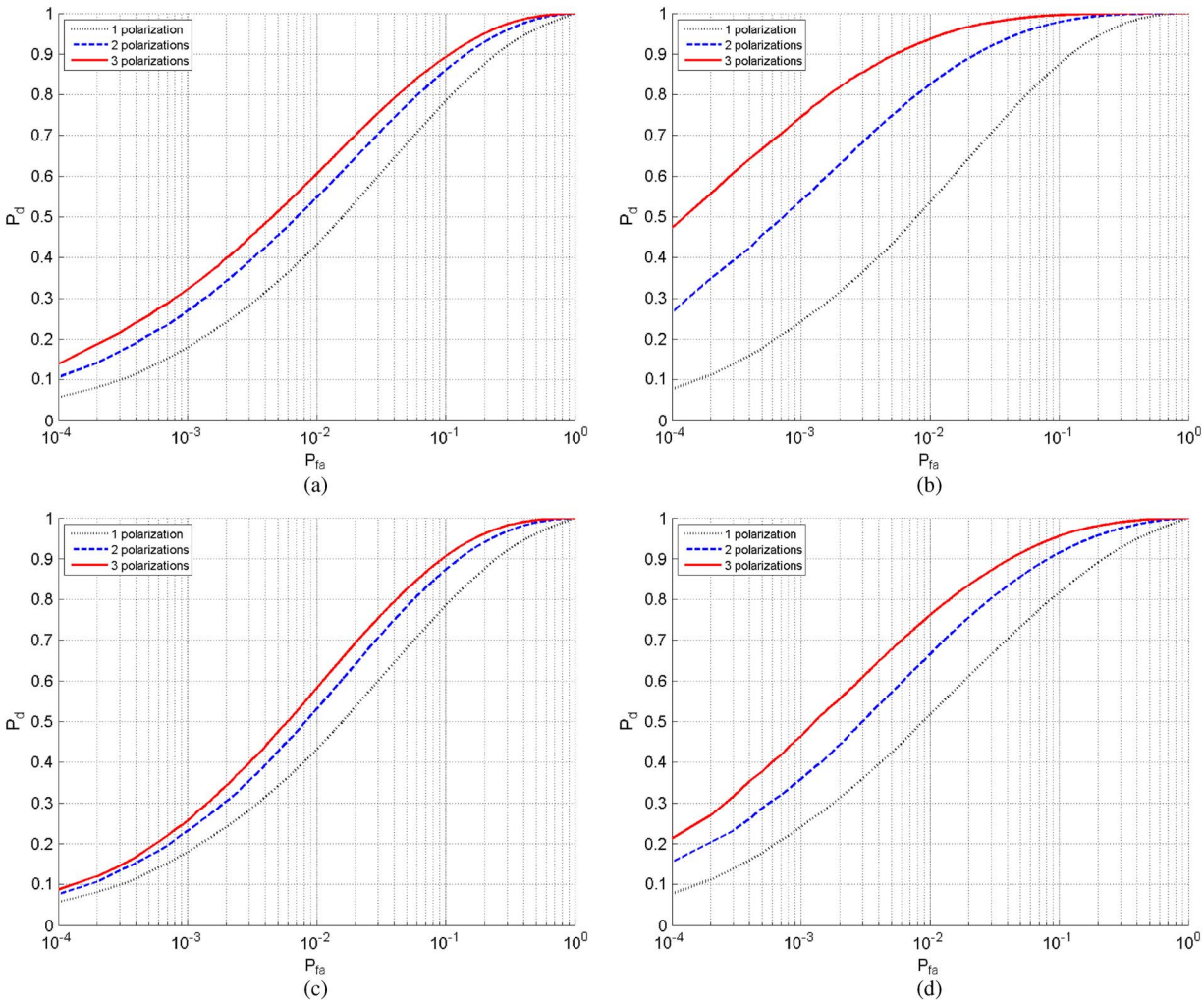


Fig. 5. P_d versus P_{fa} for $W = 5$. (a) Detector (20): $\prod_{i=1}^{N}((1+\lambda_i)^2/\lambda_i)$. (b) Detector (22): $\sum_{i=1}^{N}(1/\lambda_i)$. (c) Detector (25): $\max(\lambda_1\,,\,(1/\lambda_N))$. (d) Detector (26): $\sum_{i=1}^{N}((1/\lambda_i)-\ln(1/\lambda_i))$.

The results are displayed in Fig. 6 where, for conciseness, only the full-polarization case is shown. As expected, when the induced maximal invariant is simulated according to case 1, the ROC curves agree with those shown in Fig. 4 with detector (22) performing better than the counterparts and (21) providing the poorest performance.

On the contrary, when case 2 is considered, a different behavior of the ROC curves is observed. In particular, in the hierarchy, detectors (21) and (22) switched their positions compared with the previous case (this agrees with the theoretical considerations at the beginning of Section V-B). Finally, when case 3 is simulated, the ROC curves result in between the two previous situations. This can be explained observing case 3; the performance can be obtained as the average of that in cases 1 and 2. Summarizing this last analysis confirms that there is not a detector able to able to uniformly outperform the others. This implies that a battery of detectors could be considered with a suitable aggregation logic or some a priori knowledge (if any) about the characteristics of the operating polarimetric environment could be exploited to select the most suitable decision rule. Otherwise stated, based on the operating scenario (urban, sea, and foliage) and the targets (car, tank, and boats) of tactical interest for change detection, it is possible to exploit a priori models for the polarimetric covariance matrices, which allow to gain some information about the expected induced maximal invariant. Based on this last vector, the most suitable detector can be selected within the class of proposed invariant rules (for instance, using the plots in Fig. 2 if $N = 2$).

VI. Testing on Real Data

Here, the performance analysis of the algorithms proposed in Section IV is presented. The analysis is performed using real X-band data; the data set used is the Coherent Change Detection Challenge data set acquired by the AFRL [8]. The airborne SAR used to acquire the data set employed a coherent receiver with 640-MHz bandwidth and dual-polarized mode. The data are in the form of focused complex images with a range and cross-range resolution of 0.3 m. The overall data set provides ten complex images for each of the three available polarizations (HH, VV, and HV), acquired on the same day. The original image size is 4501 $\times$ 4501 pixels, and the images have been

