Scalar and Representative Observables, and Their Associated Statistical Model, for Polarimetric SAR Data

Journal:	*Journal of Selected Topics in Applied Earth Observations and Remote Sensing*
Manuscript ID:	Draft
Manuscript type:	Regular
Date Submitted by the Author:	n/a
Complete List of Authors:	Le, Thanh-Hai; Nanyang Technological University, School of Computer Engineering Mcloughlin, Ian; University of Science and Technology of China, School of Information Science and Technology Vun, Chan Hua; Nanyang Technological University, School of Computer Engineering
Keywords:	Synthetic aperture radar, Polarimetric radar

Scalar and Representative Observables, and Their Associated Statistical Model, for Polarimetric SAR Data

Thanh-Hai Le, Ian Vince McLoughlin, and Chan-Hua Vun

***Abstract*—This paper proposes novel scalar and representative observables, including several statistically consistent discrimination measures, that would be useful in the interpretation and processing of multi-dimensional POLSAR data. Specifically, the statistical behaviour of the POLSAR covariance matrix determinant is used to derive a scalar and generic statistical model for multi-dimensional POLSAR data, which is shown applicable to the two and three dimensional versions of partial and full monostatic polarimetric SAR data. In addition, the proposed model is also able to subsume the traditional SAR intensity model under the umbrella of a unified model, where the SAR intensity is shown to be a one dimension version of the multi-dimensional data. Consequently, the main beneficial implications of the proposed approach are that it provides a consistent theory unifying the currently disconnected proposals for SAR and POLSAR discrimination measures, and it simplifies the adaptation of existing SAR data processing techniques for POLSAR data.**

***Index Terms*—Polarimetric Synthetic Aperture Radar, Electromagnetic Modeling, Multidimensional Signal Processing**

I. Introduction

Relentless growth in computing power has allowed the once computationally-demanding Synthetic Aperture Radar (SAR) to now become a feasible and preferred technique for earth observation. The basic SAR technology has also been extended in a few directions, one of which is polarimetric SAR (POLSAR). POLSAR exploits the natural polarization property of Electro-Magnetic (EM) waves, encoded in multiple channels, compared to traditional one-channel SAR.

While several models have been proposed for POLSAR data, they tend to be complex and unintuitive due to the multidimensional and stochastic nature of the data. It is therefore necessary to establish a simple and intuitive model for the understanding of the data, and statistical models are undoubtedly suitable due to its stochastic nature. Furthermore, practical POLSAR data processing also makes heavy use of scalar discrimination measures, which should be based on statistically consistent (i.e. signal-independent) models. It is thus important to establish scalar and representative observables for multi-dimensional POLSAR data.

A few scalar POLSAR observables with accompanying statistical models have been proposed [1–7], but none of them is able to provide widely-accepted scalar discrimination measures. As such no observable can be considered as being highly representative of this multi-dimensional data, which severely limits their applicability in practical data processing applications. Alternatively, a few popular POLSAR discrimination measures have been proposed [8–12], but all of them are based on likelihood ratio statistics (see Subsection II-B). This statistical test should ideally be based on an exact and consistent distribution but so far only asymptotic distributions have been demonstrated [1].

This paper presents scalar and representative observables, and an associated generic statistical model, that describes multi-dimensional POLSAR data to provide a consistent foundation for the derivation of discrimination measures. The development is based on the common assumption that POLSAR data follows the complex Wishart distribution. Subsequently, the generic statistical model for the covariance matrix determinant is presented as being just a scalar projection of the multidimensional data. This model is then used to derive several scalar and consistent statistical distributions suggesting that their associated observables are capable of being used as discrimination measures for POLSAR.

This paper also shows that the specific one dimensional (1-D) version of the proposed model matches the traditional statistical model used for SAR intensity. This effectively incorporates the existing SAR theory under the umbrella of the proposed scalar approach for multi dimensional POLSAR. The different discrimination proposals for SAR and POLSAR are reviewed in light of this and the proposal in this paper is shown to provide a strong, unifying and consistent foundation. The applicability of these theoretical models will be illustrated by experiments where the specific 1, 2 and 3-D versions of the proposed models are validated against real-life captured data. The other beneficial implication of the proposed approach will also be demonstrated and discussed, when the proposed model is applied, as an example illustrating their effectiveness, to the evaluation of POLSAR speckle filters.

The remainder of this article is organized as follows. After the next section reviews existing discrimination measures as well as scalar observables models for POLSAR, Section III derives the generic statistical model for the determinant of the POLSAR covariance matrix and proposes several new discrimination measures for the multi-dimensional data. Section IV then illustrates the correspondence between the specific one dimensional model, and traditional models, for SAR intensity. Following this, Section V links the disparate discrimination

Thanh-Hai Le and Chan-Hua Vun are with School of Computer Engineering, Nanyang Technological University, Singapore. Ian McLoughlin is with the School of Information Science and Technology, University of Science and Technology of China.

Manuscript received March, 2014; revised March, 2014.

measure proposals for traditional SAR with more recent POLSAR data. The applicability of these models against practical data is illustrated in Section VI. Section VII then presents a simple demonstration where an existing technique of SAR processing is extended for POLSAR data before Section VIII finally concludes the paper.

II. Related Work in Literature

Section II-A reviews various statistical models used for different scalar observables for POLSAR, where it is shown that none are able to provide statistically consistent discrimination measures. Section II-B further strengthens those findings by discussing discrimination measures that have been proposed for POLSAR, and demonstrates that almost all of them are based on the likelihood statistical test for complex Wishart distribution. While an exact statistical distribution is expected to be necessary for a test, only an asymptotic distribution is used in their underlying foundation [1].

