Circularity and Sphericity of Complex Stochastic Models in Multivariate High-Resolution SAR Images
Gabriel Vasile, Nikola Besic, Andrei Anghel

To cite this version:
Gabriel Vasile, Nikola Besic, Andrei Anghel. Circularity and Sphericity of Complex Stochastic Models in Multivariate High-Resolution SAR Images. [Research Report] GIPSA-LAB. 2016. <hal-01402263>

HAL Id: hal-01402263
https://hal.archives-ouvertes.fr/hal-01402263
Submitted on 24 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Circularity and Sphericity of Complex Stochastic Models in Multivariate High-Resolution SAR Images

Gabriel VASILE\(^{(1)}\), Nikola BESIC\(^{(1)}\) and Andrei ANGHEL\(^{(1)}\)

\(^{(1)}\): Grenoble-Image-sPeech-Signal-Automatics Lab, CNRS / Grenoble INP GIPSA-lab DIS/SIGMAPHY, Grenoble INP - BP 46, 38402 Grenoble Cedex, FRANCE
Tel: +33 476 826 334 - Fax: +33 476 574 790 - Email: gabriel.vasile@gipsa-lab.grenoble-inp.fr

ABSTRACT

This paper presents a new methodological framework to assess the conformity of multivariate high-resolution SAR data in terms of asymptotic statistics. Three important statistical properties are studied by applying statistical hypotheses testing, successively: circularity, sphericity and spherical symmetry. Starting from the classical tests designed for the multivariate Gaussian case, these tests are extended to the Spherically Invariant Random Vector (SIRV) stochastic model. A zero-mean test is proposed for both Gaussian and SIRV stochastic processes. The link between the spherical symmetry property and the conformity to the SIRV model is established asymptotically by the specific structure of the quadricovariance matrix. Two high and very high resolution datasets are used to illustrate departures from the standard model assumptions: TerraSAR-X multi-pass InSAR and ONERA RAMSES POLSAR images. As well, the derived tests are applied on the appropriate synthetic dataset. The detection results are qualitatively and quantitatively analysed and some important inferences are drawn regarding these two datasets.

I. INTRODUCTION

Multidimensional Synthetic Aperture Radar (SAR) data, like POLSAR and multi-pass InSAR images, describe the interaction between the electromagnetic waves and the scatterers inside each resolution cell. Multivariate SAR images, acquired by either spaceborne or airborne sensors, are currently analyzed to extract useful information concerning the physical properties of the illuminated target. For distributed targets, this analysis relies on the stochastic properties of the SAR data. In general, multivariate SAR data can be locally modelled by a multivariate zero-mean circular Gaussian stochastic process, which is completely determined by its covariance matrix.

With the improved resolution of modern SAR platforms, the number of scatterers inside each resolution cell decreases considerably. The higher scene heterogeneity may eventually lead to non-Gaussian clutter modelling. More complex stochastic models, such as the SIRP (Spherically Invariant Random Process), are then required. Several special cases of univariate stochastic processes (K-compound, Weibull, etc.) have been extensively studied over the years (for example in coastal radar applications) before being reunited under the common umbrella of SIRP \([1]\).

Polarimetry and multi-pass interferometry extend the dimensionality of the remotely sensed SAR data. Multivariate versions of SIRP distributions, namely SIRV (Spherically Invariant Random Vectors) \([2]\), are employed for modelling high-resolution...
POLinSAR [3–8] and InSAR [9–13] data. This multiplicative model is expressed as a product between the square root of a scalar positive quantity (texture) and the description of an equivalent homogeneous surface (speckle) [14]. For an m-dimensional SAR system, the single channel model [15] has been extended as follows: in each azimuth / range location, k is the m × 1 complex target vector corresponding to the same area on the ground.

For distributed targets, the corresponding k vector is considered non-deterministic and may be written, under the SIRV assumption as $k = \sqrt{\tau} \cdot z$, where $\tau$ is a positive random variable (texture) while $z$ is a complex-valued, centred and Gaussian distributed, random vector with covariance matrix $[M] = E\{zz^H\}$. The two random variables are statistically independent. Characterizing k reduces to the writing of the probability distribution $p_\tau(\tau)$ of $\tau$ and the normalized covariance matrix $[M]$ of $z$, with $\text{trace}([M]) = m$. It is important to notice that in the SIRV definition, the texture probability density function is not explicitly specified. As a consequence, SIRVs describe a whole class of stochastic processes [16]. If $p_\tau(\tau)$ is a Dirac pulse, then the Gaussian (multivariate) model is retrieved for k. In many field applications, $p_\tau(\tau)$ is assumed to have a predefined analytical form. In this case, a specific SIRV is employed such as the multivariate $K$ distribution (Gamma texture) [17], the $G^0$ distribution (inverse Gamma texture) [18] or the KummerU distribution (Fisher texture) [4]. Nevertheless, at fixed number of samples, as the number of parameters increases, the variance of their estimates will increase.

The four components of the POLSAR k target vector (in lexicographic order) describing the polarimetric characteristics of a given resolution cell are the elements of the Sinclair matrix: $S_{hh}$, $S_{hv}$, $S_{vh}$ and $S_{vv}$. For mono-static configurations, where the reciprocity theorem applies, $S_{hv} = S_{vh}$, only three components remain: $S_{hh}$, $S_{hv}$ and $S_{vv}$. In this case, the dimension of the target vector k becomes $m = 3$.