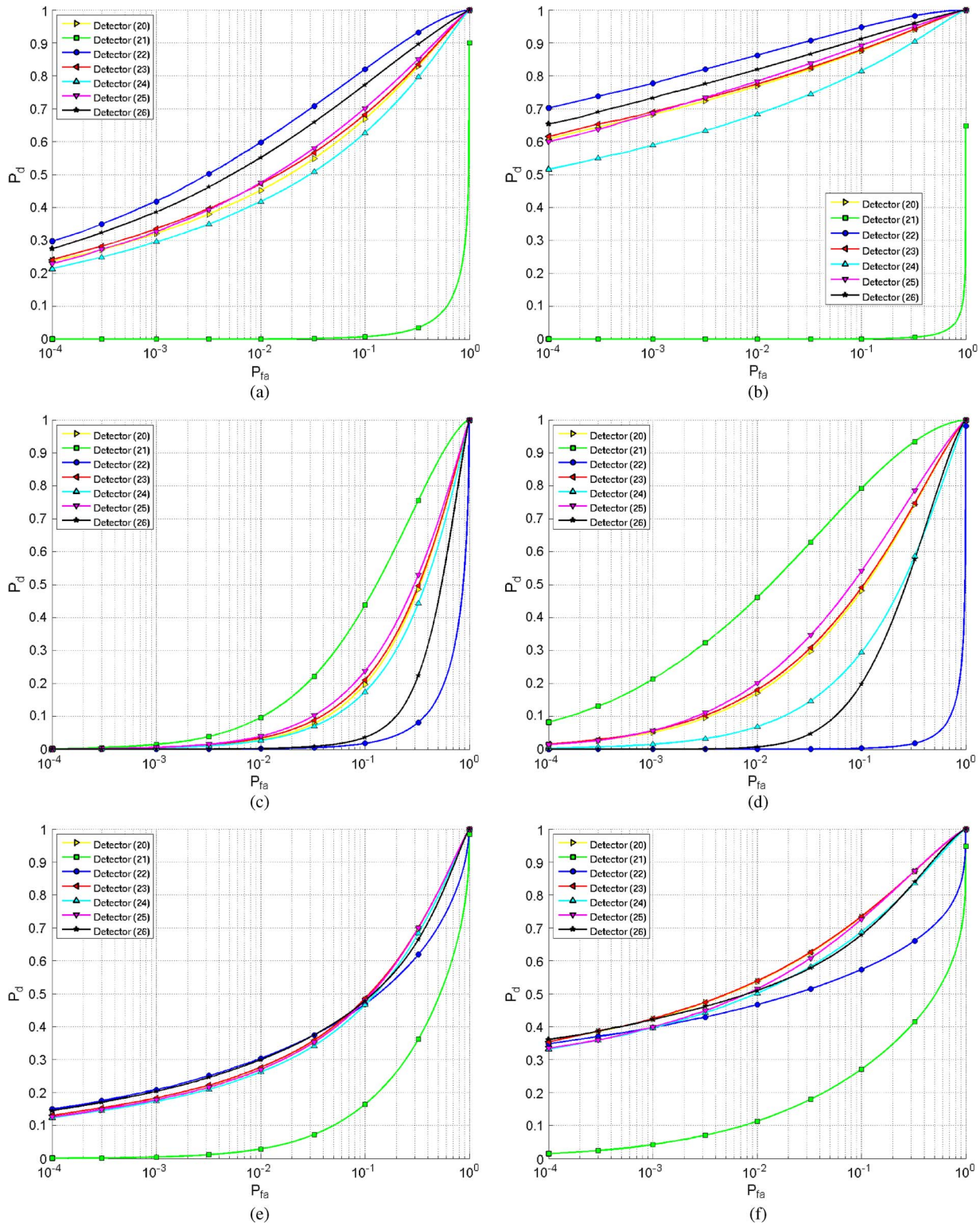
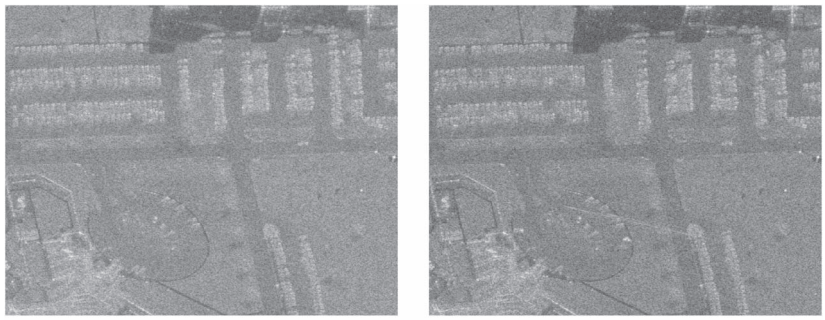


Fig. 6. P_d versus P_{fa}. Detector (20): $\prod_{i=1}^{N}((1+\lambda_i)^2/\lambda_i)$; detector (21): $\sum_{i=1}^{N}\lambda_i$; detector (22): $\sum_{i=1}^{N}(1/\lambda_i)$; detector (23): $\sum_{i=1}^{N}((1/\lambda_i)+\lambda_i)$; detector (24): $(\lambda_1+(1/\lambda_N))$; detector (25): $\max(\lambda_1, (1/\lambda_N))$; detector (26): $\sum_{i=1}^{N}((1/\lambda_i)-\ln(1/\lambda_i))$. (a) $N=3$, $W=3$, case 1. (b) $N=3$, $W=5$, case 1. (c) $N=3$, $W=3$, case 2. (d) $N=3$, $W=5$, case 2. (e) $N=3$, $W=3$, case 3. (f) $N=3$, $W=5$, case 3.

coherently aligned to a single reference (per polarization) with the help of Digital Elevation Map information [8].

For our analysis, we focus on two acquisitions from the entire data set. Unfortunately, the ground truths of the data are not available (e.g., the actual changes between two different acquisitions); hence, the selection of two passes providing the opportunity to generate a sufficiently accurate ground truth is required. Hence, two passes satisfying this requirement have



(a) (b)

Fig. 7. Reference (a) and test (b) images gathered in HH mode.

been identified: the acquisition named "FP0124" is used as reference pass, whereas the acquisition "FP0121" is used as a test pass. From the two passes, the area with the higher activity (changes) between the passes has been selected; this area is represented by a subimage of 1000×1000 pixels (i.e., $L = M = 1000$) and is composed of several parking lots that are occupied by numerous parked (i.e., stationary) vehicles. Fig. 7 shows the reference and test subimages for the HH mode.

For this particular scenario, the changes between the reference and test images (denoted by $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively), occurred during the time interval between the two acquisitions, can be distinguished in two cases as follows.

- A vehicle is present in $\boldsymbol{X}$ but is not present in $\boldsymbol{Y}$, i.e., the vehicle has departed from its parking space (pixels relative to this kind of event will be referred to in the following as departures).
- A vehicle is not present in $\boldsymbol{X}$ but is present in $\boldsymbol{Y}$, i.e., the vehicle has arrived in an empty parking space (pixels relative to this kind of event will be referred to in the following as arrivals).