A. Scalar Observables for POLSAR Data and their Statistical Models

Different target decomposition theorems have identified many possible scalar observables for complex POLSAR data. In [2], the performance of different scalar POLSAR observables is evaluated for classification purposes. While many scalar observables for POLSAR were presented, their corresponding statistical models and classifiers were not available. Furthermore, at its conclusion, the paper indicated that it is impossible to identify a single best representation. Although, to be fair, the observables were identified for describing a decomposed portion of the complex POLSAR data, rather than providing a unified representation of the POLSAR data.

Using a different approach, given that the joint distribution for POLSAR is known to be the multi-variate complex Wishart distribution, it is possible to derive the scalar statistical models for some univariate POLSAR observables. However, such derivations are not trivial tasks, and so far, only a handful of such statistical models have been proposed, including the following:

1) cross-pol ratio $r_{HV/HH} = |S_{HV}|^2/|S_{HH}|^2$ [3],
2) co-pol ratio $r_{VV/HH} = |S_{VV}|^2/|S_{HH}|^2$ [3],
3) co-pol phase difference $\phi_{VV/HH} = arg(S_{VV}S^*_{HH})$ [3] [4],
4) magnitude $g = |avg(S_{pq}S^*_{rs})|$ [4],
5) normalized magnitude $\xi = \frac{|avg(S_{pq}S^*_{rs})|}{\sqrt{avg(|S_{pq}|^2)avg(|S_{rs}|^2)}}$ [4],
6) intensity ratio $w = avg(|S_{pq}|^2)/avg(|S_{rs}|^2)$ [4],
7) and the Stokes parameters $S_i, 0 \le i \le 3$ [5].

More recently, statistical models for each element of the POLSAR covariance matrix, i.e. $S_{pq}S^*_{rs}$, [6] as well as for the largest eigen-value of the covariance matrix λ_1 [7] have been proposed.

While these models undoubtedly help to further our understanding of POLSAR data, none of the underlying observables have been shown to meet the dual criteria of (i) resulting in statistically consistent discrimination measures and thus (ii) being representative of the complex POLSAR data.

B. POLSAR Discrimination Measures

Commonly used measures of dissimilarity for matrices are either the Euclidean or Manhattan distances, defined respectively as:

$$d(C_x, C_y) = \sum_{i,j} |\mathbb{R}(C_x - C_y)_{i,j}| + \sum_{i,j} |\mathbb{I}(C_x - C_y)_{i,j}| \quad (1)$$

$$d(C_x, C_y) = \sqrt{\sum_{i,j} |C_x - C_y|^2_{i,j}} \quad (2)$$

where $C_{i,j}$ denotes the (i, j) elements of the POLSAR covariance matrix C, || denotes absolute values and $\mathbb{R}, \mathbb{I}$ denote the real and imaginary parts respectively. However, in the context of POLSAR, these dissimilarity measures are not widely used mainly due to the multiplicative nature of the noisy data.

In the field of POLSAR, the Wishart distance is probably the most widely used, as part of the well-known Wishart classifier [13]. It is defined [8] as:

$$d(C_x, C_y) = \ln |C_y| + tr(C_x C_y^{-1}) \quad (3)$$

where $tr(C)$ denotes the trace of the matrix C. As a measure of distance, its main disadvantage is that $d(C_y, C_y) = \ln |C_y| \neq 0$.

Recent works have suggested alternative dissimilarity measures including the symmetric and asymmetric refined Wishart distance [9],

$$d(C_x, C_y) = \frac{1}{2} tr(C_x^{-1} C_y + C_y^{-1} C_x) - d \quad (4)$$

$$d(C_x, C_y) = \ln |C_x| - \ln |C_y| + tr(C_x C_y^{-1}) - d \quad (5)$$

the Bartlett distance [10],

$$d(C_x, C_y) = 2 \ln |C_{x+y}| - \ln |C_x| - \ln |C_y| - 2d \ln 2 \quad (6)$$

the Bhattacharyya distance [11],

$$r(C_x, C_y) = \frac{|C_x|^{1/2} |C_y|^{1/2}}{|(C_x + C_y)/2|} \quad (7)$$

and the Wishart Statistical test distance [12]

$$d(C_x, C_y) = (L_x + L_y) \ln |C| - L_x \ln |C_x| - L_y \ln |C_y| \quad (8)$$

Closer examination of these dissimilarity measures reveals that most of them are related in some ways. The Bhattacharyya distance is easily shown to be related to the Barlett distance. At the same time the Barlett distance can be considered a special case of the Wishart Statistical Test distance, when the two data sets have the same number of looks, i.e. $L_x = L_y$. The close relation among the measures is further supported by the fact that all of their publications refered the same statistical model developed in [1] as the foundation. In [1], to determine if the two scaled multi-look POLSAR covariance matrixes Z_x and Z_y, which have L_x and L_y as the corresponding number of looks, come from the same underlying stochastic process, the likelihood ratio statistics for POLSAR covariance matrix is considered:

$$Q = \frac{(L_x + L_y)^{d \cdot (L_x + L_y)}}{L_x^{d \cdot L_x} L_y^{d \cdot L_y}} \frac{|Z_x|^{L_x} |Z_y|^{L_y}}{|Z_x + Z_y|^{(L_x + L_y)}}$$

Taking the log-transformation of the above equation, and denoting $C_{vx} = Z_x/L_x$, $C_{vy} = Z_y/L_y$ and $C_{vxy} = (Z_x + Z_y)/(L_x + L_y)$ then:

$$Q = \frac{|C_{vx}|^{L_x} \cdot |C_{vy}|^{L_y}}{|C_{vxy}|^{L_x+L_y}} \tag{9}$$

$$\ln Q = L_x \ln|C_{vx}| + L_y \ln|C_{vy}| - (L_x + L_y)\ln|C_{vxy}| \tag{10}$$

To detect changes, a test statistic is developed for this discrimination measure. This means a distribution is to be derived for the dissimilarity measure. However, in the original work [1], only an *asymptotic* distribution was used. By contrast, this paper will propose a statistical model for the determinant of the POLSAR covariance matrix $|C_v|$ which is capable of providing an *exact* distribution for the test.