In the case of conventional two-pass InSAR system [19], only two channels are involved and $m = 2$. By denoting with $c = \rho \exp(j\phi)$ the complex correlation coefficient, the target relative displacement $d_{12}$ between the two acquisitions can be retrieved from the exact knowledge of SAR antenna phase center positions, terrain height, acquisition geometry, and an estimate of the differential interferometric phase $\phi_{12}$. $\rho_{12}$ is called interferometric coherence and it describes both the local phase stability and the amplitude decorrelation of the InSAR pair. The phase information $\phi_{12}$ allows phase differences (interferograms) to be computed in order to measure topography or target displacements between repeated pass acquisitions. In the general case, the $m$-dimensional interferometric target vector k will contain information about the relative displacements between each combination of 2 passes. The main parameter to be estimated is the speckle covariance matrix, from which normalized correlation coefficients can be easily derived.

The question if the SIRV models, despite their flexibility, are appropriate for describing any multidimensional SAR dataset, still remains. More exactly, "what is to be gained, and through which strategies, in pursuit of higher-order statistics, and non-Gaussian models"? One recommendation has been released by the POLinSAR scientific community, namely to "Increase reliance on quantitative objective norms for comparing the performance of alternative analysis". This fact motivated us to propose a general framework which allows quantitative evaluation of fitting SIRV stochastic models with respect to a given multidimensional SAR dataset. Moreover, the illustrated multivariate high-resolution SAR datasets show that in this case, it can be worth challenging two specific and very important stochastic properties: the circularity and the sphericity.
This paper is organized as follows. In Section II we introduce classical circularity and sphericity tests along with their extensions to the SIRV stochastic model. Section III contains in detail elaborated method of quantitative assessment of the SIRV conformity, relying on testing spherical symmetry. In Section IV, the results obtained using the proposed robust tests are presented and analyzed with synthetic, multi-pass InSAR and POLSAR datasets. Section V presents some general remarks regarding the proposed methodological framework. In Section VI, some conclusions and perspectives are presented. Eventually, Appendix I introduces a zero-mean test, which is proposed for both Gaussian and SIRV stochastic processes.

In this paper, we assume that the covariance matrix $[M]$ is of full rank $m$. In the specific case of a non-invertible matrix, the same considerations can be applied using only the non-zero signal subspace.

II. CIRCULARITY AND SPHERICITY

In the context of multivariate SAR data, we introduce two stochastic properties, often assumed in case of a Gaussian model: circularity and sphericity. After adequately defining the properties, we present the classical tests (Gaussian model) and propose the ones extended to the SIRV stochastic model.

A. Circularity

A complex-valued random variable ($r_0 = x_0 + iy_0$) is circular [20] if its distribution remains invariant to multiplication with complex numbers on the unity circle i.e. if the real random vector $r_0 = (x_0, y_0)^T$ is spherically symmetric with respect to the origin. When dealing with the circularity of complex random vectors, we ought to rely on the second order properties [21] by introducing the complex extended target vector $\mathbf{z} = [\mathbf{k}^T, \mathbf{k}^H]^T$. Starting from $\mathbf{z}$, the second-order statistical properties of the complex target vector $\mathbf{k}$ can be analysed using the extended covariance matrix:

$$[R] = E\{\mathbf{z}\mathbf{z}^H\} = \begin{bmatrix} [C] & [P] \end{bmatrix} \in C^{2m},$$

(1)

where $[C] = E\{\mathbf{k}\mathbf{k}^H\}$ is the complex positive semi-definite Hermitian symmetric covariance matrix and $[P] = E\{\mathbf{k}\mathbf{k}^T\}$ is the complex symmetric pseudo-covariance matrix\(^3\) of the target vector $\mathbf{k}$ [22].

Schreier et al. proposed the circularity Generalized Likelihood Ratio Test (GLRT) by employing the previously defined extended covariance matrix [23]. The considered hypotheses are:

$$\begin{cases} H_0 : [P] = 0, & \text{\(\mathbf{k}\) is circular}, \\ H_1 : [P] \neq 0, & \text{\(\mathbf{k}\) is not circular} \end{cases}$$

(2)

This circularity test is checking for the block-diagonality of the extended covariance matrix with respect to the covariance and pseudo-covariance matrix.

1) Gaussian random processes: The probability density function (PDF) of a zero-mean complex circular Gaussian target vector $\mathbf{k}$ can be generalized with respect to its associated extended target vector $\mathbf{z}$ as [21, 24]:

$$p_G(\mathbf{k}) = \pi^{-m}(\det[R])^{-\frac{1}{2}} e^{-\frac{||\mathbf{z}||^2}{2}}.$$  

(3)

\(^3\)Also known as "relation matrix" or "complementary covariance matrix".
The maximum-likelihood (ML) estimator of the extended covariance matrix is the Sample Extended Covariance Matrix (SECM) obtained by replacing the statistical mean from Eq. 1 with the spatial average:

$$\hat{R}_{SECM} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{z}_i \mathbf{z}_i^H = \begin{bmatrix} \hat{C}_{SCM} & \hat{P}_{SPM}^* \\ \hat{P}_{SPM} & \hat{C}_{SCM} \end{bmatrix},$$