Using the cases defined above, a total of 34 changes between $\boldsymbol{X}$ and $\boldsymbol{Y}$ can be visually identified (by flickering the two images). In the analysis, the straight line crossing the test image has not been considered, as its nature does not represent an arrival. However, as it is visible in the test image but not visible in the reference image, we expect it to be detected as an arrival. The obtained ground truth is shown in Fig. 8, where the black regions represent the departures, and the white regions indicate the arrivals. In particular, denoting by $\mathcal{K}$ the set of pixels that correspond to changes, the ground truth can be represented as a matrix $\boldsymbol{G}$ whose entries are given by

$$\boldsymbol{G}(l,m) = \begin{cases} 1 & \text{if } (l,m) \in \mathcal{K} \quad l = 1, \ldots, L \\ 0 & \text{otherwise} \qquad\quad m = 1, \ldots, M\,. \end{cases} \tag{27}$$

In Fig. 9(a), the ground truth mask $\boldsymbol{G}(l,m)$ is shown. Although the acquisitions were performed during the same day and the images were registered, the returns from a scatterer contribute differently to neighbor pixels, for example, a slightly different aspect angle can produce a different amount of energy spillover. These relative differences in the imaged data can lead to false alarms in the change detection results. In order to prevent false alarm caused by pixel contamination by target returns, we consider a guard area around each arrival–departure. This allows the definition of an extended ground truth [see Fig. 9(b)] used in the following to compare the performance of the considered detection algorithms.

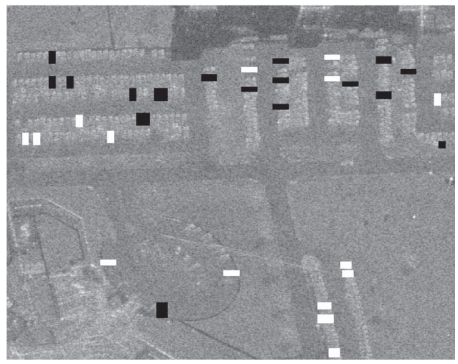
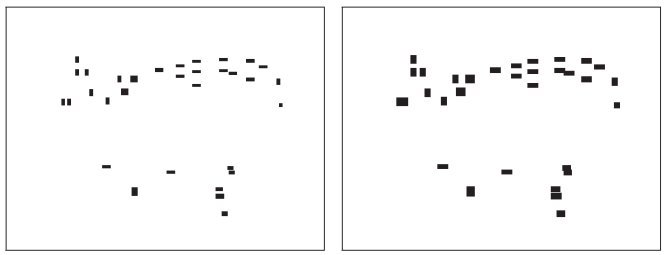


Fig. 8. Ground truth superimposed to the reference image.



(a) (b)

Fig. 9. Ground truth without (a) and with (b) the addition of guard cells.

In order to assess the performance of the detectors, both the number of detected changes and the change detection maps are presented. For the ith receiver, the corresponding map of changes $\boldsymbol{C}_i$ is a $L \times M$ matrix whose (l,m)th entry is the ith decision statistic considering the $N \times N$ sample Grammian matrix $\boldsymbol{S}_{X_{l,m}}$ ($\boldsymbol{S}_{Y_{l,m}}$) evaluated considering a square neighbourhood[7] with size $W \times W$ of the pixel (l,m) of $\boldsymbol{X}$ ($\boldsymbol{Y}$).

[7]We notice that, in order to obtain $\boldsymbol{C}_i$ of size $L \times M$, we include a frame of ε pixel width of both reference and test images with $\varepsilon = (W-1)/2$, in order to be able to compute the statistics on the image borders. By doing so, W must be odd.

TABLE I
NUMBER OF CORRECT DETECTIONS IN THE EXTENDED GROUND TRUTH. ALL THE COMPARISONS HAVE BEEN DONE USING THE SAME P_{fa} LEVEL, NAMELY, THE NUMBER OF THRESHOLD CROSSINGS IN THE COMPLEMENT OF THE GROUND TRUTH IS EXACTLY THE SAME FOR ALL THE DETECTORS. THE BEST DETECTOR, FOR THE CONSIDERED DATA SET AND P_{fa} LEVEL, IS THUS THE ONE ENSURING THE HIGHEST NUMBER OF DETECTIONS WITHIN THE GROUND TRUTH REGION (NUMBER OF PIXELS IN THE GROUND TRUTH REGION WHERE THE DECISION STATISTIC IS OVER THE THRESHOLD)

Detector			(20)	(21)	(22)	(23)	(24)	(25)	(26)
Statistic			$\prod_{i=1}^{N}\frac{(1+\lambda_i)^2}{\lambda_i}$	$\sum_{i=1}^{N}\lambda_i$	$\sum_{i=1}^{N}\frac{1}{\lambda_i}$	$\sum_{i=1}^{N}\left(\frac{1}{\lambda_i}+\lambda_i\right)$	$\left(\lambda_1+\frac{1}{\lambda_N}\right)$	$\max\left(\lambda_1,\frac{1}{\lambda_N}\right)$	$\sum_{i=1}^{N}\left(\frac{1}{\lambda_i}-\ln\frac{1}{\lambda_i}\right)$
$W=5$	$N=1$	HH	4512	3977	2301	4512	4512	4512	2301
		VV	5190	3941	2485	5190	5190	5190	2759
		HV	5557	4043	2796	5557	5557	5557	2818
	$N=2$	HH-VV	5372	5129	2337	5103	4877	4980	2300
		HH-HV	6088	4932	3019	5969	5597	5778	2946
		VV-HV	6425	4962	3396	6540	6165	6553	3274
	$N=3$	HH-VV-HV	6492	5513	2884	5901	5463	5644	2819
$W=9$	$N=1$	HH	5671	5426	2765	5671	5671	5671	2797
		VV	7712	5627	3684	7712	7712	7712	5564
		HV	7854	5653	3798	7854	7854	7854	5055
	$N=2$	HH-VV	6544	6591	2664	5852	5334	5433	2601
		HH-HV	8034	6381	3692	7525	6889	7140	3593
		VV-HV	9009	6773	4580	8994	8355	9005	5086
	$N=3$	HH-VV-HV	8300	7049	3389	7120	6233	6386	3255

The detection map corresponding to $\boldsymbol{C}_i$ is then defined as

$$\boldsymbol{D}_i(l,m)=\begin{cases}1, & \text{if } \boldsymbol{C}_i(l,m)>T_i \qquad l=1,\ldots,L\\ 0, & \text{otherwise} \qquad\quad m=1,\ldots,M\end{cases} \tag{28}$$

where T_i denotes the detection threshold. In the analysis presented in this section, the thresholds are set to ensure $P_{fa}=10^{-3}$ in the complement of the extended ground truth area, namely, in the region where no changes occur (there are no true positives). This means that, for each detector, after computing the decision statistics (for each pixel belonging to the complement of the extended ground truth), the threshold has been selected in order that

$$10^{-3}\times\text{total number of available statistics (trials)}$$

are greater than the threshold. This ensures that all the comparisons refer to the same P_{fa} level, namely, the number of threshold crossings in the complement of the extended ground truth is exactly the same for all the analyzed detectors.