III. The Generic Scalar Statistical Model for POLSAR

In this section, after the basic foundations of POLSAR statistical analysis are introduced, the generic scalar statistical model for the multi-dimensional POLSAR will be presented.

Note that the POLSAR scattering vector is denoted as s in this paper. For partial polarimetric SAR (single polarization in transmit and dual polarization in receive), the vector is two-dimensional ($d = 2$) and is normally written as:

$$s_{part} = \begin{bmatrix} S_h \\ S_v \end{bmatrix} \tag{11}$$

For full and monostatic POLSAR data, the vector is three-dimensional ($d = 3$) and is presented as:

$$s_{full} = \begin{bmatrix} S_{hh} \\ \sqrt{2}S_{hv} \\ S_{vv} \end{bmatrix} \tag{12}$$

Let $\Sigma = E[ss^{*T}]$ denote the population expected value of the POLSAR covariance matrix, where s^{*T} is the complex conjugate transpose of s. Assuming s is jointly circular complex Gaussian with the expected covariance matrix Σ, then the probability density function (PDF) of s can be written as:

$$pdf(s; \Sigma) = \frac{1}{\pi^d|\Sigma|} e^{-s^{*T}\Sigma^{-1}s} \tag{13}$$

where $||$ denotes the matrix determinant.

The sample POLSAR covariance matrix is formed as the mean of the Hermitian outer product of independent single-look scattering vectors,

$$C_v = \langle ss^{*T} \rangle = \frac{1}{L}\sum_{i=1}^{L} s_i s_i^{*T} \tag{14}$$

where L is the number of looks and s_i denotes the single-look scattering vector, which equals s_{part} for the case of partial POLSAR or s_{full} for the case of full polarimetry.

Complex Wishart distribution statistics are normally used for the scaled covariance matrix $Z = LC_v$, whose PDF is given as:

$$pdf(Z; d, \Sigma, L) = \frac{|Z|^{L-d}}{|\Sigma^L|\Gamma_d(L)} e^{-tr(\Sigma^{-1}Z)} \tag{15}$$

with $\Gamma_d(L) = \pi^{d(d-1)/2}\prod_{i=0}^{d-1}\Gamma(L-i)$ and d is the dimensional number of the POLSAR covariance matrix.

The approach taken in this paper differs by applying the homoskedastic log transformation on a less-than-well-known result for the determinant of the covariance matrix. Goodman [14] proved that the ratio between the observable and expected values of the sample covariance matrix determinants behaves like a product of d chi-squared random variables with different degrees of freedom:

$$\chi_L^d = (2L)^d \frac{|C_v|}{|\Sigma_v|} \sim \prod_{i=0}^{d-1}\chi(2L-2i) \tag{16}$$

This result is used here to develop the generic scalar statistical model for POLSAR. From Eqn. 16 we have:

$$|C_v| \sim |\Sigma_v| \cdot \frac{1}{(2L)^d} \cdot \prod_{i=0}^{d-1}\chi(2L-2i) \tag{17}$$

In a given homogeneous POLSAR area, the parameters Σ_v, d and L can be considered constant. Thus Eqn. 17 indicates that in the original POLSAR domain, a multiplicative speckle noise pattern is present.

Moreover, since the average and variance of these chi-squared distributions are known to be constant, i.e. $avg[\chi(2L)] = 2L$ and $var[\chi(2L)] = 4L$, the product and summation of these random variables also have fixed summary statistics. Specifically:

$$avg\left[\prod_{i=0}^{d-1}\chi(2L-2i)\right] = 2^d \cdot \prod_{i=0}^{d-1}(L-i),$$

$$var\left[\prod_{i=0}^{d-1}\chi(2L-2i)\right] = \prod_{i=0}^{d-1}4(L-i)(L-i+1) - \prod_{i=0}^{d-1}4(L-i)^2,$$

Combining these results with Eqn. 17 , we have:

$$avg[|C_v|] = \frac{|\Sigma_v|}{L^d}\prod_{i=0}^{d-1}(L-i) \tag{18}$$

$$var[|C_v|] = \frac{|\Sigma_v|^2\left[\prod_{i=0}^{d-1}(L-i)(L-i+1) - \prod_{i=0}^{d-1}(L-i)^2\right]}{L^{2d}} \tag{19}$$

For a real world captured image, while the parameters d and L do not change for the whole image, the underlying Σ_v is expected to differ from one region to the next. Thus over a heterogeneous scene, the stochastic process for $|C_v|$ and $\ln|C_v|$ vary depending on the underlying signal Σ_v. In such context, Eqn. 19 implies that the variance of $|C_v|$ also differs depending on the underlying signal Σ_v (i.e. it is heteroskedastic).

Similar to the way intensity-ratio is proposed as the discrimination measure for the multiplicative and heteroskedastic SAR intensity [15], in this paper, the determinant-ratio and the change-ratio are also proposed as a discrimination measure for the POLSAR data: Which is shown above to also suffer from the multiplicative and heteroskedastic phenomena.

For cases where the true value of the underlying signal Σ_v is known *a priori*, then the determinant-ratio of the signal random variable ($\mathbb{R}_\Sigma$) is defined as:

$$\mathbb{R}_\Sigma = \frac{|C_v|}{|\Sigma_v|} \tag{20}$$

For cases where the POLSAR data is known to have come from a homogeneous area, but the true value of the underlying signal Σ_v is *unknown*, then a random variable called the change-ratio ($\mathbb{R}_C$) is defined as:

$$\mathbb{R}_C = \frac{|C_1|}{|C_2|} \tag{21}$$

where C_1 and C_2 are samples of the covariance matrix determinant in an assumed homogeneous area.