(4)

where $N$ is the number of samples, $\hat{C}_{SCM}$ and $\hat{P}_{SPM}$ are the Sample Covariance Matrix (SCM) and the Sample Pseudo-covariance Matrix (SPM) estimators, respectively. Under the constraint $[P] = [0]_m$ imposed by $H_0$, the ML estimator is [23]:

$$\hat{R}_{SECM 0} = \begin{bmatrix} \hat{C}_{SCM 0} & [0]_m^* \\ [0]_m & \hat{C}_{SCM} \end{bmatrix}.$$ 

(5)

By introducing Eqs. 4 and 5 to the Likelihood Ratio Test (LRT) associated to Eq. 2, Schreier et al. derived in [23] the Generalized Likelihood Ratio Test (GLRT) for $N$ i.i.d. observed samples:

$$\Lambda(k_1,...,k_N) = \frac{\det[\hat{R}_{SECM}]}{\det[\hat{C}_{SCM}]} \overset{H_0}{\gtrless} \lambda.$$ 

(6)

The asymptotic distribution of the decision statistic under the null hypothesis is $H_0 : -N \ln \Lambda \rightarrow \chi^2_{m(m+1)}$ (Theorem 2 in [25]). Schreier et al. showed that the GLRT from Eq. 6 is invariant with respect to invertible linear transforms.

2) Spherically Invariant Random Vectors: The GLRT makes use of the Gaussian assumption when inserting $\hat{C}_{SCM}$ and $\hat{P}_{SPM}$ estimators in the LRT. Therefore, it ought to be modified before being applied on SIRV data. By directly applying Tyler’s Corollary 1 from [26], Ollila and Koivunen [25] showed that the GLRT $(\hat{C}_{ML}, \hat{P}_{ML})$ for SIRV can be derived by dividing the logarithm of the GLRT statistics from Eq. 6 ($\ln \Lambda$) by the correction factor $\hat{\gamma}$, estimated as:

$$\hat{\gamma} = \sum_{i=1}^{m} \frac{E\{|k_i|^4\}}{E\{|k_i|^2\}^2} r_i(k_i)^2 + 2.$$ 

(7)

with respect to the circularity coefficients, representing the "amount of circularity":

$$r_i(k_i) = \frac{|E\{|k_i|^2\}|}{E\{|k_i|^2\}}.$$ 

(8)

Following the same procedure as Ollila and Koivunen, $N$ is replaced by $(N - m)$ from the Box approximation of the GLRT [27]. Finally, according to Tyler’s Corollary 1, the decision statistic under the null hypothesis is $H_0 : -(N - m) \frac{\ln \Lambda}{\hat{\gamma}} \rightarrow \chi^2_{m(m+1)}$.

B. Sphericity

The sphericity of a $m$-dimensional complex circular random vector $\mathbf{z}_0$ is defined with respect to the augmented real random vector. Let $\mathbf{z}_0 = [x_0^T, y_0^T]^T$ be the augmented $2m \times 1$ real random vector, where $x_0$ and $y_0$ are its real and imaginary parts, respectively. The augmented real target vector can be obtained from the complex extended random vector using the following transform [21, 24]:

$$\mathbf{z} = \frac{1}{2} \begin{bmatrix} |f|^m_m & |f|^m_m \\ -j|f|^m_m & j|f|^m_m \end{bmatrix}.$$ 

(9)
The sphericity GLRT was introduced by Mauchly for real-valued m-dimensional Gaussian random processes [28]. If the random process is circular, it can be directly extended to complex-valued random processes also. The hypotheses are:

\[
\begin{align*}
H_0 & : [C] = \varsigma[I]_m, \quad \text{k is spherical}, \\
H_1 & : [C] \neq \varsigma[I]_m, \quad \text{k is not spherical}.
\end{align*}
\]  

(10)

where \(\varsigma \in \mathbb{R}\) is unknown.

For the sake of simplicity and without loss of generality, the test is build such that \([M] = [I]_m\). As a consequence, the derived sphericity test must present invariance to linear transforms, also.

Given a set of \(N\) i.i.d. observed samples, the Mauchly’s sphericity LRT is:

\[
\Lambda_s(k_1, ... k_N) = m \frac{(\det[M])^{\dagger} H_1}{\text{tr}[M]} \lambda_s.
\]

(11)

In [29], \(\Lambda_s\) is reported to be invariant with respect to scale and invertible linear transforms (see Chapter 10.7 on page 431):

\[
\Lambda_s(k_1, ... k_N) = \Lambda_s([V]k_1, ..., [V]k_N),
\]

(12)

with \([V][M][V]^H = [I]_m\). One can observe that this linear transform is not the same as the whitening transform changing the SIRV covariance matrix to \([I]_m\). Consequently, a SIRV exhibiting sphericity is a particular case of a SIRV with equal variances of the differences between all combinations of target vector components.

In particular, the sphericity property needs to hold when processing conventional multi-pass interferometric SAR data by interferometric pair. If the sphericity is not respected, fully multivariate processing should be applied (for example when jointly using the InSAR coherence and phase for 3D phase unwrapping [30] or relevant scatterer detection [31]): the InSAR channels are not independent with equal variance (homoscedastic).

