On the Use of the $l_2$-Norm for Texture Analysis of Polarimetric SAR Data

Xinping Deng, and Carlos López-Martínez, Senior Member, IEEE

Abstract—In this paper, the use of the $l_2$-norm, or Span, of the scattering vectors is suggested to texture analysis of PolSAR data, with the benefits that neither we need to analyze the polarimetric channels separately, nor a filtering of the data is required to analyze the statistics. Based on the product model, the distribution of the $l_2$-norm is studied. Closed expressions of the probability density functions under the assumptions of several texture distributions are provided. To utilize the statistical properties of the $l_2$-norm, quantities including normalized moments and log-cumulants are derived, along with corresponding estimators and estimation variances. Results on both simulated and real SAR data show that the use of the statistics based on the $l_2$-norm brings advantages in several aspects with respect to the normalized intensity moments and matrix variate log-cumulants.

Index Terms—Polarimetric SAR, Texture Analysis, $l_2$-Norm, Span, Statistical Model, Log-Cumulant.

I. INTRODUCTION

Knowledge of the exact statistical properties of the signal plays an important role in the applications of Polarimetric Synthetic Aperture Radar (PolSAR) data, such as speckle filtering [1], segmentation [2], ground cover classification [3], [4], etc. Gaussian statistics for the radar return signals have been frequently assumed [5]–[7]. However, the analysis of real PolSAR images reveals that non-Gaussian models give a better representation of the data [8], [9]. In the last two decades, a considerable research effort has been dedicated to find accurate and efficient non-Gaussian models for PolSAR data [10]–[15].

As there are many models proposed in the literature, how to choose a proper one is a critical challenge. Some comparative statistics, as well as tools to visualize the fit of models to data, are required. There are three widely used approaches.

The most intuitive way is to calculate the histogram, or empirical Probability Density Function (PDF) of the PolSAR data, and then compare it with the theoretical distribution channel by channel [6], [11]. To quantify the difference, measures such as Kolmogorov-Smirnov (KS) distance and correlation coefficient can be employed [14], [16]. The KS distance is an estimate of the discrepancy between the empirical distribution function of the samples and the cumulative distribution function of the reference distribution, and the correlation coefficient provides a simple quantitative measure of the similarity between two distributions. Methods based on the empirical PDF, however, depend on the bin size of the histogram, and the comparison of PDFs is not effective.

The Method of Moments (MoM) is another important approach. Especially, the Normalized Intensity Moments (NIM) are used as a measure of heterogeneity in many works by scaling the mean value of the intensities to unity [8], [17]–[19]. One can calculate the NIMs of different orders from data samples, and compare them with those of a specific model to see if the model is proper. The comparison is accomplished for each polarimetric channel separately. Therefore, the method exploits only the intensities of the data, regardless of the correlation between polarimetric channels.

At last, the Mellin Kind Statistics (MKS, also known as log-cumulants) can be also used to examine the fit of a distribution to the SAR data [3], [14], [15], [20], [21]. They were first proposed by Nicolas [20] to analyze the compounded distributions used to model single-channel SAR data, and later was extended to the case of covariance matrix by Anfinsen et al. [21]. It is demonstrated that MKS are of great value for the analysis of PolSAR data, and that they can be used to derive estimators for distribution parameters with low bias and variance. However, to obtain the sample covariance matrices, a multilooking process of the scattering vectors is required, which may change the original statistics, especially when it comes to high spatial resolution data.

As it shows, there are some limitations of the mentioned approaches due to the multidimensionality of the PolSAR data. In this paper, the use of the $l_2$-norm [22], or Span, of the scattering vectors is suggested for texture analysis, with the benefits:

1) The polarimetric channels are not analyzed separately, and the correlations between polarimetric channels are considered.
2) No filter is required, which may average out the texture or give rise to mixture in certain scenes like urban areas.
3) No estimation of the Equivalent Number of Looks (ENL) is needed.
4) The Span is polarimetrically invariant, the statistics of the Span are consistent when the scattering vectors are expressed in different polarisation basis.

The idea is to map the multidimensional signal to a scalar using the $l_2$-norm, of which the statistics take into account both the intensities and the correlations between polarimetric channels. Results on both simulated data and real data show that this approach brings advantages in several aspects compared with the method of NIM or matrix variate log-cumulants.
The remainder is organized as follows. First of all, related knowledge is reviewed, including the product model [23], the NIM, as well as univariate log-cumulants and matrix variate ones. Then, the distribution of the $l_2$-norm is studied based on the product model. Statistics such as normalized moments and log-cumulants are computed. At last, experiments exploiting the $l_2$-norm are implemented on both simulated data and real SAR data. Conclusions are given at the end.

II. BACKGROUND

A. Product Model

Gaussian statistics generally provide a good fit to PolSAR data when the surface roughness is relatively low, the spatial resolution is moderate, and a large number of scatterers are present [6], [24]. The scattering vector $\mathbf{z}$, can be modeled as having a $d$-dimensional complex Gaussian distribution with zero mean, with PDF given by [5]

$$p(\mathbf{z}; \Sigma) = \frac{1}{\pi^{d/2} |\Sigma|} \exp(-\mathbf{z}^\dagger \Sigma^{-1} \mathbf{z}) \quad (1)$$

where $| \cdot |$ is the determinant operation, $(\cdot)^\dagger$ means the transpose conjugate. The distribution is totally determined by the covariance matrix $\Sigma = E\{\mathbf{z} \mathbf{z}^\dagger\}$, where $E\{\cdot\}$ denote the expectation operator. The thermal noise of PolSAR system is not considered in this model, as for most recent systems, the noise level is very low, and the thermal noise can be neglected compared with the received signal from scenes like agriculture, forest and urban areas [19]. The multidimensional Gaussian distribution belongs to the elliptical family, which consists of a wide range of distributions with characteristic functions of the quadratic form [25].

As the image resolution increases, other models from the elliptical family, such as product models [10], [12]–[14], Wishart-Kotz-type distributions [26], [27], and multivariate generalized Gaussian distributions [28] are reported to give a better description of the statistical behavior. It is now widely accepted that the randomness of the data is commonly due to two unrelated factors, texture and speckle. The texture models the natural spatial variation of the radar cross section, whereas the speckle, following a complex Gaussian distribution, represents the polarimetric information. The texture and speckle are incorporated with a product operation which leads to a doubly stochastic model called product model [23].

Under the assumption of product model, the scattering vector can be written as [23]:

$$\mathbf{k} = \sqrt{\tau} \mathbf{z} \quad (2)$$

where $\tau$ is the texture parameter, a positive random variable with unity mean, and $\mathbf{z}$ is the speckle vector characterized by (1). The product model is also called Spherically Invariant Random Vector (SIRV) model [4], [29]–[31], or scale mixture of Gaussian [13]. Assuming that the texture is independent from the speckle [23], the PDF of the scattering vector can be calculated using

$$p(\mathbf{k}; \Sigma, \theta) = \int_0^\infty \frac{1}{\pi^{d/2} |\Sigma|} \exp\left(-\frac{\mathbf{k}^\dagger \Sigma^{-1} \mathbf{k}}{\tau}\right) p(\tau; \theta) d\tau \quad (3)$$

where $\theta$ represents the parameter of the texture distribution. A number of models have been proposed in the literature by introducing different texture distributions, including the gamma distribution [9], [10], the inverse gamma distribution [11], [12], the Fisher distribution [2], [14], and the beta distribution [15], see Table I.