Using the results from Eqn. 17, we have

$$\mathbb{R}_\Sigma \sim \frac{1}{(2L)^d} \cdot \prod_{i=0}^{d-1} \chi(2L-2i) \tag{22}$$

$$\mathbb{R}_C \sim \prod_{i=0}^{d-1} \frac{\chi(2L-2i)}{\chi(2L-2i)} \tag{23}$$

Since the elementary components follow fixed distributions (i.e. $\chi^2(2L)$), it is natural that this variable also follows a fixed distribution. Moreover, it is independent of the underlying signal Σ_v, indicating its statistically consistent properties and hence its applicability as a POLSAR discrimination measure. This claim is extrapolated from the widely-used intensity-ratio as a SAR discrimination measure and will be further discussed in the next sections.

IV. SAR as a one-dimensional case of POLSAR

This section shows that the proposed generic model is equally applicable to the 1-dimensional ($d = 1$) case, which is physically equivalent to collapsing the multi-dimensional POLSAR dataset into conventional single dimensional SAR data. Mathematically, the sample covariance matrix C_v is reduced to the sample variance while the determinant $|C_v|$ is equivalent to the scalar variance value. As variance is equal to intensity I for SAR data, our result is consistent with previous results for SAR intensity data. Thus the proposed generic model for POLSAR, collapsed into one dimension will be shown to incorporate the traditional SAR intensity as a natural case of the unified model.

The results for our models can be summarized using the following equations:

$$|C_v| \sim |\Sigma_v| \cdot \frac{1}{(2L)^d} \cdot \prod_{i=0}^{d-1} \chi(2L-2i)$$

$$\mathbb{R}_\Sigma = \frac{|C_v|}{|\Sigma_v|} \sim \frac{1}{(2L)^d} \prod_{i=0}^{d-1} \chi(2L-2i)$$

$$\mathbb{R}_C = \frac{C_1}{C_2} \sim \prod_{i=0}^{d-1} \frac{\chi(2L-2i)}{\chi(2L-2i)}$$

Upon setting $d = 1$ into the above equations, the equations become:

$$|C_v| \sim \frac{|\Sigma_v|}{(2L)} \cdot \chi(2L)$$

$$\mathbb{R}_\Sigma = \frac{|C_v|}{|\Sigma_v|} \sim \frac{1}{(2L)} \cdot \chi(2L)$$

$$\mathbb{R}_C = \frac{C_1}{C_2} \sim \prod_{i=0}^{d-1} \frac{\chi(2L-2i)}{\chi(2L-2i)}$$

Since the PDF of chi-squared distribution can be written as:

$$\chi(2L) \sim pdf\left[\frac{\chi^{L-1}e^{-\chi/2}}{2^L\Gamma(L)}\right]$$

Applying the variable change theorem into the above equations results in:

$$|C_v| \sim pdf\left[\frac{L^L x^{L-1} e^{-Lx/|\Sigma_v|}}{\Gamma(L)|\Sigma_v|^L}\right]$$

$$\mathbb{R}_\Sigma = \frac{|C_v|}{|\Sigma_v|} \sim pdf\left[\frac{L^L x^{L-1} e^{-Lx}}{\Gamma(L)}\right]$$

$$\mathbb{R}_C = \frac{C_1}{C_2} \sim pdf\left[\frac{\Gamma(2L-1)x^{L-1}}{\Gamma^2(L-1)(1+x)^{2L}}\right]$$

These equations match exactly with the following traditional model for multi-look SAR intensity:

$$I \sim pdf\left[\frac{L^L x^{L-1} e^{-Lx/\bar{I}}}{\Gamma(L)\bar{I}^L}\right] \tag{24}$$

$$\mathbb{R}_{\bar{I}} = \frac{I}{\bar{I}} \sim pdf\left[\frac{L^L x^{L-1} e^{-Lx}}{\Gamma(L)}\right] \tag{25}$$

$$\mathbb{R}_I = \frac{I_1}{I_2} \sim pdf\left[\frac{\Gamma(2L-1)x^{L-1}}{\Gamma^2(L-1)(1+x)^{2L}}\right] \tag{26}$$

considering that $|C_v| \mapsto I$ and $|\Sigma_v| \mapsto \bar{I}$ when multi-dimensional POLSAR collapses into single-dimensional SAR.

V. Unifying the different discrimination measures proposals for both SAR and POLSAR

Statistical models have always been the foundation for deriving discrimination measures. This is true for both fields of SAR and POLSAR. For SAR, where the field is more or less settled on the issue, the statistical model for SAR intensity has been used to derive the most widely used intensity-ratio discrimination measure [15]. For POLSAR, where the research field is much less mature, the same case should apply, except that so far only asymptotic distributions have been derived for the most common foundation, i.e. the likelihood test statistics [1].

With the insight gained from section IV, this section presents a few results. First, similar to the way that the statistical model for SAR intensity has been used to provide the foundation for SAR discrimination measures, e.g. intensity-ratio [15], the statistical model proposed for POLSAR covariance matrix determinant is to be reviewed as providing the foundation for POLSAR discrimination measures, i.e. the likelihood test statistics. And second, new discrimination measures for POLSAR may be derived by learning from the existing SAR discrimination measures.

As for the first matter, in view of the models given in Eqn 17, the likelihood test statistics presented in [1] and rewritten in Eqns 9 & 10 can be expressed as:

$$\ln Q \sim k + L_x\Lambda^d_{L_x} + L_y\Lambda^d_{L_y} - (L_x+L_y)\Lambda^d_{(L_x+L_y)}$$

$$Q \sim e^k \frac{(\chi^d_{L_x})^{L_x} \cdot (\chi^d_{L_y})^{L_y}}{(\chi^d_{L_x+L_y})^{L_x+L_y}}$$

where $k = d\left[(L_x+L_y)\ln(L_x+L_y) - L_x \ln L_x - L_y \ln L_y\right]$. This, in essence, derives an exact statistical distribution for the likelihood test statistics (as opposed to an asymptotic distribution [1]).