1) Gaussian random processes: When replacing the covariance matrix by the SCM (ML estimate under Gaussian assumption), the GLRT is obtained as:

\[
\Lambda_s(k_1, ... k_N) = m \frac{(\det[M]_{SCM})^{\dagger} H_1}{\text{tr}[M]_{SCM}} \lambda_s.
\]

(13)

Asymptotically, \(H_0 : -N \ln \Lambda_s \rightarrow \chi^2_{(m+2)(m-1)}\) [28, 32].

2) Spherically Invariant Random Vectors: Following Muirhead and Waternaux original studies on the robustness of the GLRT from Eq. 11 when sampling from a SIRV [32], Tyler proposed two different robust approximated GLRTs for sphericity [26, 33].

The first method (Corollary 1 from [26]) is derived using the same reasoning as in the case of circularity. The derived approximated GLRT is reported to be inefficient for moderate departures from Gaussianity.

In this paper, we adopt the second method for constructing the approximated GLRT for sphericity (Corollary 4 from [26]):

\[
a \Lambda_s(k_1, ... k_N) = (\det[M]_{FP})^{\dagger} \lambda_s,
\]

(14)

where \([M]_{FP}\) is the iterative Fixed-Point covariance matrix estimator (already employed with multivariate SAR data [3]):

\[
[M]_{FP} = m \frac{1}{N} \sum_{i=1}^{N} \frac{k_i k_i^H}{k_i^H [M]^{-1}_{FP} k_i}.
\]

(15)
When originally introducing this estimator, Tyler showed in [34] that \( [M]_{FP} \) is an affine-invariant covariance matrix M-estimator [35]. It has been demonstrated by Tyler that the Fixed-Point estimator is an AML estimator for SIRVs [36]. Hence, all conditions required to apply Corollary 4 from [26] are now fulfilled and \( H_0 : -N^{-m+1}_m \ln \alpha \Lambda_s \rightarrow \chi^2(m+2)(m-1) \) in distribution. Due to the Theorem V.4 in [37], \( N \) is replaced by \( N^{-m+1}_m \).

Finally, the formerly introduced circularity test complements the sphericity test presented here: the test for circularity checks that there is no correlation between the real and the imaginary part of the complex random vector, while the sphericity test checks for equal independence between the target vector components (independence and homoscedasticity). Hence, circularity is a requirement before testing for sphericity, as \( k \) and \( k^* \) cannot be independent.

### III. Spherical symmetry

According to Vershik’s definition [38], \( \zeta_0 \) is spherically invariant if and only if the characteristic function of the augmented random vector \( \zeta_0 \) can be written as:

\[
\Phi_{\zeta_0}(\zeta_0) = \phi \left( \frac{\zeta_0^T [C] \zeta_0}{2} \right), \quad \phi(v) = \int_0^\infty e^{-\tau v} p_r(\tau) d\tau,
\]

where \([C]\) is a positive definite characteristic matrix and \( p_r(\tau) \) is a probability measure on \([0, \infty)\). More intuitively, by applying \( \zeta'_0 = [C]^{-1/2} \zeta_0 \) linear transform on the augmented random vector, the density generator function \( \phi(v) \) in Eq. 16 takes the form:

\[
\phi(\zeta_0^T \zeta'_0 / 2) = f(\|\zeta_0\|^2),
\]

where \( \|\cdot\| \) is the Euclidian norm. This implies that the SIRV \( \zeta_0 \) is reduced to a new SIRV \( \zeta'_0 \) with its new covariance matrix equal to \([I]_m \) (identity matrix of order \( m \)). \( \zeta'_0 \) clearly exhibits spherical invariance.

Let now \( \phi \) be defined in Eq. 16 so that the matrix \([C]\) is the speckle normalized covariance matrix [6] and the conditions from the Yao’s representation theorem are respected [2]. In consequence, for a given texture pdf \( p_r(\tau) \), the SIRV process describing the complex target vector \( k \) can be defined as:

\[
p_k(k) = \frac{1}{\pi^m \det[M]} h_m \left( k^1 [M]^{-1} k \right),
\]

with \( h_m(q) = \int_0^\infty \tau^{-m} \exp \left( -\frac{q}{2} \right) p_r(\tau) d\tau \). \( \det[M] \) denotes the determinant of the matrix \([M]\). Yao (Lemma 4.1 in [2]) also demonstrated SIRV closure under invertible linear transform (e.g. as suggested for Eq. 17). All target vectors \( k \) satisfying these conditions accept a product model stochastic representation [39].

Note also that a random target vector \( k_E \) has an elliptically symmetric distribution [40] if it is affinely equivalent in distribution to a spherically symmetric target vector \( k_S \):

\[
k_E = [A] k_S + b.
\]

Therefore, spherical symmetry represents a particular case of elliptical symmetry when \([A] = [I]_m \), \( b = [0]_{m \times 1} \). In our case, by identifying in the SIRV whitening transform from Eq. 17, it yields \([A] = [C]^{-1/2} \) and \( b = [0]_{m \times 1} \) (due to the zero mean). Hence, on one side, a spherically symmetric white SIRV vector, being characterized with unity covariance matrix, is necessarily spherical. On the other side, the previously introduced sphericity test is no longer valid in the absence of the elliptical symmetry, making the last property a mandatory constraint in testing the sphericity.
The spherical symmetry property in multivariate statistics is defined with respect to a SIRV with covariance matrix $[I]_m$ (white SIRV). The general SIRV case, with unknown covariance matrix, was studied under the "elliptical symmetry" property. Since the family of spherical symmetric distributions - SSD can be considered as the standardized form of the family of elliptically contoured distributions - ECD (by employing the whitening transform from Eq. 17), we adopt in this paper the original Yao’s notation by considering zero mean elliptically contoured distributions as SIRVs. Hence, we use the term "spherical symmetry" for both SSD and ECD goodness-of-fit testing. See [41] for an exhaustive presentation on SSD and ECD.