For speckle reduction, the scattering vectors are frequently multilook processed and represented by sample covariance matrices. Assume that the texture has a higher spatial correlation than the speckle, and the texture parameter is constant over the multilook processing window [10], then the sample covariance matrix can be expressed as

$$C_L = \frac{1}{L} \sum_{i=1}^L \mathbf{k}_i \mathbf{k}_i^\dagger = \frac{\tau}{L} \sum_{i=1}^L \mathbf{z}_i \mathbf{z}_i^\dagger. \quad (4)$$

Here $L$ denotes the number of looks. Since the pixels are usually correlated, $L$ is always smaller than the actual number of pixels, and it needs to be estimated from the data. The estimation, however, is difficult [32]. The conditional distribution of $C_L$ on $\tau$ is a Wishart distribution [10], therefore, the marginal PDF of $C_L$ is obtained as

$$p(C_L; L, \Sigma, \theta) = \int_0^\infty \frac{L^d |C_L|^{L-d}}{\Gamma_d(L) |\Sigma|^L} \exp\left(-\frac{L}{\tau} \text{Tr}(\Sigma^{-1} C_L)\right) p(\tau; \theta) d\tau \quad (5)$$

where $\text{Tr}(\cdot)$ refers to the matrix trace operation and $\Gamma_d(L)$ is a normalization factor given by

$$\Gamma_d(L) = \pi^{d(d-1)/2} \prod_{i=1}^d \Gamma(L - i + 1). \quad (6)$$

with $\Gamma(\cdot)$ representing the gamma function. To obtain covariance matrices over textured areas, care needs to be taken to ensure that the window size of the multilook processing must be smaller than the texture correlation length.

B. Normalized Intensity Moments

One way to quantitatively evaluate the non-Gaussian behavior of the data is to compute the NIM [8], [17]

$$\text{nim}_v \{I\} = \frac{m_v \{I\}}{m_1 \{I\}} \quad (7)$$

where $m_v \{I\}$ refers to the $v$th order moment of the intensity, and $m_1 \{I\}$ is the first power of the mean intensity. On the one hand, for Gaussian distributed scattering vectors, the intensity, denoted by $I_g$, follows an exponential distribution [33], and the NIM of the $v$th order is given by

$$\text{nim}_v \{I_g\} = \Gamma(v + 1). \quad (8)$$

The exponential distribution can be obtained from the gamma distribution by letting the shape parameter equal to 1, which is widely used to model the multilook intensity in SAR data [33]. On the other hand, over textured areas modeled by the product model (2), the NIM can be written as

$$\text{nim}_v \{I\} = \frac{m_v \{\tau\} m_v \{I_g\}}{m_1 \{\tau\} m_1 \{I_g\}} = \frac{m_v \{\tau\} \text{nim}_v \{I_g\}} {m_1 \{\tau\}} \quad (9)$$
because the texture component and speckle one are independent, and the mean value of the texture parameter is equal to 1. It can be shown that \( m_v(\tau) \geq 1 \) holds for all the texture distributions in Table I. And the \( \text{nim}_v \) of textured data is larger than that of Gaussian distributed data when \( v > 1 \).

The NIM of the second order is used in many works. It follows that the second order NIM of the exponential distribution is \( \text{nim}_2 \{ I_g \} = 2 \), and \( \text{nim}_2 \{ I \} > 2 \) for all texture models in Table I. As a result, we can determine whether the data can be modeled by a Gaussian distribution or not simply by comparing the \( \text{nim}_2 \{ I \} \) estimated from data with 2; if the discrepancy between the estimated value and 2 is very large, the data shows non-Gaussianity. The method of NIM requires to analyze different polarimetric channels separately.

### C. Univariate Log-Cumulants

It is demonstrated that the statistics derived from the Mellin transform can be employed to design estimators for the distribution parameters with low bias and variance [20], [21]. Let \( p(x) \) be a function defined on the positive real axis \( 0 < x < \infty \), the Mellin transform is the operation mapping the function \( p \) into the function \( \phi \) defined on the complex plane by the relation [34]

\[
\phi_x(s) = \int_0^\infty x^{s-1}p(x)dx. \tag{10}
\]

The function \( \phi_x(s) \) is also known as the Mellin kind characteristic function [21], which can be viewed as the \((s-1)\)th order moment of a positive random variable. Expanding \( x^{s-1} = e^{(s-1)\ln x} \) using the Taylor series [35], the Mellin Transform can be further written as a formal power series [36]

\[
\phi_x(s) = \sum_{v=0}^\infty \frac{(s-1)^v}{v!} \mu_v \{ x \} \] \tag{11}

where the coefficients are known as log-moments [20], [21]

\[
\mu_v \{ x \} = \int_0^\infty (\ln x)^vp(x)dx. \tag{12}
\]

Let \( \varphi_x(s) = \ln \phi_x(s) \), then \( \varphi_x(s) \) is called the Mellin kind characteristic function of the second kind [21], that can also be formulated as a formal power series according to the composition of formal power series [36]

\[
\varphi_x(s) = \sum_{v=0}^\infty \frac{(s-1)^v}{v!} \kappa_v \{ x \} \] \tag{13}

with coefficients as log-cumulants [20], [21]

\[
\kappa_v \{ x \} = \mu_v \{ x \} - \sum_{k=1}^{v-1} \binom{v-1}{k} \mu_k \{ x \} \kappa_{v-k} \{ x \}. \tag{14}
\]

Equation (14) is known as the combinatorial version of Faà di Bruno’s formula, which is denoted by the function \( g_v(\mu_1 \{ x \}, \cdots, \mu_v \{ x \}) \) in the remainder of this text. It shows that the log-cumulants are polynomials of the log-moments.

From (11) and (13), the log-moments and the log-cumulants can be calculated using

\[
\mu_v \{ x \} = \left. \frac{d^v}{ds^v} \phi_x(s) \right|_{s=1} \tag{15}
\]

\[
\kappa_v \{ x \} = \left. \frac{d^v}{ds^v} \varphi_x(s) \right|_{s=1} \tag{16}
\]

and with them, we can avoid analyzing the functions such as \( \phi_x(s) \) directly. Table I shows the log-cumulants of several well known texture distributions.

#### D. Matrix Variate Log-Cumulants

The univariate log-cumulant is extended to the matrix case by Anfinsen et al. to analyze the sample covariance matrices of PolSAR data [21]. Define the Mellin kind matrix variate characteristic function as

\[
\phi_X(s) = \int_{\Omega^+} |X|^{s-d} p(X) dX \tag{17}
\]

with \( d \) denoting the dimension of the sample covariance matrix and \( \Omega^+ \) the space of positive semi-definite Hermitian matrices, then, the \( v \)th order log-moment and log-cumulant are given by

\[
\mu_v \{ X \} = \left. \frac{d^v}{ds^v} \phi_X(s) \right|_{s=d} \tag{18}
\]

\[
\kappa_v \{ X \} = \left. \frac{d^v}{ds^v} \ln \phi_X(s) \right|_{s=d}. \tag{19}
\]

The properties of the univariate log-cumulants and the matrix variate log-cumulants are very similar. For the same distribution parameters, they are related by [21]

\[
\kappa_v \{ X \} = d^v \kappa_v \{ x \}. \tag{20}
\]

From independent and identically distributed samples \( \{ X_i, i = 1, \cdots, N \} \), the log-cumulants can be estimated using the Faà di Bruno formula, \( \hat{\kappa}_v \{ X \} = g_v(\hat{\mu}_1 \{ X \}, \cdots, \hat{\mu}_v \{ X \}) \)

where \( \hat{\mu}_v \{ X \} \) is the estimated log-moments

\[
\hat{\mu}_v \{ X \} = \frac{1}{N} \sum_{i=1}^N (\ln |X_i|)^v. \tag{21}
\]

The covariance of the estimated log-moments is given by [37]

\[
\text{Cov}(\hat{\mu}_v, \hat{\mu}_\nu) = \frac{1}{N} (\mu_{v+\nu} - \mu_v \mu_\nu) \tag{22}
\]

where \( \hat{\mu}_v \) can be either the univariate log-moments or the matrix variate ones. As the log-cumulants are the polynomials of the log-moments, therefore, once we know the covariances of the log-moments, the variance of the estimated log-cumulant can be calculated using [37]

\[
\text{Var} \{ \hat{\kappa}_v \} = \nabla g_v^T M_v \nabla g_v \tag{23}
\]

where \( M_v \) is the covariance matrix of the log-moments with entries \( [M_v]_{ij} = \text{Cov}(\hat{\mu}_i, \hat{\mu}_j) \), and \( \nabla g_v \) is the vector of the partial differentials of \( g_v \)

\[
\nabla g_v = \left[ \frac{\partial g_v}{\partial \mu_1}, \cdots, \frac{\partial g_v}{\partial \mu_v} \right]^T. \tag{24}
\]
It is necessary to note that the estimator of log-cumulants based on the Faa di Bruno’s formula is biased [38, 39].