As a by-product of this exact derivation, several discrimination measures for the common case of $L_x = L_y$ are further proposed. They are the determinant-ratio and the change-ratio presented in Section III. Compared to existing discrimination measures for POLSAR reviewed in Section II, the proposed dissimilarity measures are simpler in both concept and computation. They, at the same time, are multi-dimensional extensions of the intensity-ratio discrimination measure widely used in SAR.

In short, with the newly gained insight, the approach proposed in this paper provides a new bridge between the two fields of SAR and POLSAR. This invites the adaptation of many established SAR data processing techniques towards POLSAR data, where Section VII will describe an example demonstration. However, before further discussion ensues, it is necessary to validate the presented theoretical model with real-life practical data.

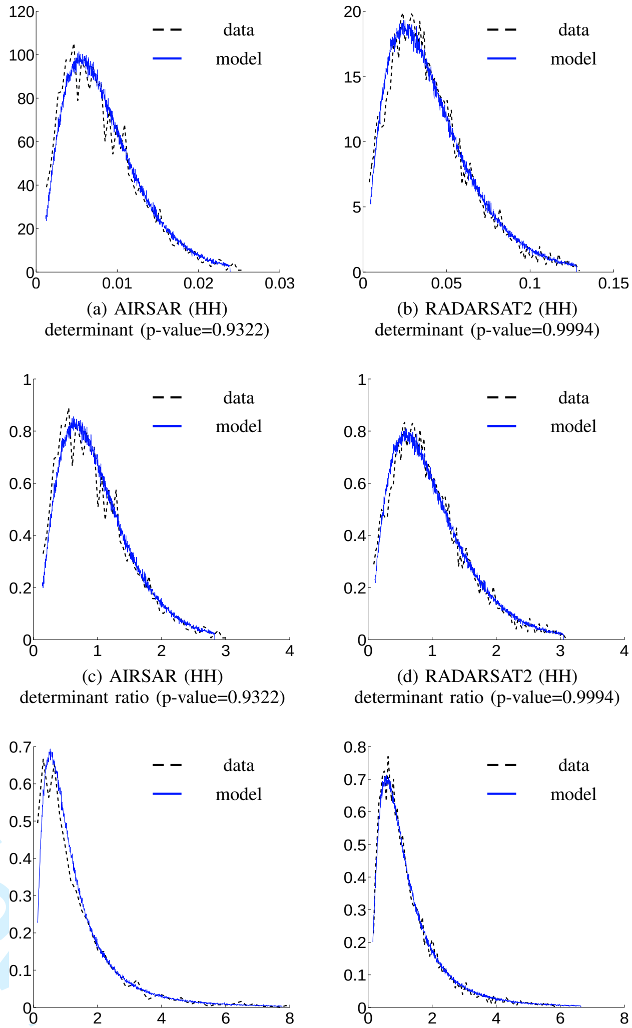
VI. Validating the Proposed Models Against Real-Life Data

This section aims to verify the theoretical models given in Eqns. 17, 22 and 23 against practical data. All of these models require the estimation of two parameters from the captured data. These are (i) the dimensional number d and (ii) the look number L. The dimensional number d is related to the type of (POL)SAR data captured, with $d = 1, 2, 3$ corresponding to the cases of SAR, partial and full polarimetric SAR data, respectively. The look-number L can be taken either as nominally stated by the data provider, or it can be estimated using the technique proposed by [16].

To show the robustness of the proposed models, their validations are also put to test on different types of POLSAR sensors. Specifically in this paper, the air-borne four-look ($L = 4$) AIRSAR image of Flevoland and a fine-quad single-look complex RADARSAT2 image ($L = 1$) are used. Since the determinant of the covariance matrix is only significant on multi-look data, nine-look processing is first applied on the single-look RADARSAT2 data ($L = 9$).

A. The Traditional case of SAR ($d = 1$)

Since it has been shown in Section IV that the model for the case of $d = 1$ matches exactly with the traditional model for SAR intensity, the validation of the proposed $d = 1$ model is quite straightforward. Fig. 1 presents the results of a test performed for the stated purpose. In the test, the intensity of single-channel SAR data (HH) for sample homogeneous areas in both datasets is extracted. The histograms for both the intensity and intensity ratio are then plotted against the theoretical PDF given above. The plots are obtained having set ENL to an estimated value obtained using the technique presented in [16]. In all cases, the good match between the actual data and model distribution lends support to the proposed model.



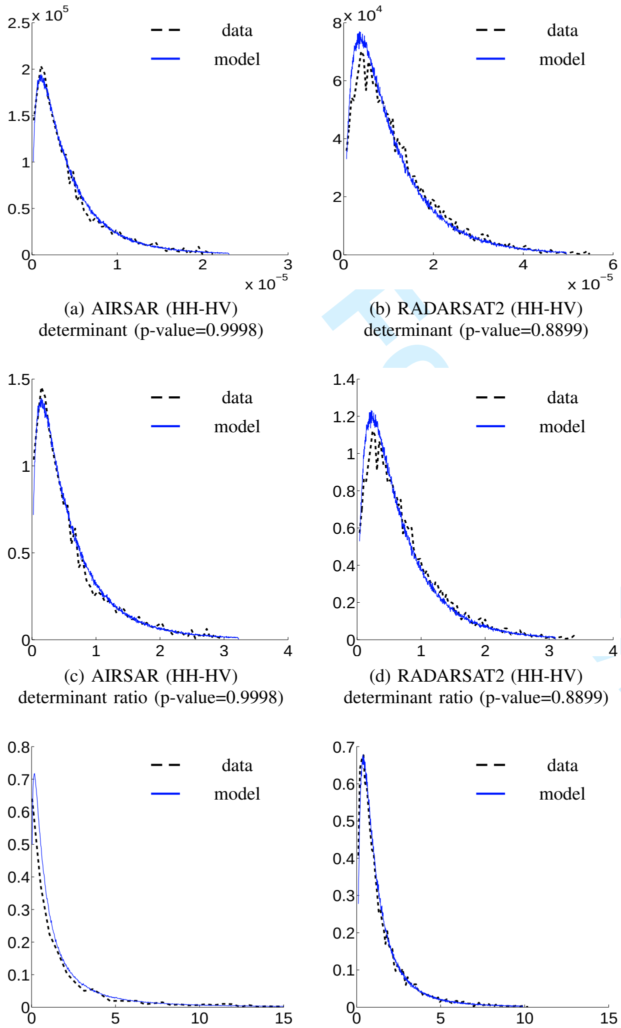
(a) AIRSAR (HH) determinant (p-value=0.9322)
(b) RADARSAT2 (HH) determinant (p-value=0.9994)
(c) AIRSAR (HH) determinant ratio (p-value=0.9322)
(d) RADARSAT2 (HH) determinant ratio (p-value=0.9994)
(e) AIRSAR (HH) change ratio (p-value=0.8425)
(f) RADARSAT2 (HH) change ratio (p-value=0.9923)