Spherical symmetry testing was firstly introduced by Kariya and Eaton [42], using an alternative form of the Lehmann and Stein lemma [43] with known covariance matrix. In the common situation where the covariance matrix is estimated from data, several strategies for robust spherical symmetry testing have been proposed by Beran [44], King [45], Baringhaus [46], Fang et al. [47], Manzotti et al. [48] and Huffer et al. [49] among the most recent publications. Li et al. proposed in [50] a graphical method for spherical symmetry testing: the Q-Q probability plots. This method has been applied in [51] to hyperspectral image analysis.

According to [49], one of the most powerful spherical symmetry tests was proposed in [52] for real random vectors. In this section, we have adapted the Schott test for multidimensional complex SAR data analysis.

A. The Schott test for circular complex random vectors

Assuming the existence of the forth order moment (or quadricovariance) matrix:

$$[M]_4 = k_ik_i^H \otimes k_ik_i^H,$$ (20)

Schott proposed the Wald test [53] for verifying that the structure of $[M]_4$ corresponds to a SIRV (as originally given by Tyler in [26]). This structure holds for circular complex random vectors, also [40].

With complex random vectors, the first modification is the sample quadricovariance estimator. According to [54], the sample complex quadricovariance estimator can be expressed in terms of the Kronecker product $\otimes$ as:

$$\hat{[M]}_4 = \frac{1}{N} \sum_{i=1}^{N} k_ik_i^H \otimes k_ik_i^H,$$ (21)

where the transposed operator $^T$ is replaced by the conjugate and transpose operator $^H$. Its corresponding standardized form is:

$$\hat{[M]}_4^* = \left([M]^{-\frac{1}{2}} \otimes [M]^{-\frac{1}{2}}\right) \hat{[M]}_4 \left([M]^{-\frac{1}{2}} \otimes [M]^{-\frac{1}{2}}\right).$$ (22)

According to the Schott’s theorem, the Wald test statistic for spherical symmetry can be expressed as:

$$T_{Schott} = 2N \left\{ \beta_1 \text{tr} \left(\hat{[M]}_4^2\right) + \beta_2 \text{vec} \left(\hat{[I]}_m\right)^H \hat{[M]}_4^2 \text{vec} \left(\hat{[I]}_m\right) - 3\beta_1 + (m + 2)\beta_2 \right\} m(m + 2)(1 + \tilde{\kappa})^2 - 24,$$ (23)

where

$$\beta_1 = (1 + \tilde{\theta})^{-1}/24,$$ (24)

$$\beta_2 = -3\alpha[24(1 + \tilde{\theta})^2 + 12(m + 4)\alpha(1 + \tilde{\theta})]^{-1},$$ (25)
\[ a = (1 + \hat{\theta}) + (1 + \hat{\kappa})^3 - 2(1 + \hat{\kappa})(1 + \hat{\eta}), \]  

(26)

with the Mardia’s kurtosis \( \hat{\kappa} \) and the generalized higher order scalar moments \( \hat{\theta}, \hat{\eta} \) given by:

\[ (1 + \hat{\kappa}) = \frac{1}{m(m+2)N} \sum_{i=1}^{N} \left[ k_i^H \hat{M} \right]^{-1} k_i, \]  

(27)

\[ (1 + \hat{\theta}) = \frac{1}{m(m+2)(m+4)N} \sum_{i=1}^{N} \left[ k_i^H \hat{M} \right]^{-1} k_i^3, \]  

(28)

\[ (1 + \hat{\eta}) = \frac{1}{m(m+2)(m+4)(m+6)N} \sum_{i=1}^{N} \left[ k_i^H \hat{M} \right]^{-1} k_i^4. \]  

(29)

Asymptotically, \( T_{Schott} \to \chi^2_{v_{m,\text{complex}}} \) with \( v_{m,\text{complex}} = \frac{m^2(m+1)(m+5)}{12} - 1 \). This represents the second modification with respect to the Wald test from [52]. According to Schott, the degrees of freedom is set according to the number of unknowns of the quadricovariance matrix:

\[ v_{m,\text{complex}} = 2(v_{m,\text{real}} + 1) - \left[ \frac{m(m-1)}{2} + m \right] - 1, \]  

(30)

where \( v_{m,\text{real}} = m^2 + \frac{m(m-1)(m^2+7m-6)}{24} - 1 \) as in [52]. The second term in Eq. 30 comes from the number of real elements of \( \hat{M} \): \( \frac{m(m-1)}{2} \) elements of the form \( x_i^2 x_j^2 \) and \( m \) elements of the form \( x_i^4 \). It has been proven in [55] that the Wald test and the LRT are asymptotically equivalent.

**IV. RESULTS AND DISCUSSIONS**

To illustrate the proposed tests, results obtained with the synthetic POLSAR dataset, high-resolution TerraSAR-X multi-pass InSAR and very high-resolution ONERA RAMSES [56] POLSAR data are reported. These data sets are shown in Fig. 1.