To see if a statistical model is applicable, we can compare the theoretical log-cumulants calculated from the PDF ($\hat{I}_k$) and those estimated from the sample data ($\hat{I}_k$) [20, 21]. The comparison is illustrated in Figure 1, where the second order log-cumulants $\hat{I}_2$ and the third order log-cumulants $\hat{I}_3$ are employed. By varying the parameter values, different texture distributions shown in Table I take up certain parts in the plot, and they make up a complete coverage of the diagram [15].

![Fig. 1. A complete coverage of log-cumulant $\kappa_2$-$\kappa_3$ space. Theoretical log-cumulants calculated from texture distributions in Table I take up different parts in the diagram. From data with texture following a specific distribution, the statistics of the $I_k$ can be avoided. Nevertheless, the statistics of the $I_k$ are listed as follows.](image)

III. NORM STATISTICS

**A. Distribution**

The $l_2$-norm, denoted by $\|z\|_2 = \sqrt{\sum |z|^2}$, is a non-negative mapping from the $d$-dimensional complex space $\mathbb{C}^d$ to the set of non-negative real numbers $\mathbb{R}^+$ [22]. The square of the $l_2$-norm, which can be written as

$$I_k = k^\dagger k$$  

(24)

is mainly studied in this work, since the square root operation can be avoided. Nevertheless, the statistics of the $l_2$-norm can be obtained from those of $I_k$ by transformations. $I_k$ is also known as the Span, which can be interpreted as the total scattered power. It is polarimetrically invariant, independent from the choice of the polarization basis in which the scattering matrix is decomposed. Therefore, the statistics of the Span are consistent when the scattering vectors are expressed in different polarization basis. As a matter of fact, the Span is also employed in some other works [1], [31]. Assuming that the scattering vector can be modeled by the product model (2), the square of the $l_2$-norm can be further rewritten as

$$I_k = \tau I_z$$  

(25)

where $I_z = z^\dagger z$, independent from $\tau$, is the $l_2$-norm square of the speckle vector.

Based on the assumption that the speckle vector $z$ follows a multivariate complex Gaussian distribution with PDF as (1), it can be proved that the PDF of $I_z$ is given by (see Appendix A)

$$p(I_z) = \sum_{i=1}^d p_i \frac{1}{\lambda_i^{d-1}} \exp \left( -\frac{I_z}{\lambda_i} \right)$$  

(26)

where $\lambda_i, i = 1, \ldots, d$ are the eigenvalues of the covariance matrix $\Sigma$, and $p_i$ are the corresponding weights

$$p_i = \frac{\lambda_i^{d-1}}{\prod_{j=1, j \neq i}^d (\lambda_i - \lambda_j)}.$$  

(27)

This distribution is known as hyperexponential distribution [40], or mixed exponential distribution, which is a weighted sum of exponential distributions. The eigenvalues are assumed to be distinct. When there exist repeated eigenvalues, we have to substitute the corresponding exponential distribution for a gamma distribution. One can refer to [41] for more details. In this paper, the case where there are distinct eigenvalues is mainly discussed.

Provided that the distribution of the texture parameter $p(\tau; \theta)$ is known, the PDF of $I_k$ can be calculated by

$$p(I_k; \theta) = \sum_{i=1}^d \int p_i \frac{1}{\lambda_i^{d-1}} \exp \left( -\frac{I_k}{\lambda_i} \right) \frac{1}{\tau} p(\tau; \theta) d\tau.$$  

(28)

For instance, when the texture parameter is gamma, inverse gamma or Fisher distributed, the PDFs can be calculated as shown in Table II. The distributions given by (28) are univariate ones. It is much easier to use them for texture analysis than the distributions of the sample covariance matrix or scattering vector.

**B. Normalized Moments**

It is known that the moments of a mixture model is a weighted average of those of the mixing components [42]. The $v$th order moment of the speckle $I_z$, therefore, can be written as the combination of the $v$th order moments of exponential distributions [43]

$$m_v \{ I_z \} = \Gamma(v + 1) \sum_{i=1}^d p_i \lambda_i^v$$  

(29)

with $\lambda_i$ denoting the eigenvalues, and $p_i$ given by (27). Let $h_v(\lambda_1, \cdots, \lambda_d) = \sum_{i=1}^d p_i \lambda_i^v$, as shown in [44], it can be rewritten as

$$h_v(\lambda_1, \cdots, \lambda_d) = \sum_{1 \leq i_1 \leq \cdots \leq i_v \leq d} \lambda_{i_1} \cdots \lambda_{i_v}$$  

(30)

which is known as the Complete Homogeneous Symmetric Polynomials (CHSP) [45]. Various calculations become easier based on (30) as there are only multiplications and additions. For instance, the asymptotic limits when there are repeated eigenvalues can be easily computed. Some examples of the CHSP when $d = 3$ are listed as follows

$$h_1(\lambda_1, \lambda_2, \lambda_3) = \lambda_1 + \lambda_2 + \lambda_3$$  

(31)

$$h_2(\lambda_1, \lambda_2, \lambda_3) = \lambda_1^2 + \lambda_2^2 + \lambda_3^2 + \lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3$$  

(32)

$$h_3(\lambda_1, \lambda_2, \lambda_3) = \lambda_1^3 + \lambda_2^3 + \lambda_3^3 + \lambda_1 \lambda_2^2 + \lambda_1^2 \lambda_3 + \lambda_2^2 \lambda_1 + \lambda_2 \lambda_3^2 + \lambda_3 \lambda_2^3 + \lambda_1 \lambda_2 \lambda_3$$  

(33)

...
In the case of which is found to be the fraction of the moment of the and substitute them for the eigenvalues in (29), the normalized of different covariance matrices comparative, we normalize the 

\[ \psi(v-1)(\alpha) = - (\alpha - 1)^{v-1} \psi(v-1)(\alpha) \]

\[ + (\alpha - 1) \psi(v)(\alpha) \]

\[ + \delta(v-1) \ln \frac{\zeta}{\tau} \]