Fig. 1: The specific d=1 models are validated for both RADARSAT2 and AIRSAR datasets. Each histogram plots the signal under investigation (along the x-axis) against its probabilistic distribution across the patch (on the y-axis), with p-value being the significance level of the Kolmogorov-Smirnov goodness-of-fit test between the two histograms

B. The Multi-dimensional case of POLSAR ($d = 2, 3$)

The remainder of this section now focuses on validating the models for both partial ($d = 2$) and full ($d = 3$) POLSAR. Similar tests are carried out for both types of polarimetric SAR on both AIRSAR and RADARSAT2 datasets. The look-number is estimated for each type of data and for each image [16]. The practical and predicted distributions are again plotted to determine whether their visual match is close in each case. Fig. 2 shows the $d = 2$ plots for determinant, determinant ratio and change ratio for both datasets. It is very clear that the model histogram matches the data well, in fact appearing similar to a smoothed data response. Evidently the match is good for $d = 2$.

Similarly, Fig. 3 explores the $d = 3$ case for the same data, same model parameters. Again, although the histogram is much tighter, the match is visually obvious in all cases. This qualitative observation is further supported by high significance level (or p-value) in the standard Kolmogorov-Smirnov goodness-of-fit test between curves [17], noted under each graph. While a good p-value may not lead to automatic acceptance of the model, consistently high p-values lend a high level of confidence to the proposal.



(a) AIRSAR (HH-HV) determinant (p-value=0.9998)

(b) RADARSAT2 (HH-HV) determinant (p-value=0.8899)

(c) AIRSAR (HH-HV) determinant ratio (p-value=0.9998)

(d) RADARSAT2 (HH-HV) determinant ratio (p-value=0.8899)

(e) AIRSAR (HH-HV) change ratio (p-value=0.9182)

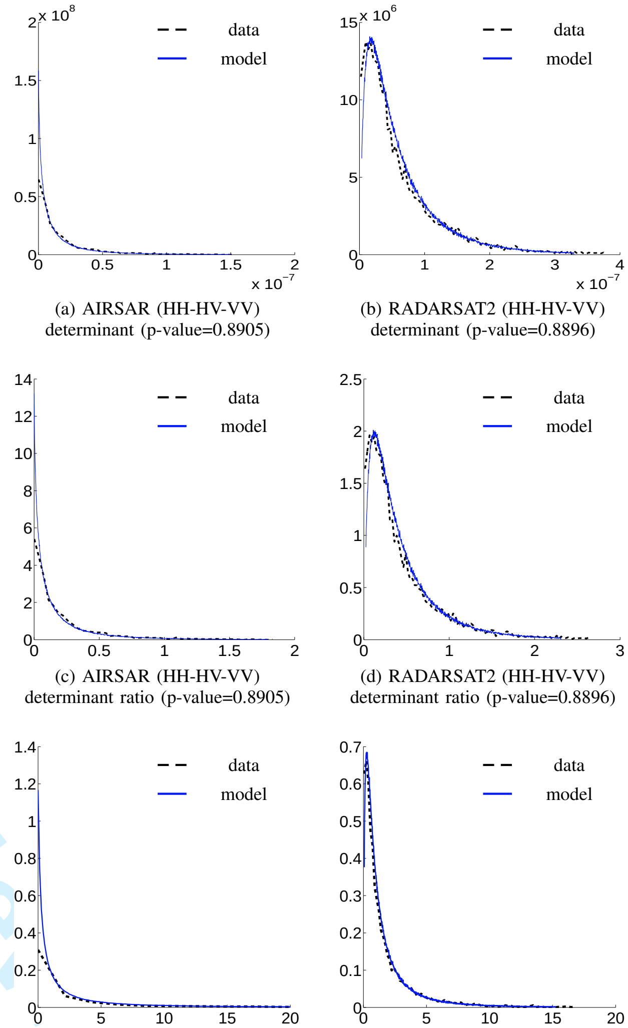
(f) RADARSAT2 (HH-HV) change ratio (p-value=0.9824)

Fig. 2: The specific $d = 2$ models are validated on both RADARSAT2 and AIRSAR datasets. Each histogram plots the signal under investigation (along the x-axis) against its probabilistic distribution across the patch (on the y-axis), with p-value being the significance level of the Kolmogorov-Smirnov goodness-of-fit test between the two histograms.

VII. Extending SAR data processing techniques towards POLSAR: A Short Example

One of the main beneficial implications for this proposal is that it allows the extension of many of existing SAR data processing technique toward POLSAR. One example has already been presented in this paper, where existing SAR discrimination measures are extended towards POLSAR. This section illustrates another example for a common technique in SAR, namely speckle filter evaluation.