[Figure 1 about here.]

This section is dedicated to the analysis of the three data sets in terms of circularity, sphericity and spherical symmetry. Since the mean equal to zero is a requirement for all the derived tests, we equally show results obtained by applying the zero-mean test, introduced in Appendix I, on the augmented real random vector \( \zeta_0 \). This test is valid for both Gaussian and SIRV stochastic models.

Finally, the discussions associated to the presented results concern these high-resolution multivariate SAR data sets, only. The estimation neighbourhood is the \( 13 \times 13 \) boxcar and the false alarm probability threshold is \( p_{fa} = 0.01 \).

**A. Synthetic data**

The synthetic dataset is composed out of nine different regions (Fig. 2a). Six of them are characterized with the SIRV PolSAR clutter, while for the remaining three the clutter is Gaussian. In four regions the additive thermal noise (circular or non-circular) is present. As well, we introduce a coherent scattering through the simulated elementary reflectors (two trihedral, dihedral and dipole). In the derivation we relied on the ML deterministic texture estimator:

\[ \hat{\tau}_{FP} = \frac{k_i^H \hat{M}_{FP}^{-1} k_i}{m}. \]  

(31)
Aside from the assumed detection of the coherent scattering sub-regions (elementary reflectors), circularity test rejects successfully the regions corrupted with additive non-circular thermal noise. The heterogeneity (borders between different regions) does not appear to influence this test significantly.

Sphericity test detects properly the Gaussian clutter with additive circular thermal noise as well as the Gaussian clutter without noise. As it is the case with the circularity test, the influence of non-stationarity is negligible.

However, the spherical symmetry test does not seem to be immune to the heterogeneity, but still quite successfully rejects the coherent scattering (deterministic target).

B. High-resolution multi-pass InSAR data

The 3-pass interferometric stripmap HH images were acquired in 2009, at 11-day interval, over the Argentire village, France with a mean incidence angle of $5^\circ$, an azimuth-resolution of 3.3 m and a slant-range resolution of 1.8 m. This data set has been used for SAR tomography over this area as it exhibits a high coherence level over the main buildings from Fig. 1-(a) [57]. The background image from Fig. 3-(a), (b), (c), (d) is the fixed point Interferometric Whitening Filter (IWF) span (Eq. 31).

The pixels illustrated in magenta on Fig. 3-(a) indicate where the zero-mean test from Appendix I is rejected. It can be observed from Table I that this dataset is zero-mean: the percentage of rejected pixels is much less than the significance level a priori set (1% in all cases).

Fig. 3-(b) shows in red the pixels where the adjusted circularity test is rejected. These pixels should be processed as $2m$ real random vectors. The percentage of rejected non circular pixels (cf. Table I) is, although larger than the significance level, still small enough (3.14%) for us to conclude that this dataset is circular.

Fig. 3-(c) and Fig. 3-(d) illustrate in blue the pixels where the sphericity test from Eq. 14 is accepted, and in green the pixels where the spherical symmetry test from Eq. 23 is rejected. The quantitative results summed up in Table I indicate that both sphericity and spherical symmetry properties are significant for this dataset. Qualitatively, it can be observed that:

- localized mainly in regions with high density of strong scatterers, the nonspherical pixels should be treated as a fully multivariate process inside the local neighborhood when estimating the InSAR coherence and phase parameters;
- mostly located in the same areas of the image, the nonspherically symmetric pixels indicate where the SIRV (and consequently the compound Gaussian) model fails to properly describe the multivariate clutter.

This specific behavior may be linked with the presence of strong deterministic scattering in urban areas: the analyzed target is not distributed but deterministic. Nevertheless, the proposed tests build a methodological framework to study this effect with respect to the SIRV model.
Spherical symmetry is a prerequisite for sphericity. For quantitatively validating the obtained results, the percentage pixels detected as spherical and rejected for spherical symmetry was computed in Table I: the value is quite small considering the significance level. This condition is met for non-circular and non-zero mean pixels, also.

Additionally, about 18% of non-circular pixels are rejected by the Schott test for spherical symmetric circular complex random vectors. This illustrates that circularity should be tested before testing for spherical symmetry.

C. Very high resolution POLSAR data

Illustrated in Fig. 1-(b), this data set was acquired over Toulouse, France with a mean incidence angle of 50°. It represents a fully polarimetric (monostatic mode) X-band acquisition with a spatial resolution of approximately 0.5 m in range and azimuth.

The background image from Fig. 4-(a), (b), (c), (d) is the fixed point Polarimetric Whitening Filter (PWF) span obtained by applying Eq. 31 with polarimetric target vectors. We have used here the same representation and the same color coding as in the previous section, while Table II sums up the POLSAR results. It can be observed that this dataset is globally zero-mean and not spherical: the corresponding percentages are less than the imposed significance level. The SIRV model (spherical symmetry) holds in about three-quarters of all realizations, while the non-sphericity confirms that multivariate statistical modeling is correctly employed since the clutter is not spherical.

In other words, Fig. 4-(d) underlines the same phenomena as in Fig. 3-(d): the spherical symmetry is rejected over the urban areas exhibiting strong deterministic scattering. Quantitatively, one can notice in Table II that the Schott test for spherical symmetric circular complex random vectors is rejecting non-circular pixels in about 31% of all cases, only. Hence circularity testing is mandatory prior to testing the SIRV model conformity.