Notes: \( \psi(v)(z) \) is the polygamma function defined as \( \frac{d^{v+1}}{dz^{v+1}} \ln \Gamma(z) \), and \( \delta(\cdot) \) is the Dirac function.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>PDF of the ( l_2 )-Norm, ( p(I_k; \Sigma) )</th>
<th>Moments ( m_v(\tau) )</th>
<th>Log-Cumulants ( \kappa_v(\tau) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamma</td>
<td>( p(\tau; \alpha) = \frac{1}{\Gamma(\alpha)} \left( \frac{\alpha}{\tau} \right)^{\alpha \cdot \tau} \exp(-\alpha \tau) )</td>
<td>( \left( \frac{1}{\alpha} \right)^{\Gamma(\nu+\alpha) / \Gamma(\alpha)} )</td>
<td>( \psi(v-1)(\alpha) - (\nu(v-1) + \ln \alpha) )</td>
</tr>
<tr>
<td>Inverse Gamma</td>
<td>( p(\tau; \alpha) = \frac{1}{\Gamma(\alpha)} \left( \frac{\alpha}{\tau} \right)^{\alpha \cdot \tau} \exp(-\alpha \cdot \tau) )</td>
<td>( \frac{1}{\Gamma(\alpha)} \left( \frac{\alpha}{\tau} \right)^{\alpha \cdot \tau} \exp(-\alpha \cdot \tau) )</td>
<td>( \psi(v-1)(\alpha) )</td>
</tr>
<tr>
<td>Fisher</td>
<td>( p(\tau; \xi, \zeta) = \frac{\Gamma(\xi+1) \left( \xi + \zeta - 1 \right)}{\Gamma(\xi) \Gamma(\zeta + 1)} \left( \frac{\xi}{\zeta} \right)^{\xi \cdot \tau} \left( 1 - \frac{\xi}{\zeta} \right)^{\zeta \cdot \tau} )</td>
<td>( \frac{1}{\Gamma(\xi) \Gamma(\zeta + 1)} \left( \frac{\xi}{\zeta} \right)^{\xi \cdot \tau} \left( 1 - \frac{\xi}{\zeta} \right)^{\zeta \cdot \tau} )</td>
<td>( \psi(v)(\xi + (\nu-1)(\zeta + 1)) + (\nu(v-1)) \ln \frac{\xi}{\zeta} )</td>
</tr>
<tr>
<td>Beta</td>
<td>( p(\tau; \xi, \zeta) = \frac{1}{\Gamma(\xi+1) \Gamma(\zeta + 1)} \left( \frac{\xi}{\zeta} \right)^{\xi \cdot \tau} \left( 1 - \frac{\xi}{\zeta} \right)^{\zeta \cdot \tau} )</td>
<td>( \frac{1}{\Gamma(\xi) \Gamma(\zeta + 1)} \left( \frac{\xi}{\zeta} \right)^{\xi \cdot \tau} \left( 1 - \frac{\xi}{\zeta} \right)^{\zeta \cdot \tau} )</td>
<td>( \psi(v-1)(\xi + (\nu-1)(\zeta + 1)) + (\nu(v-1)) \ln \frac{\xi}{\zeta} )</td>
</tr>
<tr>
<td>Beta Prime</td>
<td>( p(\tau; \xi, \zeta) = \frac{1}{\Gamma(\xi+1) \Gamma(\zeta + 1)} \left( \frac{\xi}{\zeta} \right)^{\xi \cdot \tau} \left( 1 - \frac{\xi}{\zeta} \right)^{\zeta \cdot \tau} )</td>
<td>( \frac{1}{\Gamma(\xi) \Gamma(\zeta + 1)} \left( \frac{\xi}{\zeta} \right)^{\xi \cdot \tau} \left( 1 - \frac{\xi}{\zeta} \right)^{\zeta \cdot \tau} )</td>
<td>( \psi(v-1)(\xi + (\nu-1)(\zeta + 1)) + (\nu(v-1)) \ln \frac{\xi}{\zeta} )</td>
</tr>
</tbody>
</table>

Notes: \( K_\alpha \) is the modified Bessel function of the second kind of order \( \alpha \), \( U \) is the confluent hypergeometric function of the second kind (Kummer U function), \( W \) is the Whittaker W function, and \( M \) is the confluent hypergeometric function of the first kind (Kummer M function).

When the covariance matrix \( \Sigma \) is scaled by a factor, the moments expressed in (29) will change accordingly. To eliminate this effect and make the moments under conditions of different covariance matrices comparative, we normalize the eigenvalues by

\[ \tilde{\lambda}_i = \frac{\lambda_i}{\sum_{k=1}^{d} \lambda_k} \]

and substitute them for the eigenvalues in (29), the normalized moment of the \( v \)th order is obtained

\[ nm_v(I_k) = \Gamma(v + 1) b_v(\tilde{\lambda}_1, \cdots, \tilde{\lambda}_d) \]

which is found to be the fraction of the \( v \)th order moment and the \( v \)th power of the mean

\[ nm_v(I_k) = m_v(I_k) \]

In the case of \( d = 1 \), the normalized moments are reduced to the normal intensity moments of single channel data (7).

For Gaussian distributed speckle, the \( v \)th order normalized moment is not a fixed value, instead, it varies in a range. The value reveals the polarimetric information to some extent. Let \( \tilde{\lambda}_i = 1/d, i = 1, \cdots, d \), which means that the polarimetric scattering Entropy is equal to 1 [33], the smallest value of the normalized moment can be obtained from (30)

\[ \min\{nm_v(I_k)\} = \frac{\Gamma(v + d)}{d^v \Gamma(d)}. \]

When the polarimetric scattering Entropy is equal to 0 (\( \tilde{\lambda}_i = 0 \) and \( \lambda_j = 0, \forall j \neq i \), the normalized moments will reach the upper boundary

\[ \max\{nm_v(I_k)\} = \Gamma(v + 1). \]

Taking into account the texture part, the \( v \)th order moment of the \( l_2 \)-norm of the total scattering vector can be written as

\[ m_v(I_k) = m_v(\tau) m_v(I_k) \]

since the texture and the speckle are independent. Under the assumption that the mean value of texture parameter is equal to 1, the normalized moments of \( I_k \) are then

\[ nm_v(I_k) = \frac{m_v(I_k)}{m_v(I_{\kappa})} = m_v(\tau) nm_v(I_k). \]

Different distributions for the texture parameter will result into different values. As a result, the normalized moments could be employed to reveal the texture information, besides the polarimetric information of PolSAR data.

Given \( N \) independent and identically distributed samples, \( \{k_i, i = 1, \cdots, N\} \), the sample moments can be estimated using

\[ \tilde{m}_v(I_k) = \frac{1}{N} \sum_{i=1}^{N} (k_i)^v \]

and the normalized moments by

\[ \tilde{nm}_v(I_k) = \frac{\tilde{m}_v(I_k)}{\tilde{m}_1(I_k)}. \]

This estimator of the normalized moment is biased [46], [47]. As shown in Appendix B, the bias is of the order \( O(1/N) \)

\[ B = \frac{1}{2N} \left( v(v+1) \cdot nm_2(I_k) nm_v(I_k) - 2v \cdot nm_{v+1}(I_k) - v(v-1) \cdot m_v(I_k) \right). \]
When the sample size \( N \) is large, the mean value and the variance of \( \hat{n}_{m}\{I_k\} \) can be approximated by (Appendix B)
\[
E\{\hat{n}_{m}\{I_k\}\} \approx \frac{m_{v}\{I_k\}}{m^{(v)}_1\{I_k\}} = nm_v\{I_k\} \tag{44}
\]
\[
\text{Var}\{\hat{n}_{m}\{I_k\}\} \approx \frac{1}{N}[nm_{2v}\{I_k\} - (v - 1)^2 \cdot nm_v^2\{I_k\} - 2v \cdot nm_v\{I_k\}nm_{v+1}\{I_k\} + v^2 \cdot nm_2\{I_k\}nm_v^2\{I_k\}] \tag{45}
\]