The intensity ratio is routinely used in SAR to evaluate speckle filters. Specifically, the ratio of the filtered output to



(a) AIRSAR (HH-HV-VV) determinant (p-value=0.8905)

(b) RADARSAT2 (HH-HV-VV) determinant (p-value=0.8896)

(c) AIRSAR (HH-HV-VV) determinant ratio (p-value=0.8905)

(d) RADARSAT2 (HH-HV-VV) determinant ratio (p-value=0.8896)

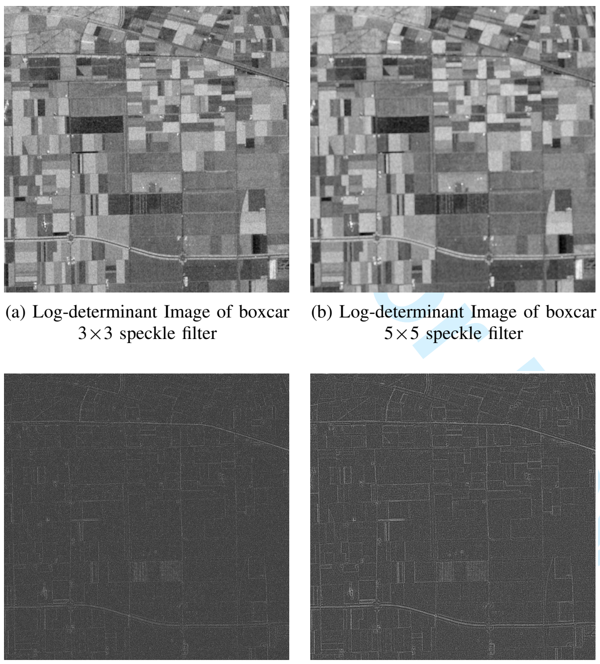
(e) AIRSAR (HH-HV-VV) change ratio (p-value=0.9254)

(f) RADARSAT2 (HH-HV-VV) change ratio (p-value=0.8883)

Fig. 3: The specific $d = 3$ models are validated on both RADARSAT2 and AIRSAR datasets. Each histogram plots the signal under investigation (along the x-axis) against its probabilistic distribution across the patch (on the y-axis), with p-value being the significance level of the Kolmogorov-Smirnov goodness-of-fit test between the two histograms.

the noisy input image represents the amount of noise removed. The multiplicative ratio, however, is not well suited to digital image presentation, which is linear and additive in nature [18]. Thus logarithmic transformation is applied to convert the multiplicative intensity-ratio into linear subtractive residuals. Apparently if the speckle filter is perfect, the ratio or residual images should only contain random noise. Thus the evaluation method is simply confirming that the speckle filter has not removed any image features.

Since the POLSAR determinant ratio has been shown to be equivalent to the SAR intensity ratio, this technique can be equally applied to evaluate POLSAR speckle filters. To demonstrate this feature, the technique is used to evaluate the performance of 3×3 and 5×5 boxcar POLSAR filters[1]. A square 700×700 pixel patch is first extracted from the AIRSAR Flevoland data (partial polarimetry HH-HV), and the two POLSAR speckle filters are then applied to the patch. The log-determinant images of the filtered outputs are displayed in Fig. 4. At the same time, the residual is computed for both filters, and the original image patches are also displayed in the same figure.



(a) Log-determinant Image of boxcar 3×3 speckle filter

(b) Log-determinant Image of boxcar 5×5 speckle filter

(c) Image of Log-determinant Residual for 3×3 filter

(d) Image of Log-determinant Residual for 5×5 filter

Fig. 4: Visual evaluation of POLSAR boxcar 3×3 vs. 5×5 speckle filters on AIRSAR Flevoland part-pol data (HH-HV).

From a visual examination of Fig 4, the evaluation technique also appears workable for POLSAR speckle filters, since the residual images clearly reveal the distortion impact of both POLSAR speckle filters.

Many similar SAR data processing techniques may be adaptable to POLSAR in this way, however they are not explored here due to space constraints.

VIII. Conclusion

This paper has proposed a generic statistical model for the POLSAR covariance matrix determinant based on the complex Wishart POLSAR target vector distribution. The model has been validated for specific $d = 2$ and $d = 3$ cases using real-life partial and full POLSAR data. The paper also establishes two new POLSAR discrimination measures: the determinant-ratio and the change-ratio, which essentially are the generic versions of the SAR intensity-ratio. We have also shown an excellent match between the specific $d = 1$ model and that for traditional SAR intensity, which effectively brings the existing SAR theories under the umbrella of this new model.

The main emphasis of this paper is to consider the POLSAR covariance matrix determinant, and its derivatives, as scalar and representative observables for the multi-dimensional POLSAR data. Compared to other published scalar observables, the determinant is highly representative of the multi-dimensional data. Its representative power is justified for the following reasons. First, the covariance matrix determinant, when collapsed into 1-D, transforms neatly into the representative SAR intensity. Second, the statistical model for this observable is shown to be the generic multi-dimensional extension of the traditional model for the one dimensional SAR intensity. Third, this observable also leads to the establishment of new discrimination measures for POLSAR data. And finally, similar to eigen-values, the determinant of the POLSAR covariance matrix is invariant to any changes in polarization basis.

The detailed and mathematical explanation linking the SAR intensity and the POLSAR determinant has been presented in Section IV but consider another high-level and intuitive linkage: The POLSAR scalar determinant can be used to represent and compare the multi-dimensional data, much in the same way that the scalar magnitude is used to represent and compare complex numbers. Similar to the determinant, the magnitude observable is also scalar and is not lossless. Besides the magnitude, another observable would be needed to fully represent a complex number. For example, in a polar representation, the angle would be required in addition to magnitude. Despite this, magnitude is still widely used when comparing two complex numbers, as it can be shown to be invariant to a change in the reference frame. In the same way, the determinant of the POLSAR covariance matrix, also scalar and invariant to a change of polarization basis, is useful despite not being *fully* representative of the POLSAR data. Thus this scalar observable is proposed to represent and distinguish between multi-dimensional datasets.