Finally, both Fig. 4-(c) and Table II indicate a relatively high percentage of noncircular pixels in this POLSAR dataset. These pixels are mainly located in the weak backscattering image areas (shadowing). Since the data are zero-mean, this effect is not induced by any non-centered thermal noise additive component. However, it may be introduced by the specific calibration of this airborne very high resolution POLSAR data: the motion compensation module tracks and rectifies the signature of specific calibrated point targets on the ground.

V. GENERAL REMARKS

When dealing with multivariate high-resolution SAR data, it is crucial to decide if a specific stochastic model is properly fitting the experimental dataset inside the estimation neighborhood. The accepted stochastic process with the smallest number of parameters should be selected. In other words, if both the SIRV model and the Gaussian model are fitting the data, the latter will have better estimation performances with a finite number of samples. Based on the results presented in this paper, we provide the methodological framework to assess multivariate SAR data conformity:
1) **Zero-mean** test from Appendix I. STOP if the test is rejected: the mean parameter should be included in the SIRV stochastic model.

2) **Circularity** test: STOP if the test is rejected: the augmented (or extended) random vectors should be used instead of the complex target vectors.

3) **Sphericity** test: STOP if the test is accepted: it is better to use one complex circular random variable for each channel, separately.

4) **Spherical symmetry** test: STOP if the test is rejected: SIRVs are not stationary inside the neighborhood and improved statistical (or simply deterministic) modeling should be addressed.

5) Testing for **non-Gaussianity** (for example the one proposed in [6]). STOP if the test is rejected: the dataset is more likely to be locally Gaussian and multivariate zero-mean complex circular Gaussian statistics should be employed.

By successively performing the proposed tests for a specific multivariate SAR dataset, it is possible to asymptotically evaluate the pertinence of various model-based statistical processing schemes (filtering, segmentation or detection).

In terms of theoretical performance analysis, the adjusted generalized LRT is asymptotically uniformly most powerful according to the Neyman-Pearson lemma. This “optimality” holds provided the estimators plugged into the LRT (or the Wald test) are consistent and unbiased, which is the case for our study.

Special care must be taken when applying the tests with multivariate SAR data. In theory, the proposed conformity testing holds as long as the observed number of samples (estimation neighborhood) is large enough with respect to the dimension of the target vector, especially for the spherical symmetry test (based on the specific structure of the quadricovariance).

Finally, it is important to stress that no predefined analytical form was imposed on the texture probability function when establishing the conformity tests. Therefore, they can be directly applied for a wide class of stochastic processes currently used for describing multivariate high-resolution SAR data.

**VI. Conclusions**

This paper has presented a new methodological framework to asymptotically assess the conformity of multivariate high-resolution SAR data. The proposed approach consist of applying successively three statistical hypotheses tests for verifying three important statistical properties: circularity, non sphericity and spherical symmetry. The latter is asymptotically equivalent, under certain hypotheses, to the conformity of the experimental data with respect to the SIRV product model. In addition, the zero-mean (from Appendix I) and the non-Gaussianity [6] tests can be used to decide which model is better suited to asymptotically fit the experimental data.

The proposed framework, aside from the fact that it is gathering the most notable advances in the field of signal processing, is introducing the extension of both the sphericity and the spherical symmetry tests with zero-mean complex circular SIRV assumption.

The effectiveness of the proposed detection schemes was illustrated by high-resolution TerraSAR-X multi-pass InSAR and very high resolution ONERA RAMSES POLSAR data. The conclusions driven from the analysis of the obtained results are important with respect to the two tested datasets: non-sphericity can be an important issue for spaceborne multi-pass InSAR,
while circularity is important for airborne POLSAR data.

It has been illustrated that in strong heterogeneous clutter, such as the urban environment, the SIRV model can fail. The bottom line is that characterization of urban regions is much more complex (and difficult) - since a more complex model, with more parameters, may be required. In the light of the results shown in this paper, SIRV models may be less appropriate for urban areas characterization. However, alternative explanations are possible. As an example, the root of this inappropriateness might as well be the assumed ergodicity / stationarity (in spatial sense) for the backscattered signal and, also, even in the hypothesis of randomness: targets exhibit a deterministic behavior.

First, the very use of a sliding analysis window for estimating the stochastic parameters of the scattered signal may be questioned, as it implicitly assume that the considered signal is ergodic / stationary. While this hypothesis holds for distributed and uniform targets, where the physical parameters (and, thus, the electromagnetic scattering behavior) differs very little from one resolution cell to another, in urban areas the physical structure (and, as such, its electromagnetic behavior) may change considerably from one resolution cell to the next. This makes the hypothesis of ergodicity / stationarity less applicable.

Second, one should note that even the randomness of the radar echo is not given, but assumed. This is mainly a way to deal with the inherent complexity of the signal. Anyway, for identical measuring conditions, the recorded radar data is perfectly identical. Even if small differences in measuring conditions lead to strong discrepancies in the recorded data, this is not an evidence for randomness, as such behavior can be fully explain under a deterministic paradigm - the chaotic models. Various parameters, such as meteorological conditions and, even more important, the changes that the target suffers in time (between two succeeding acquisitions, for example), account for the observed randomness of the recorded data. However, these changes of the target are more significant for green targets (such as forests and agricultural fields), where humidity and wind modify both their physical structure and their electromagnetic behavior. On the other hand, those changes are less significant for urban targets and, as such, randomness is less likely for the latter.