\section*{C. Log-Cumulants}

As shown in the previous section, the moments can be viewed as functions of the order \( v \), and these functions are distinct for different distributions. Therefore, they can be employed to identify data models. Letting \( v = s - 1 \) and computing the logarithm on both sides of (39) gives the relation:
\[
\varphi_{I_k}(s) = \varphi_{\tau}(s) + \varphi_{I_k}(s) \tag{46}
\]
where \( \varphi_{\tau}(s) \) is the Mellin kind characteristic function of the second kind (Section II-C). The log-cumulant can be obtained by calculating the \( v \)th order differential and setting \( s = 1 \)
\[
\kappa_v\{I_k\} = \kappa_v\{\tau\} + \kappa_v\{I_k\}. \tag{47}
\]
It is an addition of the values from the texture part and the speckle part. The log-cumulants of the texture part are well studied in [20], [21], and the results of some widely studied distributions are listed in Table I. For the speckle part, it can be shown from (29) that the Mellin kind characteristic function of the second kind is
\[
\varphi_{I_k}(s) = \ln \Gamma(s) + \ln h_{s-1}(\lambda_1, \cdots, \lambda_d), \tag{48}
\]
and the log-cumulants can be further calculated by (Appendix C)
\[
\kappa_v\{I_k\} = \psi^{(v-1)}(1) + \beta_v \tag{49}
\]
where \( \psi^{(v)}(\cdot) \) is the polygamma function, and \( \beta_v \) is solved recursively using
\[
\beta_v = \hat{h}(v) - \sum_{k=1}^{v-1} \binom{v-1}{k} \hat{h}(k)\beta_{v-k} \tag{50}
\]
with \( \hat{h}(v) = \sum_{i=1}^{d} P_i(\ln \lambda_i)^v \). A great advantage of the log-cumulants for texture analysis is that the contribution of the texture and the speckle can be separated using the logarithm operation [21]. The log-cumulants of the \( l_2 \)-norm are calculated from the scattering vectors directly. As a result, no filtering is required to obtain the sample covariance matrices, where the procedure may change the statistical properties of the original data, especially when it comes to very high resolution data.

To test the fit of a model to PolSAR data using the log-cumulants, we can employ the biased estimator based on the Fià di Brunos formula as shown in Section II-D, or the unique unbiased estimator with minimum variance, k-statistics [38], [48]. Given \( N \) independent and identically distributed samples \( \{k_i, i = 1, \cdots, N\} \), the k-statistics of the first three orders are given by [48]
\[
\hat{k}_1\{I_k\} = \frac{S_1}{N} \tag{51}
\]
\[
\hat{k}_2\{I_k\} = \frac{NS_2 - S_1^2}{N(N-1)} \tag{52}
\]
\[
\hat{k}_3\{I_k\} = \frac{2S_1^3 - 3NS_1S_2 + N^2S_3}{N(N-1)(N-2)} \tag{53}
\]
where \( S_v \) is the sums of the \( v \)th powers
\[
S_v = \sum_{i=1}^{N} [\ln(k_i)]^v. \tag{54}
\]

The variance of the first three k-statistics are given by [48]
\[
\text{Var}\{\hat{k}_1\{I_k\}\} = \frac{\kappa_2\{I_k\}}{N} \tag{55}
\]
\[
\text{Var}\{\hat{k}_2\{I_k\}\} = \frac{\kappa_4\{I_k\}}{N} + \frac{2\kappa_2^2\{I_k\}}{N-1} \tag{56}
\]
\[
\text{Var}\{\hat{k}_3\{I_k\}\} = \frac{9\kappa_2\{I_k\}\kappa_4\{I_k\}}{N-1} + \frac{9\kappa_2^2\{I_k\}}{N-1} + \frac{\kappa_6\{I_k\}}{N} + \frac{6N\kappa_2^2\{I_k\}}{(N-1)(N-2)} \tag{57}
\]
In addition, the covariance of the second order and third order k-statistics is [48]
\[
\text{Cov}\{\hat{k}_2\{I_k\}, \hat{k}_3\{I_k\}\} = \frac{\kappa_5\{I_k\}}{N} + \frac{6\kappa_2\{I_k\}\kappa_3\{I_k\}}{N-1}. \tag{58}
\]

The covariance matrix of \( \hat{k}_2\{I_k\} \) and \( \hat{k}_3\{I_k\} \) then can be written as
\[
\mathbf{K} = \begin{bmatrix}
\text{Var}\{\hat{k}_2\{I_k\}\} & \text{Cov}\{\hat{k}_2\{I_k\}, \hat{k}_3\{I_k\}\} \\
\text{Cov}\{\hat{k}_2\{I_k\}, \hat{k}_3\{I_k\}\} & \text{Var}\{\hat{k}_3\{I_k\}\}
\end{bmatrix}. \tag{59}
\]
Letting \( \kappa = [\kappa_2\{I_k\}, \kappa_3\{I_k\}]^T \) and \( \hat{\kappa} = [\hat{k}_2\{I_k\}, \hat{k}_3\{I_k\}]^T \), the statistic \( Q_p \)
\[
Q_p = (\hat{\kappa} - \kappa)^T \mathbf{K}^{-1}(\hat{\kappa} - \kappa) \tag{60}
\]
will asymptotically follow a \( \chi^2 \) distribution with degrees of freedom equal to 2 [37]. With a significance level \( \alpha_c \), we can calculate the acceptance region for the \( \kappa, \hat{\kappa}_c \), according to \( p(Q_p > z_{\alpha_c}) = \alpha_c \). The comparison of the theoretical log-cumulants with the estimated ones thus can be quantitatively evaluated.

\section*{IV. Experimental Results}

Experiments on both simulated data and real SAR data are designed. With the former, we can validate the distributions and statistics of the \( l_2 \)-norm, whereas the latter aims to demonstrate the applications as well as advantages of the \( l_2 \)-norm to SAR data analysis.

\subsection*{A. Simulated Data}

First, experiments are carried out to validate the PDFs in Table II. Simulated data with different texture distributions are tested, including the gamma distribution (Gamma), the inverse gamma distribution (InvGamma), the Fisher distribution (Fisher), the beta distribution (Beta) and the beta prime
distribution (Beta Prime), as well as constant values (Dirac) [49]. The parameters for the gamma and inverse gamma distributions are $\alpha = 8$, and $\xi = 8$, $\zeta = 12$ for the other three ones. All the data share the same covariance matrix

$$C_1 = 10^{-3} \times \begin{bmatrix} 161 & -7 - 4j & 39 - j \\ -7 + 4j & 82 & -4 + 4j \\ 39 + j & -4 - 4j & 100 \end{bmatrix} \quad (61)$$

which is estimated from a forest area of an AirSAR data acquired over the Netherlands.

The PDFs in Table II are calculated, and compared with the histograms of the simulated data. The result is shown in Figure 2, where the solid lines stand for the PDFs, and the points represent the histograms. It shows that the histograms conform to the corresponding PDFs closely, and different distributions can be discriminated by the PDFs, though not effectively. This is also validated by the quantitative assessment shown in Table III, where the KS distance is employed. The first row refers to the distances between the histograms and the corresponding PDFs (KS1), and as a reference, the second row shows the distances from the histograms to Gaussian PDF (KS2). As it shows, the KS1 is smaller than 0.01 for all the simulated data, while KS2 is larger than 0.02 for textured data. The values confirm that the histograms can be employed to texture analysis. It is necessary to remind that the KS distance depends on the bin size of the histogram, therefore, the results are meaningful only when the bin size is same.