Finally, considered are the cases of $N=1$, $N=2$, and $N=3$, which correspond, respectively, to processing the single polarimetric channel (HH, VV, or HV), to jointly considering a polarization pair (HH-VV, HH-HV, or VV-HV), or to jointly considering all the available channels (HH, VV, and HV). In particular, for each detector, starting from $\boldsymbol{D}_i$, and denoting by $\mathcal{K}_G$ the set of the pixels that correspond to changes in the extended ground truth area, performances are given in terms of the number of detections belonging to $\mathcal{K}_G$, i.e., considering the cardinality of the set

$$\mathcal{D}_i=\{\boldsymbol{D}_i(l,m)\,:\,\boldsymbol{D}_i(l,m)\in\mathcal{K}_G,\}$$

where G represents the number of guard cells used to generate the extended ground truth map. In this analysis, $G=5$ is selected, which means that around each set of pixels representing a change, a 5-pixel frame is included.

The number of detections for each receiver corresponding to an actual change present in the extended ground truth is summarized in Table I. As expected, the common trend is that the performance improves by increasing W. Moreover, as to the effects of polarization, the following remarks are in order.

- Processing either VV or HV channel leads, in general, to better performances than processing the HH polarization; the cross-polarized returns provide the highest number of correct detections for single-channel processing (the sole exception is receiver (26) for $W=5$; in this case, the VV channel leads to the best choice). The high discrimination capability of the cross-polarized returns can be explained observing that in the presence of a car target, there are dihedrals that significantly depolarize the incident radiation. On the contrary, the depolarization effect is much limited in the absence of a car since the parking area is much more similar to a flat plane. Moreover, the intuitive reason for better performances achievable with VV polarization

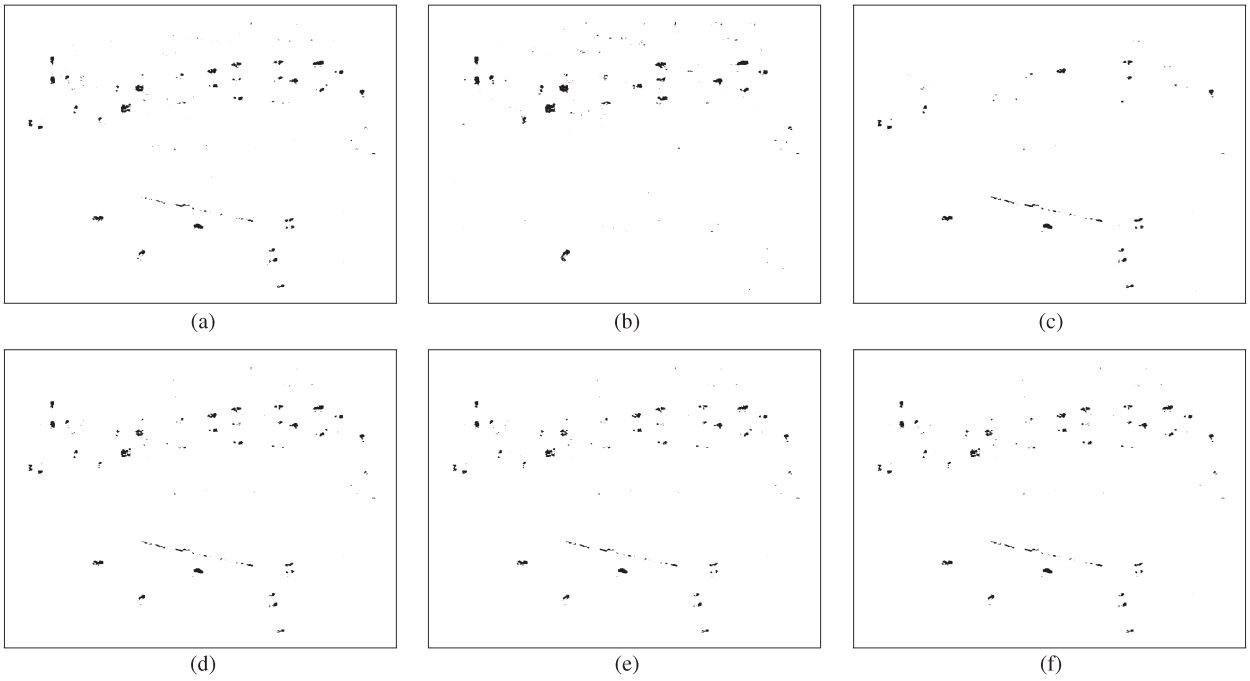


Fig. 10. Change detection maps for $W = 5$ and $N = 3$. (a) Detector (20): $\prod_{i=1}^{N}((1+\lambda_i)^2/\lambda_i)$. (b) Detector (21): $\sum_{i=1}^{N}\lambda_i$. (c) Detector (22): $\sum_{i=1}^{N}(1/\lambda_i)$. (d) Detector (23): $\sum_{i=1}^{N}((1/\lambda_i)+\lambda_i)$. (e) Detector (24): $(\lambda_1 + (1/\lambda_N))$. (f) Detector (25): $\max(\lambda_1\ ,\ (1/\lambda_N))$.

compared with those obtained with HH polarization resides in the dependence of the polarization return on the specific change considered (absence or presence of the car object in the image) and on its geometry (height extension and acquisition geometry).

- With reference to two polarimetric channels, jointly processing the pair VV-HV seems the best choice for almost all the considered decision rules (with exceptions represented by detectors (21), for $W = 5$, for which the HH-VV pair is the most suitable choice, and detector (26), for $W = 9$, for which polarization diversity does not seem to provide detection improvements).
- A suitable two-channel polarization diversity can outperform the single-channel processing (the sole exception is detector (26) for $W = 9$).
- Processing three polarimetric channels does not always improve the results obtained using two polarizations; the possible improvement depends both on parameter W and on the specific decision rule. In particular, for $W = 9$, the use of polarization diversity of order two with the pair VV-HV appears more convenient (the sole exception is detector (21) whose correct detections increases exploiting the full polarimetric information).

The maps obtained for the case with $W = 5$ and $N = 3$ for the detectors (20)–(25) are shown in Fig. 10. Based on them as well as on the results in Table I, it can be observed that detector (25) ensures the best performance for $W = 5$, whereas detector (20) outperforms the counterparts for $W = 9$. Moreover, for $W = 3$, additional results (not reported in the table for conciseness) highlights that the best performance is obtained exploiting detector (25) processing with the pair VV-HV.

