

# Statistically Consistent Measures of Distance for POLSAR Data

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

**Abstract**—In this paper, a homoskedastic and additive model for polarimetric SAR (POLSAR) data is introduced. The statistical properties of this model subsequently lead to a few different consistent measures of distance for the POLSAR covariance matrix. The dis-similarity measurements are initially developed for both partial (2x2) and full (3x3 monostatic case) polarimetric SAR data. They are computed from the determinant of the POLSAR covariance matrix. Interestingly, when the multi-polarization POLSAR is collapsed into the single-polarization SAR, this covariance matrix determinant is neatly transformed into the traditional SAR intensity. Thus the SAR intensity statistics can now be considered as a special case of the theoretical model for POLSAR. While powerful theoretically, the statistical model is also robust in handling practical data. It can be validated and match reasonably well with practical data, even when the supposedly independent polarimetric components are actually highly correlated (e.g. between  $S_{hh}$  and  $S_{vv}$ ). Moreover, the match can be further improved, as better match is shown achievable, if the model validation process include an estimation of the dataset's Effective Number of Looks (ENL) instead of blind employment of the look number given by the SAR processor. Towards the end, the practical application of these dissimilarity measures is briefly demonstrated in the context of evaluating POLSAR speckle filters.

**Index Terms**—Polarimetric, Synthetic aperture radar, speckle-filtering, homoskedasticity

## I. INTRODUCTION

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

$$s_{part} = \begin{bmatrix} S_h \\ S_v \end{bmatrix} \quad (1)$$

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

$$s_{full} = \begin{bmatrix} S_{hh} \\ \sqrt{2}S_{hv} \\ S_{vv} \end{bmatrix} \quad (2)$$

Let  $\Sigma = E[ss^{*T}]$  denotes the population expected value of the POLSAR covariance matrix, where  $s^{*T}$  denotes the complex conjugate transpose of  $s$ . Assuming all the elements in  $s$  are independent and  $s$  is jointly circular complex Gaussian

with the given covariance matrix  $\Sigma$ , then the probably density function (PDF) of  $s$  can be written as:

$$pdf(s; \Sigma) = \frac{1}{\pi^d |\Sigma|} e^{-s^{*T} \Sigma^{-1} s} \quad (3)$$

where  $|M|$  denotes the determinant of the matrix  $M$ .

As covariance matrix is only defined on multiple data points, sample covariance matrix of POLSAR data is commonly presented in “ensemble” format. They are formed as the mean of Hermitian outer product of single-look scattering vectors

$$C_v = \langle ss^{*T} \rangle = \frac{1}{L} \sum_{i=1}^L s_i s_i^{*T} \quad (4)$$

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

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

$$pdf(Z; d, \Sigma, L) = \frac{|Z|^{L-d}}{|\Sigma|^L \Gamma_d(L)} e^{-tr(\Sigma^{-1} Z)} \quad (5)$$

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

Our approach differs by applying the homoskedastic log transformation on a less-than-well-known result for the determinant of the covariance matrix. In [1], it has been proven that the ratio between the observable and the expected values of the sample covariance matrix's determinants behaves like a product of  $d$  chi-squared random variables with different degrees of freedom

$$\chi_L^d = (2L)^d \frac{|C_v|}{|\Sigma_v|} \sim \prod_{i=0}^{d-1} \chi^2(2L-2i) \quad (6)$$

Its log-transformed variable consequently behaves like a summation of  $d$  log-chi-squared random variables with the same degrees of freedom

$$\Lambda_L^d = \ln \left[ (2L)^d \frac{|C_v|}{|\Sigma_v|} \right] \sim \sum_{i=0}^{d-1} \Lambda^\chi(2L-2i) \quad (7)$$

with  $\Lambda^\chi(k) \sim \ln [\chi^2(k)]$

In this paper, the above results are further explored in the context of POLSAR data processing and a few scalar, consistent and homoskedastic measures of distance is proposed. The paper is structured as follows: after the preceding introduction to the basics of POLSAR statistical analysis, the

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

Manuscript received ?, 2013; revised ?.

second section presents conclusive evidence for two related points. The first is that: POLSAR data is multiplicative and heteroskedastic in its original domain. And the second: log-transformation converts it into an additive and homoskedastic model. Consequently a few consistent measures of distance are presented in section III. After the model is validated against both partial and full polarimetric SAR data in Section IV, section V discusses how the proposed models for the multi-variate POLSAR can be used to derive the well-known models for the univariate SAR data as its special case. Besides being theoretically powerful, the model is shown, in section VI, to be robust in handling a few practical imperfections that violate conventional assumptions. As an example application of the consistent and homoskedastic dis-similarity measures, section VII explores how they can be used in evaluating POLSAR speckle filters. Finally, section VIII critically reviews related works in literature before providing some conclusional discussion.