The Entropy given by the covariance matrix in (61) is 0.9275, which is high as the correlation coefficients between polarimetric channels are very small. By keeping the intensities and increasing the correlation, another two covariance matrices, with Entropy 0.6940 and 0.5187 respectively, are obtained

$$C_2 = 10^{-3} \times \begin{bmatrix} 161 & 79.6 - 46j & 39 - j \\ 79.6 + 46j & 82 & -4 + 4j \\ 39 + j & -4 - 4j & 100 \end{bmatrix} \quad (62)$$

$$C_3 = 10^{-3} \begin{bmatrix} 161 & 79.6 - 46j & 39 - j \\ 79.6 + 46j & 82 & 51.2 + 51.2j \\ 39 + j & -51.2 - 51.2j & 100 \end{bmatrix} \quad (63)$$

Gaussian distributed data are simulated using these covariance matrices (Figure 3a), each containing 200 $\times$ 200 samples. The normalized moments up to the fifth order are shown in Figure 3b, in which the theoretical values are represented by lines, and estimated ones by asterisks. It shows that the estimated values fit the theoretical ones very well, and data with different Entropy are distinguished in the plot, laying between the two boundaries representing Entropy 0 and Entropy 1. Table IV shows the distances between the estimated normalized moments and the theoretical ones, $d = |\hat{nm}_v - \bar{nm}_v|$, as well as the square root of the estimation variance, $\sigma = \sqrt{\text{Var}\{\hat{nm}_v\}}$. Based on the $3\sigma$ rule, we can judge that the data is characterized by the corresponding distribution, as all $d < 3\sigma$. In addition, it shows that the estimation variance of higher order normalized moments, the fifth order for example, is very large, therefore, higher orders should be avoided unless the sample size is very large.

Normalized moments are also tested on simulated data with different texture distributions, including the gamma, the inverse gamma and the Fisher distribution (Figure 3a). Texture distribution parameters, see Table I, are given by $\alpha = 8$ and $\xi = 8$, $\zeta = 12$, and covariance matrix by (61). The result is shown in Figure 3c. The values of Gaussian distributed data is also plotted as a reference. The quantitative assessment based on the estimation variance is shown in Table V. Similar results are obtained and it demonstrates again that the normalized moments can be employed to determine the distribution of SAR data. Textured data have a larger normalized moment than Gaussian distributed data. In Table V, the variance of the 4th and 5th order normalized moments for the inverse gamma distributed texture is not defined, as the variance involves 8th
and 10th order moments which do not exist when $\alpha = 8$, see the expression in Table I.

In addition to the normalized moments, the log-cumulants of the $l_2$-norms are calculated. Simulated data with texture distributions in Table I, as well as non-texture data, are tested. The result is shown in Figure 4a. For each trial, 100 $\kappa_3$-$\kappa_2$ pairs are estimated using the bootstrap method [21], with each bootstrap sample consisting of 20000 pixels. The estimated pairs are plotted in the log-cumulant diagram (point clouds), and are compared with the theoretical values (the + markers). The acceptance regions with significance level $\alpha_c = 0.05$ are also depicted (the black ellipses). As we can see, the estimated log-cumulants are congruent with the theoretical ones, as there are only several points falling out of the ellipses.

Multilook processing is applied to the scattering vectors using two different window sizes, $3 \times 3$ and $5 \times 5$, and with the obtained sample covariance matrices, matrix variate log-cumulants are estimated. The results are shown in Figure 4b and Figure 4c. As the correlation length of the simulated texture is 3, the sample matrix variate log-cumulants are completely different. With a window size smaller or equal than the correlation length, the estimated statistics are in agreement with the theoretical ones. While a large window size will average out the texture significantly, resulting in Gaussian statistics. Therefore, with the matrix variate log-cumulants, the choice of the filtering window size is critical in some circumstances.

From Figure 4a and Figure 4b, we can see that the matrix variate log-cumulants are well separated for different distributions, while the log-cumulants based on the $l_2$-norm are overlapped for some distributions. This is because the matrix variate log-cumulants of the texture part are scaled by a nonlinear function of orders as shown in (19), and the difference between distributions is amplified. From this point of view, the matrix variate log-cumulants are of advantage when the multilook window size is properly chosen.

### B. Real SAR Data

Real SAR data including RADARSAT-2 Fine Quad-Pol data (RST2) and FSAR X-band full-pol data (FSAR) from the German Aerospace Center (DLR) are analyzed using the $l_2$-norm. The two data have quite different spatial resolutions, $11.1m \times 7.6m$ (Range × Azimuth) for the RST2 data, and $0.25m \times 0.25m$ for the FSAR data. Original data are in the single look complex format.

Three Regions Of Interest (ROI) over the crops area from each data are tested, see Figure 5 and Figure 6. For the RST2 data, each ROI covers $50 \times 50$ pixels. It can be seen that the ROIs are represented by very different colors from the Pauli decomposition, implying that the polarimetric information they convey are different. The ROIs in the FSAR data are much larger, each covers $200 \times 200$ pixels. Compared with the appearance of the RST2 data, the differences among these ROIs are not so significant. The Pauli decomposition shows that the ROIs in both images are very homogeneous, no appreciable texture is observed.

The normalized moments of the $l_2$-norms are computed on all ROIs, and the values are compared with those of Gaussian distributions, since the normalized moments can be used to evaluate the non-gaussianity as shown before. Because the true covariance matrices are unknown, they are estimated using the fixed point estimator [29]–[31]

$$\hat{\Sigma}_{j+1} = \frac{d}{N} \sum_{i=1}^{N} \frac{k_i k_i^\dagger}{\|k_i\Sigma_j k_i\|}$$

where $N$ is the number of pixels in the neighborhood. The estimation is accomplished by a recursive process initialized with

$$\hat{\Sigma}_1 = \frac{1}{N} \sum_{i=1}^{N} k_i k_i^\dagger.$$

And the criterion to stop the iterative process is [31]

$$\frac{\|\hat{\Sigma}_{j+1} - \hat{\Sigma}_j\|_F}{\|\hat{\Sigma}_j\|_F} < 10^{-15}$$

where $\|\cdot\|_F$ is the Frobenius norm.

The results are shown in Figure 5 and Figure 6. As we can see, the normalized moments estimated from the RST2 data fit those calculated using the covariance matrix very well, and they are well separated between the boundaries representing Entropy 0 and Entropy 1. It is rational to conclude that these ROIs can be modeled by Gaussian distributions. In contrast, the result on the FSAR data shows different behaviors. First of all, all the ROIs seem to have similar Entropy, applications such as classification based on the Entropy, as a consequence, may not work. Secondly, there are large discrepancies between the estimated values and the theoretical values for all ROIs. Apparently, Gaussian distributions are not accurate any more. Statistical information, or texture information, is of great value to identify different types of crops in this case.

Further validations using the log-cumulants based on the $l_2$-norm are performed, see Figure 7a and Figure 7b. Here only the impact from the texture part is shown by removing the contribution of the speckle part in (47). This is because