Of particular interest are tests (21) and (22): From the corresponding detection maps, it is easy to recognize that the former identifies the departures, whereas the latter identifies the arrivals. This behavior is also clearly visible in the maps shown in Fig. 11. In fact, for detector (21), the departures present higher values, and the arrivals present lower values, whereas for detector (22), the dual behavior is obtained. Additionally, although detector (25) is outperformed by detectors (20) and (23), it allows us to discriminate between the departures and the arrivals based on the argument of $\max$ in (25). In particular, if the first argument in (25) is selected, the change corresponds to a departure, whereas it corresponds to an arrival if the second argument is selected. By considering the additional information coming from the maximum argument in (25), it is thus possible to discriminate the departures from the arrivals detected by the receiver (25); in Fig. 12(b), this enhanced detection map is shown together with the ground truth.

A. Assessing the Impact of an Aggregation Procedure

Analyzed here are the effects of an aggregation strategy after single-pixel detection so as to eliminate isolated false alarms and confirm true detections, which, due to the typical car size and system resolution, appear quite clustered (see the detection maps in Fig. 10). More precisely, a window of size 5×5 slides along the horizontal and the vertical dimensions of the detection maps (which, as already highlighted are binary images: "0" no detection, "1" detection). Then, for a given pixel localized at the center of the moving window and labeled with "1," a detection is associated if the number of "1," in the window is greater that a certain integer "fill parameter" (denoted by F and complying with $1 \leq F \leq 25$).

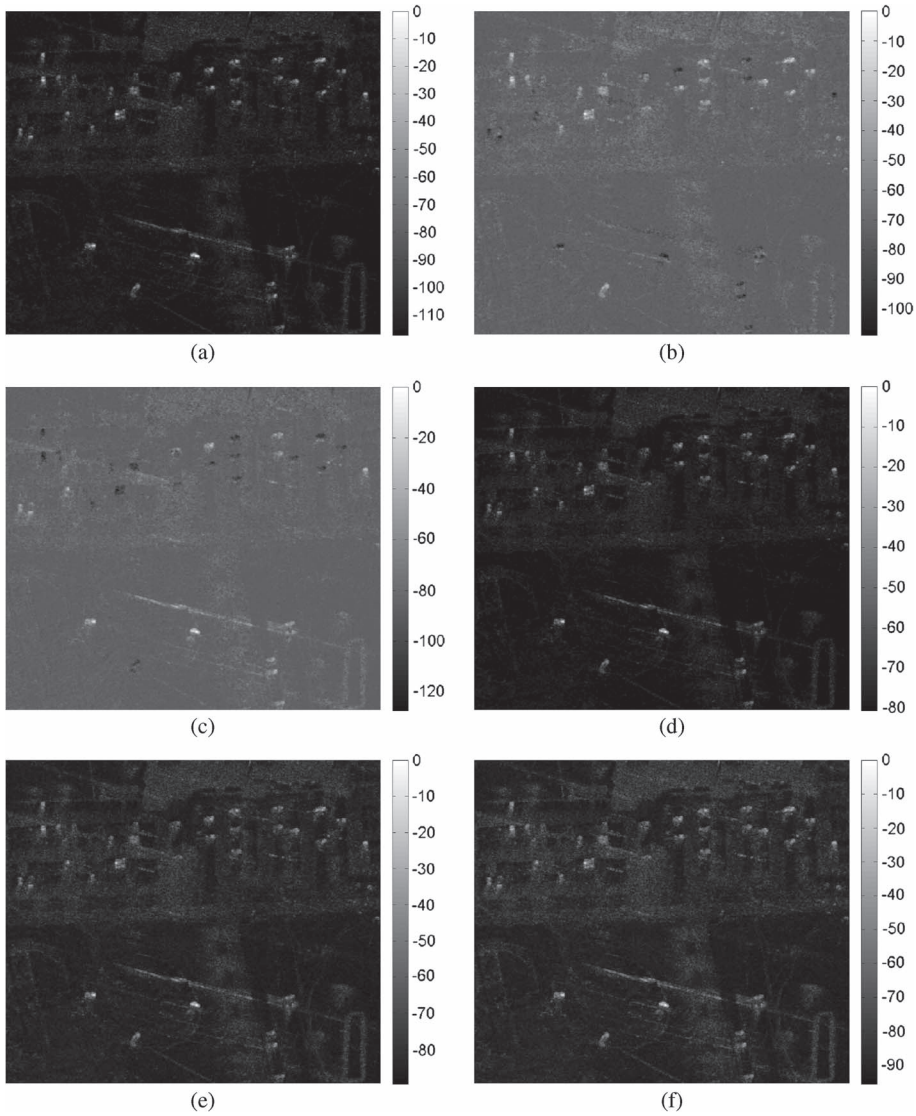


Fig. 11. Maps of changes in decibels for $W = 5$ and $N = 3$. (a) Detector (20): $\prod_{i=1}^{N}((1+\lambda_i)^2/\lambda_i)$. (b) Detector (21): $\sum_{i=1}^{N}\lambda_i$. (c) Detector (22): $\sum_{i=1}^{N}(1/\lambda_i)$. (d) Detector (23): $\sum_{i=1}^{N}((1/\lambda_i)+\lambda_i)$. (e) Detector (24): $(\lambda_1 + (1/\lambda_N))$. (f) Detector (25): $\max(\lambda_1\,,\,(1/\lambda_N))$.

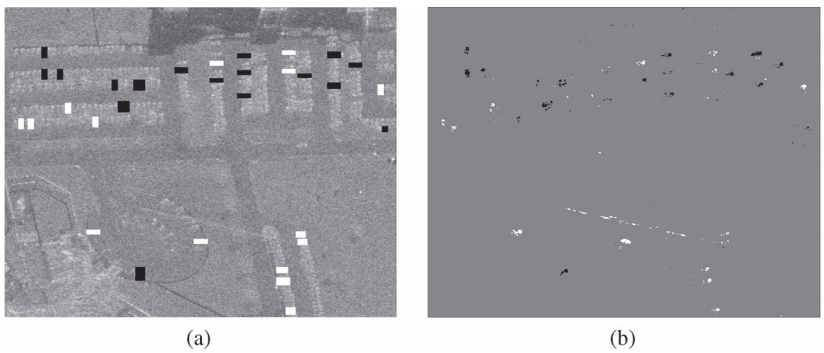


Fig. 12. Ground truth and enhanced detection map considering $W = 5$ and $N = 3$ for detector (25): $\max(\lambda_1\,,\,(1/\lambda_N))$. On both images in black and white are shown the departures and the arrivals, respectively. (a) Ground Truth. (b) Detector (25).