There are several beneficial implications of the proposed approach presented in this paper. Currently, while the field of SAR is much more developed than the field of POLSAR, the two fields remain quite distinct. Hence many techniques applicable to SAR have not been directly extended to POLSAR. The approach of this paper, by providing a scalar and representative model for POLSAR which generalizes the traditional model for SAR intensity, enables the convenient adaptation of many existing SAR data processing techniques for used with POLSAR data. In fact, two example applications of the technique were presented in this paper. In the first instance, existing SAR discrimination measures are extended towards POLSAR. While in the second case, existing technique to evaluate SAR speckle filters are shown to be extensible towards POLSAR. At the same time, this provides a consistent theory to unify the seemingly disparate discrimination measure proposals for SAR and POLSAR data.

Of course, it must be mentioned that the models proposed in this paper also have their own limitations. Firstly, they are based on the complex Wishart distribution which is only guaranteed to work for homogeneous areas. Secondly, while

[1]While boxcar filtering may not be the most modern POLSAR speckle filter, it is chosen here because the focus of the experiment is not to find the best POLSAR speckle filter, but rather to serve as an easily reproducible example of how an existing technique in SAR data processing can be simply adapted for use with POLSAR data.

it is desirable for a large number of applications to reduce the multi-dimensional POLSAR data to a scalar value, such a reduction is unlikely to be lossless in the general case.

In this paper, the theoretical model has only been validated for partial ($d = 2$) and full ($d = 3$) monostatic POLSAR data, leaving the test of its validity for other datasets such as bistatic POLSAR data ($d = 4$) or interferometric POLSAR data ($d = 6$) as possible future work. Other possible extensions include investigating the applicability of the observables into POLSAR heterogeneous areas, as well as exploring the use of these generic models in conjunction with other target decomposition techniques such as the Freeman-Durden decomposition [19] or the entropy/anisotropy decomposition [20].

References

[1] K. Conradsen, 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 Transactions on Geoscience and Remote Sensing*, vol. 41, no. 1, pp. 4 – 19, Jan 2003.

[2] V. Alberga, G. Satalino, and D. K. Staykova, "Comparison of polarimetric SAR observables in terms of classification performance," *International Journal of Remote Sensing*, vol. 29, no. 14, pp. 4129–4150, 2008. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/01431160701840182

[3] I. Joughin, D. Winebrenner, and D. Percival, "Probability density functions for multilook polarimetric signatures," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 32, no. 3, pp. 562 –574, may 1994.

[4] J.-S. Lee, K. Hoppel, S. Mango, and A. Miller, "Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 5, pp. 1017 –1028, Sep 1994.

[5] R. Touzi and A. Lopes, "Statistics of the stokes parameters and of the complex coherence parameters in one-look and multilook speckle fields," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 34, no. 2, pp. 519 –531, mar 1996.

[6] C. Lopez-Martinez and X. Fabregas, "Polarimetric SAR speckle noise model," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 41, no. 10, pp. 2232–2242, 2003.

[7] E. Erten, "The performance analysis based on SAR sample covariance matrix," *Sensors (Basel)*, vol. 12, no. 3, pp. 2766–2786, 2012.

[8] J. S. Lee, M. R. Grunes, and R. Kwok, "Classification of multi-look polarimetric SAR imagery based on complex Wishart distribution," *International Journal of Remote Sensing*, vol. 15, no. 11, pp. 2299–2311, 1994.

[9] S. N. Anfinsen, R. Jenssen, and T. Eltoft, "Spectral clustering of polarimetric SAR data with Wishart-derived distance measures," in *3rd International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry*, vol. 3, Jan 2007.

[10] P. Kersten, J. S. Lee, and T. Ainsworth, "Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and EM clustering," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 519 – 527, Mar 2005.

[11] K. Y. Lee and T. Bretschneider, "Derivation of separability measures based on central complex Gaussian and Wishart distributions," in *IEEE Internationaal Geoscience and Remote Sensing Symposium (IGARSS), 2011*, Jul 2011, pp. 3740 –3743.

[12] F. Cao, W. Hong, Y. Wu, and E. Pottier, "An unsupervised segmentation with an adaptive number of clusters using the SPAN/H/α/A space and the complex Wishart clustering for fully polarimetric SAR data analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 11, pp. 3454 –3467, Nov 2007.

[13] J.-S. Lee, M. Grunes, and G. de Grandi, "Polarimetric SAR speckle filtering and its implication for classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 5, pp. 2363 –2373, Sep 1999.

[14] N. R. Goodman, "The distribution of the determinant of a complex Wishart distributed matrix," *Annals of Mathematical Statistics*, vol. 34, no. 1, pp. 178–180, 1963.

[15] E. Rignot and J. van Zyl, "Change detection techniques for ERS-1 SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 31, no. 4, pp. 896 –906, Jul 1993.

[16] S. Anfinsen, A. Doulgeris, and T. Eltoft, "Estimation of the equivalent number of looks in polarimetric synthetic aperture radar imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 11, pp. 3795 –3809, Nov 2009.

[17] A. Kolmogoroff, "Confidence limits for an unknown distribution function," *The Annals of Mathematical Statistics*, vol. 12, no. 4, pp. 461–463, 12 1941. [Online]. Available: http://dx.doi.org/10.1214/aoms/1177731684

[18] F. Medeiros, N. Mascarenhas, and L. Costa, "Evaluation of speckle noise MAP filtering algorithms applied to SAR images," *International Journal of Remote Sensing*, vol. 24, no. 24, pp. 5197–5218, Dec 2003.

[19] A. Freeman and S. Durden, "A three-component scattering model for polarimetric SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 3, pp. 963 –973, May 1998.

[20] S. Cloude and E. Pottier, "An entropy based classification scheme for land applications of polarimetric SAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 1, pp. 68 –78, Jan 1997.