<table>
<thead>
<tr>
<th>Tests</th>
<th>Distance</th>
<th>$v = 2$</th>
<th>$v = 3$</th>
<th>$v = 4$</th>
<th>$v = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>$d$</td>
<td>0.0025</td>
<td>0.0106</td>
<td>0.0116</td>
<td>0.0143</td>
</tr>
<tr>
<td></td>
<td>$3\sigma$</td>
<td>0.0189</td>
<td>0.0638</td>
<td>0.0395</td>
<td>1.9337</td>
</tr>
<tr>
<td>$C_2$</td>
<td>$d$</td>
<td>0.0005</td>
<td>0.0126</td>
<td>0.0153</td>
<td>0.0025</td>
</tr>
<tr>
<td></td>
<td>$3\sigma$</td>
<td>0.0173</td>
<td>0.1073</td>
<td>0.0927</td>
<td>4.8918</td>
</tr>
<tr>
<td>$C_3$</td>
<td>$d$</td>
<td>0.0010</td>
<td>0.0143</td>
<td>0.0981</td>
<td>0.5775</td>
</tr>
<tr>
<td></td>
<td>$3\sigma$</td>
<td>0.0182</td>
<td>0.1379</td>
<td>0.0669</td>
<td>7.6351</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tests</th>
<th>Distance</th>
<th>$v = 2$</th>
<th>$v = 3$</th>
<th>$v = 4$</th>
<th>$v = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamma</td>
<td>$d$</td>
<td>0.0113</td>
<td>0.0080</td>
<td>0.0077</td>
<td>0.6125</td>
</tr>
<tr>
<td></td>
<td>$3\sigma$</td>
<td>0.0196</td>
<td>0.1701</td>
<td>1.5148</td>
<td>16.7464</td>
</tr>
<tr>
<td>InvGamma</td>
<td>$d$</td>
<td>0.0160</td>
<td>0.1074</td>
<td>0.3839</td>
<td>8.1539</td>
</tr>
<tr>
<td></td>
<td>$3\sigma$</td>
<td>0.0301</td>
<td>0.4614</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Fisher</td>
<td>$d$</td>
<td>0.0193</td>
<td>0.1490</td>
<td>0.7832</td>
<td>3.3680</td>
</tr>
<tr>
<td></td>
<td>$3\sigma$</td>
<td>0.0335</td>
<td>0.4588</td>
<td>9.2493</td>
<td>436.6224</td>
</tr>
</tbody>
</table>
the contribution from the speckle part is not the same for different ROIs due to the different covariance matrices of these ROIs. The acceptance regions of Gaussian distributions with significance level $\alpha_c = 0.05$ are also plotted. Figure 7a shows that most of the points fall inside the corresponding ellipse, indicating that Gaussian distributions are proper for these ROIs. For the FSAR data, it seems that the ROI 1 and ROI 3 can be modeled by beta prime distributed textures, whereas the inverse gamma distribution is proper for the ROI 2. That the size of the acceptance regions depends on the number of samples gives rise to the large difference in the sizes of the ellipses in Figure 7a and Figure 7b, as there are much more pixels in the FSAR ROIs than the RST2 ROIs.

The matrix variate log-cumulants are also computed for the FSAR data, after applying a multilook process to the scattering vectors with a $5 \times 5$ sliding window. Each bootstrap sample contains 500 sample covariance matrices. And the ENL is estimated as 6.4885. From Figure 7c, we can see that most points from ROI 2 and ROI 3 fall inside the black ellipse, which refers to the acceptance regions of a Wishart distribution with significance level $\alpha_c = 0.05$. It demonstrates again that the non-gaussianity can be removed during the multilook processing.

The calculation of matrix variate log-cumulants involves the estimation of the ENL, which is a difficult task in the analysis of PolSAR data. One needs to select a Gaussian distributed homogeneous area manually, and then use the trace moment-based estimator or log-determinant moment-based estimator [32]. Another option is the texture-invariant estimator [50] suggested by Liu et al., which can be applied on textured data. But it requires that the texture is constant in the multilook window, which is not guaranteed in all circumstances. On the contrary, the log-cumulants of the $l_2$-norms depend on the covariance matrix instead of the ENL, which can be estimated either on homogeneous area or textured area using the well studied fixed point estimator (64).

Apart from the crops areas, two forest sites from the FSAR data are analyzed, one is with short trees, and the other with...
high trees. Six ROIs of $200 \times 200$ pixels are selected for each test site, see Figure 8. The Pauli decomposition shows that the test regions are very heterogeneous due to the alternative appearing of the strong backscattering from the crowns of trees and the weak one from the shadows. Log-cumulants of the $l_2$-norms, as well as matrix variate log-cumulants are calculated. Figure 8c and Figure 8d show the values of the first site, and those of the second site are in Figure 8e and Figure 8f. It seems that the textures of all the ROIs can be modeled by Fisher distributions. We can see that different ROIs from the same test site have similar results, for example, the $\kappa_2$ of the first site is less than 1.5 and the $\kappa_3$ is less than 0, which are well separated from the second site, where the $\kappa_2$ is larger than 1.5 and $\kappa_3$ larger than 0. This means the forests can be further classified into different types using texture information. The textures are not affected too much by the multilook processing, see Figure 8c and Figure 8d for example. That the window size of the filtering, $5 \times 5$ pixels, is smaller than the size of trees could be the reason.

V. Conclusions

Analyzing the statistical properties of PolSAR data is a difficult task due to the multidimensional signal. A common way doing this is to study the statistics of each polarimetric channel separately. The correlation between different polarimetric channels are usually neglected, though it provides very useful information that makes the multidimensional SAR system differ from a single SAR system. In this paper, the $l_2$-norm, or the Span, of the scattering vector is studied, and it is found to be a useful tool for texture analysis of PolSAR data, which takes into account both the intensities and correlations. In addition, the Span is polarimetric invariant.

Two important kinds of statistics of the $l_2$-norm are studied. The normalized moments is an extension of the normalized intensity moments, which can be employed to evaluate the non-Gaussian behavior of the data easily. The values of different orders could reveal the polarimetric properties of PolSAR data, as well as the texture properties. The log-cumulants can be used to identify the texture distribution of...
the data. No multilook processing is required to compute these values, which is an advantage in the analysis of high spatial resolution data, since the statistical behavior may change and mixtures of targets may appear after a filtering. In addition, no ENL is required to obtain these quantities. Estimators for the normalized moments and log-cumulants are provided. The former is biased but the bias asymptotically vanishes as increasing the sample size. The latter is an unbiased estimator. Estimation variance of the estimators is also given. One can further develop parameter estimators and design goodness-of-fit tests using these values. We include this in the future work.

Results on real SAR data show that texture information is of great importance. When different targets belong to a same category, saying crops or forest, the scattering mechanisms may be same, and similar polarimetric information will be observed. It is difficult to discriminate these targets by the polarimetric information only. Combining information conveyed by the texture, however, could give a better identification of targets. It is possible to distinguish crops of different types, by the texture, however, could give a better identification of targets. It is possible to distinguish crops of different types, and forests with trees of different characteristics.

The statistics are derived based on the assumption that the texture for all polarimetric channels are the same. Multi-texture, where the correlations of polarimetric channels are important since the textures may be correlated, will be considered in the next step.

ACKNOWLEDGMENTS

The authors would like to thank the European Space Agency (ESA) for the RADARSAT-2 data provided in the frame of the AgriSAR 2009 campaign, and Priv.-Doz. Dr. Andreas Reigber from the Microwave and Radar Institute (HR) of the German Aerospace Center (DLR) for the FSAR data.

APPENDIX A

DISTRIBUTION OF THE NORM

Assume that \( z \) follows a complex Gaussian distribution, \( z \sim \mathcal{N}(0, \Sigma) \), where the covariance matrix can be diagonalized by \( \Sigma = P^TDP \) with \( P \) orthogonal and \( D = \text{diag}(\lambda_1, \cdots, \lambda_d) \) diagonal. Then if \( u = Pz \), by changing the variables, we have \( u \sim \mathcal{N}(0, D) \). The square of the \( l_2 \)-norm of \( u \) can be written as a sum

\[
I_u = u^T u = \sum_{i=1}^{d} I_i
\]

where \( I_i \) follows an exponential distribution

\[
p(I_i) = \frac{1}{\lambda_i} \exp \left(-\frac{I_i}{\lambda_i}\right)
\]

with Moment Generating Function (MGF) given by

\[
M_i(t) = \frac{1}{1 - \lambda_i t}
\]

Since all \( I_i \)'s are independent (due to the covariance matrix \( D \) is diagonal), the MGF of \( I_u \) is then the product of those of all components

\[
M(t) = \prod_{i=1}^{d} \frac{1}{1 - \lambda_i t} = \sum_{i=1}^{d} \frac{p_i}{1 - \lambda_i t}
\]

where

\[
p_i = \frac{\lambda_i^{d-1}}{\prod_{j=1, j\neq i}^{d} (\lambda_i - \lambda_j)}
\]

The PDF of \( I_u \), therefore, can be obtained from the MGF as

\[
p(I_u) = \sum_{i=1}^{d} \frac{p_i}{\lambda_i} \exp \left(-\frac{I_u}{\lambda_i}\right).
\]

The square of the \( l_2 \)-norm of \( z \) can be calculated by

\[
I_z = (P^{-1}u)^T(P^{-1}u) = I_u
\]

which means \( p(I_k) = p(I_u) \).
Let \( \nu = \nu \) and test areas of the second site. (c) Log-cumulants of the \( l_2 \)-norms (first site). (d) Matrix variate log-cumulants (first site). (e) Log-cumulants of the \( l_2 \)-norms (second site). (f) Matrix variate log-cumulants (second site).