In perspective, applying chaotic (or pseudo-chaotic) [58–60] models to POLSAR / InSAR data from urban areas can be a possible solution. These models should be able to take into account the deterministic features of those areas (presence of dihedral angles, straight edges, cavities, etc.), while still leaving room for some unpredictability (orientation of those elements). Using chaotic models in POLSAR and multi-pass InSAR data will make the object of our future work.

**APPENDIX I: ZERO-MEAN TEST FOR BOTH GAUSSIAN AND SIRV STOCHASTIC MODELS**

Under circularity, the zero-mean testing can be performed by using the $T^2$-statistic. When testing the hypothesis $H_0$ that a mean vector $\bar{k}$ is equal to zero, the generalized likelihood ratio criterion for the circular multivariate Gaussian model is:

$$T^2 = N|k^H[\widehat{M}]^{-1}_{SCM}k|,$$

(32)

with $T^2 \to \chi^2_m$ under $H_0$ (see Theorem 5.2.3 from [29]). In case circularity is not assumed, this test can be directly applied on the augmented real random vector $\zeta_0$. This results holds asymptotically for the SIRV models according to Theorem 5.7.1 from [29]. More details can be found in [61–63].

The GLRT from Eq. 32 is not scale invariant like the ones from Sections II-A.2 and II-B.2. Thus, Tyler’s Corollaries 1 and 4 from [26] cannot be applied directly.
ACKNOWLEDGMENT

The authors would also like to thank Dr. C. Tison (CNES, France) for providing the high-resolution POLSAR images over Toulouse and the German Aerospace Center (DLR) for providing the TerraSAR-X stripmap SAR images through the MTH0232 and the MTH0828 projects.

REFERENCES

<table>
<thead>
<tr>
<th></th>
<th>Non-zero-mean (3751 pixels)</th>
<th>Non-circularity (32882 pixels)</th>
<th>Sphericity (424856 pixels)</th>
<th>Non-spherical-symmetric (328125 pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-zero-mean</td>
<td>0.36 %</td>
<td>0.01 %</td>
<td>0.02 %</td>
<td>0.18 %</td>
</tr>
<tr>
<td>Non-circularity</td>
<td>0.01%</td>
<td>3.14%</td>
<td>1.27%</td>
<td>0.55%</td>
</tr>
<tr>
<td>Sphericity</td>
<td>0.02%</td>
<td>1.23%</td>
<td>40.52%</td>
<td>3.14%</td>
</tr>
<tr>
<td>Non-spherical-symmetric</td>
<td>0.18%</td>
<td>0.6%</td>
<td>3.14%</td>
<td>31.3%</td>
</tr>
</tbody>
</table>
TABLE II: Toulouse, RAMSES POLSAR data, X-band: detection results (percentage computed with respect to the total number of pixels).

<table>
<thead>
<tr>
<th></th>
<th>Non-zero-mean (45249 pixels)</th>
<th>Non-circularity (230104 pixels)</th>
<th>Sphericity (1283 pixels)</th>
<th>Non-spherical-symmetric (373983 pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-zero-mean</td>
<td>4.52%</td>
<td>0.91%</td>
<td>0%</td>
<td>2.3%</td>
</tr>
<tr>
<td>Non-circularity</td>
<td>0.91%</td>
<td>23.01%</td>
<td>0.02%</td>
<td>7.2%</td>
</tr>
<tr>
<td>Sphericity</td>
<td>0%</td>
<td>0.02%</td>
<td>0.13%</td>
<td>0%</td>
</tr>
<tr>
<td>Non-spherical-symmetric</td>
<td>2.3%</td>
<td>7.18%</td>
<td>0%</td>
<td>37.4%</td>
</tr>
</tbody>
</table>
Fig. 1: Data sets: (a) Synthetic POLSAR data, intensity color composition of the target vector elements $k_1-k_3-k_2$ in Pauli basis, (b) TerraSAR-X 3-pass InSAR data, amplitude color composition of the complex random vector elements $k_1-k_2-k_3$, (c) Toulouse, RAMSES POLSAR data, intensity color composition of the target vector elements $k_1-k_3-k_2$ in Pauli basis.
Fig. 2: Synthetic data set, 200x200 pixels: (a) description of the regions, (b) circularity rejection map superposed on the span, (c) sphericity map superposed on the span, (d) spherical symmetry rejection map superposed on the span.
Fig. 3: Argètierre, TerraSAR-X 3-pass InSAR data, X-band, 1024 × 1024 pixels: (a) zero-mean rejection map superposed on the IWF span in Db, (b) circularity rejection map superposed on the IWF span in Db, (c) sphericity map superposed on the IWF span in Db, (d) spherical symmetry rejection map superposed on the IWF-FP span in Db.
Fig. 4: Toulouse, RAMSES POLSAR data, X-band, 1000 × 1000 pixels: (a) zero-mean rejection map superposed on the PWF span in Db, (b) circularity rejection map superposed on the PWF span in Db, (c) sphericity map superposed on the PWF span in Db, (d) spherical symmetry rejection map superposed on the PWF span in Db.