This kind of logics looks like n-of-m aggregation techniques [24, Ch. 3] and can be interpreted as a postdetection binary integration within the reference window. As to the pixels on the image edge (namely, those laying on the first/last two rows or columns), no aggregation is performed because they never fall at the center of the moving window. Otherwise stated, it is simply confirmed the "0" or "1" value in the original detection map.

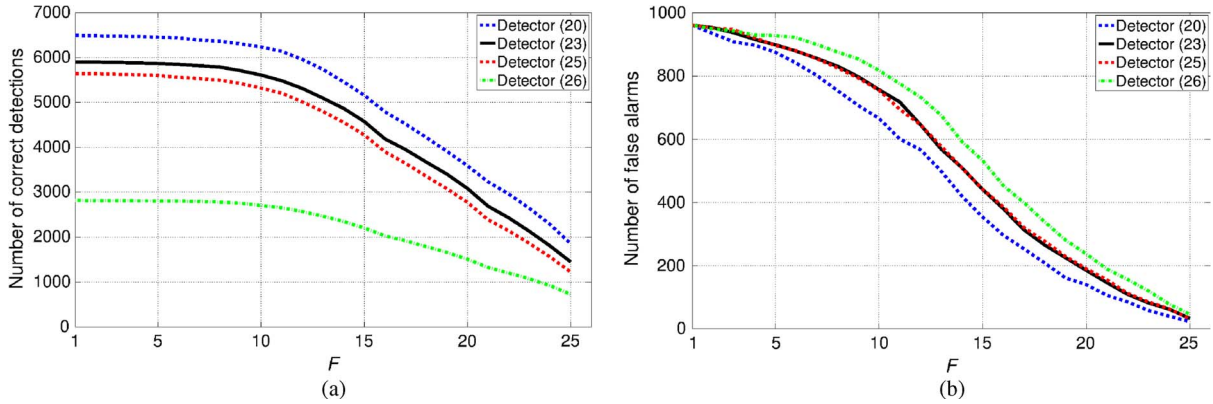


Fig. 13. (a) Number of correct detections in the extended ground truth and (b) number of false alarms in the complement of the extended ground truth versus F. Detector (20): $\prod_{i=1}^{N}((1+\lambda_i)^2/\lambda_i)$; detector (23): $\sum_{i=1}^{N}((1/\lambda_i)+\lambda_i)$; detector (25): $\max(\lambda_1\ ,\ (1/\lambda_N))$; detector (26): $\sum_{i=1}^{N}((1/\lambda_i)-\ln(1/\lambda_i))$. (a) Correct detections. (b) False alarms.

The case of three polarimetric channels is considered, and the effect of the fill parameter is studied in Fig. 13 for detectors (20), (23), (25), and (26).

The conducted deleted analysis highlights that P_d is not very sensitive to the fill parameter when it ranges between 5% and 40% of the window size. A similar result was also found in [10] with reference to single-channel change detection. Interestingly, there exists a range of values for parameter F (precisely $1 \leq F \leq 10$), where P_d is almost flat, whereas P_{fa} decreases as F increases. This suggests that a fill parameter value between 8 and 10 can be reasonably used in the aggregation procedure, reducing the number of false alarms while keeping (almost) constant the number of true positives.

VII. Conclusion

Multipolarization SAR change detection has been considered in this paper. The problem has been formulated as a binary hypothesis test, and the principle of invariance has been applied to design decision rules exhibiting a special symmetry, which is a sufficient condition to ensure the CFAR property. A maximal invariant statistic is found, and the optimum invariant detector has been computed showing that it is not UMPI and, as a consequence, not practically implementable. Hence, some suboptimum invariant tests, whose decision statistics are given in terms of the obtained maximal invariant, have been introduced and assessed both on simulated data and on real high-resolution SAR images. The conducted analysis has shown that some of them represent a viable mean to deal with the change detection problem. An important result relates to the performance obtained considering the VV-HV polarization. This particular choice (for the considered real data set) often resulted in superior performance compared with all the other single and multipolarization cases. Moreover, it is also possible to exploit suitable invariant tests for postdetection classification purposes, particularly, from the available real data, to discriminate between car arrivals and departures in the parking area without any additional RCS comparison. Possible future research tracks will consider the extension of the framework relaxing the Gaussian requirement for the data as well as the analysis on other data sets acquired by a different system, possibly at different resolutions (different performance behaviors could be observed on different data sets). Moreover, in order to deal with possible amplitude and phase calibration errors, future investigations will deal with the design of decision rules robust to miscalibration effects. The idea is to achieve this goal by imposing other invariances at the price of some detection losses. Finally, it would be interesting to investigate the joint use of multifrequency and multipolarization data to further improve the effect of diversity.

Appendix

Here, the MPI detector is derived with reference to the case of $N = 3$ polarimetric channels. To this end, it is necessary to distinguish among three cases

- **Case 1**: $\delta_1 \neq \delta_2 \neq \delta_3$. Using [25], eq. (51), ${}_1\tilde{F}_0(2K; -\boldsymbol{\Delta}^{-1}, \boldsymbol{\Lambda})$ can be recast as a rational function, i.e.,

$$ {}_1\tilde{F}_0(2K; -\boldsymbol{\Delta}^{-1}, \boldsymbol{\Lambda}) = \frac{c_3 \det\left[\left(1+\frac{\lambda_j}{\delta_i}\right)^{2-2K}\right]_{i,j=1,\ldots,3}}{\prod_{j=1}^{3}\prod_{i<j}\left(-\frac{1}{\delta_i}+\frac{1}{\delta_j}\right)(\lambda_i-\lambda_j)} \tag{29} $$

  with c_3 as a constant with respect to λ_i and δ_i. Thus, the MPI detector becomes

$$ \frac{\det\left[\left(1+\frac{\lambda_j}{\delta_i}\right)^{2-2K}\right]_{i,j=1,\ldots,3}}{\prod_{j=1}^{3}(1+\lambda_j)^{-2K}\prod_{i<j}\left(-\frac{1}{\delta_i}+\frac{1}{\delta_j}\right)(\lambda_i-\lambda_j)} \underset{H_0}{\overset{H_1}{\gtrless}} T_{A,3}\,. \tag{30} $$

- **Case 2**: $\delta_1 = \delta_2 \neq \delta_3$ or $\delta_1 \neq \delta_2 = \delta_3$ or $\delta_1 = \delta_3 \neq \delta_2$. Let us focus on $\delta_1 = \delta_2 \neq \delta_3$ (the other two cases are equivalent to the considered one). Denoting by $\delta = \delta_1 = \delta_2$ and exploiting [29, eq. (106)], the equality chain (31), (32), and (33), shown at the top of the next page, holds true. As a consequence, the MPI test becomes (34), shown at the top of the next page.