**APPENDIX B**

**ESTIMATION OF THE MEAN AND VARIANCE**

In the following text, all the moments and normalized moments are calculated on the \( l_2 \)-norm of the scattering vector. For simplicity, \( m_v \{ I_k \} \) and \( nm_v \{ I_k \} \) are replaced by \( m_v \) and \( nm_v \). The mean value of \( \hat{m}_v \) is given by [37]

\[
E\{\hat{m}_v\} = E\left\{ \frac{1}{N} \sum_{i=1}^{N} (k_i^v)^v \right\} = m_v \tag{74}
\]

and the covariance by

\[
Cov\{\hat{m}_v, \hat{m}_v\} = E\left\{ \left( \frac{1}{N} \sum_{i=1}^{N} (k_i^v)^v - m_v \right) \times \left( \frac{1}{N} \sum_{i=1}^{N} (k_i^v)^v - m_v \right) \right\} \tag{75}
\]

\[
= -m_v m_v + \frac{1}{N^2} E\left\{ \sum_{i=1}^{N} (k_i^v)^v \right\} \tag{75}
\]

\[
+ \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} (k_i^v)^v (k_j^v)^v \right\}
\]

\[
= \frac{1}{N} (m_v + m_v - m_v m_v)
\]

Let \( \nu = \nu \), the variance of \( \hat{m}_v \) is obtained as

\[
Var\{\hat{m}_v\} = \frac{1}{N} (m_2v - m_v^2). \tag{76}
\]

Define a function \( f(x, y) \) as

\[
f(x, y) = \frac{x}{y^\nu}, \tag{77}
\]

then it can be expanded at \( (\mu_x, \mu_y) \) using the Taylor series [35]

\[
f(x, y) = \mu_x + (x - \mu_x) \frac{\partial f}{\partial x}_{\mu_x, \mu_y} + (y - \mu_y) \frac{\partial f}{\partial y}_{\mu_x, \mu_y} + \\
\frac{1}{2} (x - \mu_x)^2 \frac{\partial^2 f}{\partial x^2}_{\mu_x, \mu_y} + \frac{1}{2} (y - \mu_y)^2 \frac{\partial^2 f}{\partial y^2}_{\mu_x, \mu_y} + \\
(x - \mu_x)(y - \mu_y) \frac{\partial^2 f}{\partial x \partial y}_{\mu_x, \mu_y} + \\
O\left( \left( (x - \mu_x) \frac{\partial f}{\partial x} + (y - \mu_y) \frac{\partial f}{\partial y} \right)^3 \right) \tag{78}
\]

Ignoring all the terms higher than two and applying the expectation to individual terms gives the expectation of the function \( f \)

\[
E\{f(x, y)\} = \mu_x + \mu_y \frac{\partial f}{\partial y}_{\mu_x, \mu_y} + \frac{\nu (\nu + 1) \mu_x}{2 \mu_y^{\nu+2}}. \tag{79}
\]
Similarly, the variance of \( f \) can be obtained using the first-order terms of the Taylor series expansion:

\[
\text{Var}(f(x, y)) = \text{Var}(x) \frac{1}{\mu_y^{2v}} + \text{Var}(y) \frac{v^2 \mu_x^2}{\mu_y^{2(v+1)}}
\]

\[ - \text{Cov}(x, y) \frac{2v \mu_x}{\mu_y^{2v+1}} \tag{80} \]

Let \( x = \hat{m}_v, y = \hat{m}_1, \mu_x = m_v \) and \( \mu_y = m_1 \), we have the mean value and variance of the estimator \( \hat{m}_v \) as

\[
E\{\hat{m}_v\} = \frac{m_v}{m_1} + \frac{1}{N} \left( \frac{v^2 m^2_2 m_v^2}{m_1} - (v - 1)^2 m_v^2 + \frac{2v \cdot m_v m_{v+1}}{m_1} \right) \tag{81}
\]

and

\[
\text{Var}\{\hat{m}_v\} = \frac{1}{Nm_v^2} \left[ \frac{v^2 m^2_2 m_v^2}{m_1^2} - (v - 1)^2 m_v^2 + \frac{2v \cdot m_v m_{v+1}}{m_1} \right] \tag{82}
\]

As we can see, the bias of the estimator \( \hat{m}_v \) is given by

\[
B = \frac{1}{N} \left( \frac{v(v+1) m_2 m_v}{2m_1^{v+2}} - \frac{vm_{v+1}}{m_1^{v+1}} + \frac{v(1-v)m_v}{2m_1^v} \right) - \frac{1}{2N} \left( v^2 m_2 m_v - 2v \cdot nm_{v+1} \right)
\]

which asymptotically goes to 0 when increasing the sample size \( N \).

\section*{APPENDIX C
LOG-CUMULANT OF THE SPECKLE}

The \( \varphi_{k}(s) \) shown in (48) consists of two parts. The first part is the logarithm of the gamma function, and there is a well known solution to the differential of it, the polygamma function [20], [21]. For the second part, the logarithm of the CHSP (30), we can write it as

\[
G(s) = \ln h_{s-1}(\lambda_1, \cdots, \lambda_d)
\]

which leads to

\[
h_{s-1}(\lambda_1, \cdots, \lambda_d) = e^{G(s)} \tag{84}
\]

Calculate the differential on both sides, the following is obtained

\[
h'_{s-1}(\lambda_1, \cdots, \lambda_d) = e^{G(s)} G'(s)
\]

\[ = h_{s-1}(\lambda_1, \cdots, \lambda_d)G'(s) \tag{86} \]

and according to the Leibniz’s Rule, the \((v-1)\)th order differential of the above equation is

\[
h^{(v)}_{s-1}(s) = \sum_{k=0}^{v-1} \binom{v-1}{k} h^{(k)}_{s-1}(s) G^{(v-k)}(s) \tag{87}
\]

with the differential of the CHSP is given by

\[
h^{(v)}_{s-1}(s) = \sum_{i=1}^{d} \beta_i \lambda_i^{s-1}(\ln \lambda_i)^v. \tag{88} \]

Rearrange the result we get

\[
G^{(v)}(s) = h^{(v)}_{s-1}(\lambda_1, \cdots, \lambda_d) - \sum_{k=1}^{v-1} \binom{v-1}{k} h^{(k)}_{s-1}(\lambda_1, \cdots, \lambda_d) G^{(v-k)}(s) \tag{89} \]

Let \( \beta = G^{(v)}(s) \mid_{s=1} \) and \( h^{(v)}(s) = h^{(v)}_{s-1}(\lambda_1, \cdots, \lambda_d) \mid_{s=1} \), the result shown as (47) is obtained.

\section*{REFERENCES}