$$ {}_1\tilde{F}_0(2K; -\Delta^{-1}, \Lambda) \tag{31} $$

$$ = \frac{c_3}{\prod_{j=1}^{3}\prod_{i<j}(\lambda_i - \lambda_j)} \lim_{\substack{x_1 \to -\frac{1}{\delta} \\ x_2 \to -\frac{1}{\delta}}} \frac{\det \begin{bmatrix} (1-\lambda_1 x_1)^{2-2K} & (1-\lambda_1 x_2)^{2-2K} & \left(1+\frac{\lambda_1}{\delta_3}\right)^{2-2K} \\ (1-\lambda_2 x_1)^{2-2K} & (1-\lambda_2 x_2)^{2-2K} & \left(1+\frac{\lambda_2}{\delta_3}\right)^{2-2K} \\ (1-\lambda_3 x_1)^{2-2K} & (1-\lambda_3 x_2)^{2-2K} & \left(1+\frac{\lambda_3}{\delta_3}\right)^{2-2K} \end{bmatrix}}{(x_2 - x_1)\left(-x_1 - \frac{1}{\delta_3}\right)\left(-x_2 - \frac{1}{\delta_3}\right)} \tag{32} $$

$$ = c_3 \frac{\det \begin{bmatrix} = \left(1+\frac{\lambda_1}{\delta}\right)^{2-2K} & -\lambda_1(2-2K)\left(1+\frac{\lambda_1}{\delta}\right)^{1-2K} & \left(1+\frac{\lambda_1}{\delta_3}\right)^{2-2K} \\ \left(1+\frac{\lambda_2}{\delta}\right)^{2-2K} & -\lambda_2(2-2K)\left(1+\frac{\lambda_2}{\delta}\right)^{1-2K} & \left(1+\frac{\lambda_2}{\delta_3}\right)^{2-2K} \\ \left(1+\frac{\lambda_3}{\delta}\right)^{2-2K} & -\lambda_3(2-2K)\left(1+\frac{\lambda_3}{\delta}\right)^{1-2K} & \left(1+\frac{\lambda_3}{\delta_3}\right)^{2-2K} \end{bmatrix}}{\left(\frac{1}{\delta} - \frac{1}{\delta_3}\right)^2 \prod_{j=1}^{3}\prod_{i<j}(\lambda_i - \lambda_j)} \tag{33} $$

$$ \frac{\det \begin{bmatrix} \left(1+\frac{\lambda_1}{\delta}\right)^{2-2K} & -\lambda_1(2-2K)\left(1+\frac{\lambda_1}{\delta}\right)^{1-2K} & \left(1+\frac{\lambda_1}{\delta_3}\right)^{2-2K} \\ \left(1+\frac{\lambda_2}{\delta}\right)^{2-2K} & -\lambda_2(2-2K)\left(1+\frac{\lambda_2}{\delta}\right)^{1-2K} & \left(1+\frac{\lambda_2}{\delta_3}\right)^{2-2K} \\ \left(1+\frac{\lambda_3}{\delta}\right)^{2-2K} & -\lambda_3(2-2K)\left(1+\frac{\lambda_3}{\delta}\right)^{1-2K} & \left(1+\frac{\lambda_3}{\delta_3}\right)^{2-2K} \end{bmatrix}}{\prod_{j=1}^{3}(1+\lambda_j)^{-2K}\prod_{i<j}(\lambda_i - \lambda_j)} \underset{H_0}{\overset{H_1}{\gtrless}} T_{B,3} \tag{34} $$

- **Case 3**: $\delta_1 = \delta_2 = \delta_3$. Using the splitting formula [20, eq. (92)] yields

$$ {}_1\tilde{F}_0(2K; -\boldsymbol{\Delta}^{-1}, \boldsymbol{\Lambda}) = \det\left(\boldsymbol{I} + \frac{\boldsymbol{\Lambda}}{\delta}\right)^{-2K}, \quad \delta_1 = \delta_2 = \delta_3 = \delta\,. \tag{35} $$

Hence, after standard equivalences, the MPI test (12) becomes

$$ \prod_{j=1}^{3}\left(\frac{1+\lambda_j}{\delta+\lambda_j}\right) \underset{H_0}{\overset{H_1}{\gtrless}} T_{C,3}\,. \tag{36} $$

ACKNOWLEDGMENT

The authors would like to thank the Editors and the Referees for the interesting questions and comments that have helped improve this article.

REFERENCES

[1] M. Preiss and N. J. S. Stacy, "Coherent change detection: Theoretical description and experimental results," Intell., Surveillance Reconnaissance Division, Defence Sci. Technol. Org., Canberra, ACT, Australia, Tech. Rep., DSTO-TR-1851, 2006.

[2] C. Jakowatz, Jr., D. Wahl, P. Eichel, D. Ghiglia, and P. Thompson, *Spotlight-mode Synthetic Aperture Radar: A Signal Processing Approach*. Boston, MA, USA: Kluwer, 1996.

[3] R. Touzi, A. Lopes, J. Bruniquel, and P. W. Vachon, "Coherence estimation for SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 1, pp. 135–149, Jan. 1999.

[4] I. Stojanovic and L. Novak, "Algorithms improve synthetic aperture radar coherent change detection performance," *SPIE Newsroom*, Jul. 2013.

[5] I. Stojanovic and L. Novak, "Change detection experiments using Gotcha public release SAR data," in *Proc. SPIE 8746, Algorithms Synthetic Aperture Radar Imagery XX*, Jun. 3, 2013, p. 87460I.

[6] L. Novak, "Coherent change detection for multi-polarization SAR," in *Conf. Rec. 39th Asilomar Conf. Signals, Syst. Comput.*, Oct. 2005, pp. 568–573.

[7] E. J. M. Rignot and J. J. Van Zyl, "Change detection techniques for ERS-1 SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 4, pp. 896–906, Jul. 1993.

[8] S. Scarborough *et al.*, "A challenge problem for SAR change detection and data compression," in *Proc. SPIE 7699, Algorithms Synthetic Aperture Radar Imagery XVII*, Apr. 18, 2010, p.76990U.

[9] K. Conradsen, A. A. Nielsen, J. Schou, and H. Skriver, "A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 1, pp. 4–19, Jan. 2003.

[10] L. M. Novak, "Change detection for multi-polarization, multi-pass SAR," in *Proc. SPIE Conf. Algorithms Synthetic Aperture Radar Imagery XII*, Orlando, FL, USA, Mar. 2005, pp. 234–246.

[11] A. A. Nielsen, R. Larsen, and H. Skriver, "Change detection in bi-temporal EMISAR data from Kalo, Denmark, by means of canonical correlations analysis," in *Proc. 3rd Int. Airborne Remote Sens. Conf. Exhib.*, Copenhagen, Denmark, Jul. 7–10, 1997, pp. 282–287.

[12] A. A. Nielsen, "Change detection in multispectral bi-temporal spatial data using orthogonal transformations," in *Proc. 8th Austral-Asian Sens. Conf.*, Canberra, ACT, Australia, Mar. 25–29, 1996, pp. 1–8.

[13] E. Erten, A. Reigber, L. Ferro-Famil, and O. Hellwich, "A new coherent similarity measure for temporal multichannel scene characterization,"

IEEE Trans. Geosci. Remote Sens., vol. 50, no. 7, pp. 2839–2851, Jul. 2012.
[14] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. New York, NY, USA: Wiley, 1982.
[15] E. L. Lehmann, *Testing Statistical Hypotheses*, 2nd ed. New York, NY, USA: Springer-Verlag, 1986.
[16] S. Bose and A. O. Steinhardt, "A maximal invariant framework for adaptive detection with structured and unstructured covariance matrices," *IEEE Trans. Signal Process.*, vol. 43, no. 9, pp. 2164–2175, Sep. 1995.
[17] A. De Maio and E. Conte, "Adaptive detection in Gaussian interference with unknown covariance after reduction by invariance," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 2925–2934, Jun. 2010.
[18] A. De Maio, E. Conte, and C. Galdi, "CFAR detection of multidimensional signals: An invariant approach," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 142–151, Jan. 2003.
[19] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, vol. 1. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.
[20] A. T. James, "Distribution of matrix variates and latent roots derived from normal samples," *Ann. Math. Statist.*, vol. 35, pp. 475–501, 1964.
[21] L. L. Scharf, *Statistical Signal Processing. Detection, Estimation, and Time Series Analysis*. Reading, MA, USA: Addison-Wesley, 1991.
[22] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, MA, USA: Cambridge Univ. Press, 1985.
[23] T. S. Ferguson, *Mathematical Statistics*. New York, NY, USA: Academic, 1967.
[24] M. A. Richards, J. A. Scheer, and W. A. Holm, *Principles of Modern Radar: Basic Principles*. Raleigh, NC, USA: Scitech Publishing, 2010.
[25] A. Yu. Orlov, "New solvable matrix integrals," *Int. J. Modern Phys. A*, vol. 19, pp. 276–293, 2004.
[26] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*, pt. 1. Hoboken, NJ, USA: Wiley, 1968.
[27] S. M. Kay and J. R. Gabriel, "An invariance property of the generalized likelihood ratio test," *IEEE Signal Process. Lett.*, vol. 10, no. 12, pp. 352–355, Dec. 2003.
[28] V. Akbari, S. Anfinsen, A. Doulgeris, and T. Eltoft, "The Hotelling–Lawley trace statistic for change detection in polarimetric SAR Data under the complex Wishart distribution," in *Proc. IEEE IGARSS*, Melbourne, Vic., Australia, Jul. 21–26, 2013, pp. 4162–4165.
[29] S. H. Simon, A. L. Moustakas, and L. Marinelli, "Capacity and character expansions: Moment-generating function and other exact results for MIMO correlated channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5336–5351, Dec. 2006.

Vincenzo Carotenuto (S'12) received the Laurea Specialistica degree in telecommunication engineering from the Università degli Studi di Napoli "Federico II," Napoli, Italy, in 2010. He is currently working toward the Ph.D. degree in electronic and telecommunication engineering in the Dipartimento di Ingegneria Elettrica e delle Tecnologie dell'Informazione, Università degli Studi di Napoli "Federico II."

His research interest lies in the field of statistical signal processing, with emphasis on radar signal processing.

Antonio De Maio (S'01–A'02–M'03–SM'07–F'13) was born in Sorrento, Italy, on June 20, 1974. He received the Dr. Eng. degree (with honors) and the Ph.D. degree in information engineering from the Università degli Studi di Napoli "Federico II," Napoli, Italy, in 1998 and 2002, respectively.

From October to December 2004, he was a Visiting Researcher with the U.S. Air Force Research Laboratory, Rome, NY, USA. From November to December 2007, he was a Visiting Researcher with the Chinese University of Hong Kong, Hong Kong. Currently, he is an Associate Professor with the Università degli Studi di Napoli "Federico II." His research interest lies in the field of statistical signal processing, with emphasis on radar detection, optimization theory applied to radar signal processing, and multiple-access communications.

Dr. De Maio was a recipient of the 2010 IEEE Fred Nathanson Memorial Award as the young (less than 40 years of age) AESS Radar Engineer 2010 whose performance is particularly noteworthy as evidenced by contributions to the radar art over a period of several years, with the following citation for "robust CFAR detection, knowledge-based radar signal processing, and waveform design and diversity."

Carmine Clemente (S'09–M'13) received the Laurea (cum laude) and Laurea Specialistica (cum laude) degrees in telecommunications engineering from the Universita' degli Studi del Sannio, Benevento, Italy, in 2006 and 2009, respectively, and the Ph.D. degree from the University of Strathclyde, Glasgow, U.K., in 2012.

He is currently a Research Associate with the Department of Electronic and Electrical Engineering, University of Strathclyde, working on advanced radar signal processing algorithms, multiple-input–multiple-output radar systems, and micro-Doppler analysis. His research interests include synthetic aperture radar (SAR) focusing and bistatic SAR focusing algorithm development, micro-Doppler signature analysis and extraction from multistatic radar platforms, micro-Doppler classification, and statistical signal processing.

John J. Soraghan (S'83–M'84–SM'96) received the B.Eng. (Hons.) and M.Eng.Sc. degrees from University College Dublin, Dublin, Ireland, in 1978 and 1983, respectively, and the Ph.D. degree from the University of Southampton, Southampton, U.K., in 1989, all in electronic engineering. His doctoral research focused on synthetic aperture radar processing on the distributed array processor.

After graduating, he worked with the Electricity Supply Board in Ireland and with Westinghouse Electric Corporation in the United States. In 1986, he joined the Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, U.K., as a Lecturer and became a Senior Lecturer in 1990, a Reader in 2000, and a Professor of signal processing in September 2003, within the Institute for Communications and Signal Processing (ICSP). In December 2005, he became the Head of the ICSP. He currently holds the Texas Instruments Chair in Signal Processing with the University of Strathclyde. He was a Manager of the Scottish Transputer Centre from 1988 to 1991 and a Manager of the DTI Parallel Signal Processing Centre from 1991 to 1995. His main research interests include signal processing theories, algorithms and architectures with applications to remote sensing, telecommunications, biomedicine, and condition monitoring.

Dr. Soraghan is a member of the Institution of Engineering and Technology.