## II. ORIGINAL HETOROSKEDASTIC DOMAIN AND THE HOMOSKEDASTIC LOG-TRANSFORMATION

In this section the multiplicative nature of POLSAR data is illustrated. Log-transformation is shown converting this into a more familiar additive model. Heteroskedasticity, which is defined as the dependence of variance upon the underlying signal, is proved to be the case for the original POLSAR data. In log-transformed domain, the case for a homoskedastic model, where sample variance is fixed and thus independent of the underlying signal, is demonstrated. To keep the section flowing, the mathematical derivation is only presented here in major sketches. For more detailed derivation, Appendix A is to be reviewed.

From Eqn. 6 and Eqn. 7 we have:

$$|C_v| \sim |\Sigma_v| \cdot \frac{1}{(2L)^d} \cdot \prod_{i=0}^{d-1} \chi^2(2L - 2i) \quad (8)$$

$$\ln |C_v| \sim \ln |\Sigma_v| - d \cdot \ln(2L) + \sum_{i=0}^{d-1} \Lambda(2L - 2i) \quad (9)$$

In a given homogeneous POLSAR area, the parameters  $\Sigma_v$ ,  $d$  and  $L$  can be considered as constant. Thus Eqn. 8 gives the theoretical explanation that: in the original POLSAR domain, a multiplicative speckle noise pattern is present. At the same time, Eqn. 9 shows that the logarithmic transformation converts this into a more familiar additive noise.

Since chi-squared random variables  $X \sim \chi^2(k)$  follows a known PDF:

$$pdf(x; 2L) = \frac{x^{L-1} e^{-x/2}}{2^L \Gamma(L)} \quad (10)$$

Applying the variable change theorem, its log-transformed variable follows the PDF of:

$$pdf(x; 2L = k) = \frac{e^{Lx - e^x/2}}{2^L \Gamma(L)} \quad (11)$$

As the PDFs become available, the characteristic functions (CF) of both the chi-squared and log-chi-squared random

variables can be written as:

$$CF_\chi(t) = (1 - 2it)^L \quad (12)$$

$$CF_\Lambda(t) = 2^{it} \frac{\Gamma(L + it)}{\Gamma(L)} \quad (13)$$

Subsequently their means and variances can be computed from the given characteristic functions. They are:

$$avg[\chi(2L)] = 2L \quad (14)$$

$$var[\chi(2L)] = 4L \quad (15)$$

$$avg[\Lambda(2L)] = \psi^0(L) + \ln 2 \quad (16)$$

$$var[\Lambda(2L)] = \psi^1(L) \quad (17)$$

where  $\psi^0()$  and  $\psi^1()$  stands for digamma and trigamma function respectively.

Since the average and variance of both chi-squared distribution and log-chi-squared distribution are constant, the product and summation of these random variables also has fixed summary statistics. Specifically:

$$\begin{aligned} avg \left[ \prod_{i=0}^{d-1} \chi^2(2L - 2i) \right] &= 2^d \cdot \prod_{i=0}^{d-1} (L - i), \\ var \left[ \prod_{i=0}^{d-1} \chi^2(2L - 2i) \right] &= \prod_{i=0}^{d-1} 4(L - i)(L - i + 1) - \prod_{i=0}^{d-1} 4(L - i)^2, \\ avg \left[ \sum_{i=0}^{d-1} \Lambda(2L - 2i) \right] &= d \cdot \ln 2 + \sum_{i=0}^{d-1} \psi^0(L - i), \\ var \left[ \sum_{i=0}^{d-1} \Lambda(2L - 2i) \right] &= \sum_{i=0}^{d-1} \psi^1(L - i) \end{aligned}$$

Combining these results with Eqns. 8 and 9, we have:

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

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

$$avg[\ln |C_v|] = \ln |\Sigma_v| - d \cdot \ln L + \sum_{i=0}^{d-1} \psi^0(L - i) \quad (20)$$

$$var[\ln |C_v|] = \sum_{i=0}^{d-1} \psi^1(L - i) \quad (21)$$

Over a real and captured image, while the paramaters  $d$  and  $L$  do not change for the whole image, the underlying  $\Sigma_v$  is expected to differ from one region to the next. Thus over an heterogeneous scene, the stochastic process for  $|C_v|$  and  $\ln |C_v|$  varies depending on the underlying signal  $\Sigma_v$ . In such context, Eqn. 19 implies that the variance of  $|C_v|$  also differs depending on the underlying signal  $\Sigma_v$ , which indicates its heteroskedastic property. At the same time, in the log-transformed domain, Eqn. 21 shows that the variance of  $\ln |C_v|$  is invariant and independent of  $\Sigma_v$  manifesting its homoskedastic nature.

### III. CONSISTENT MEASURES OF DISTANCE FOR POLSAR

Similar to the way dispersion and contrast is defined in our previous work [2], this section introduces the consistent sense of distance in a few different perspective. First assuming that the true value of the underlying signal  $\Sigma_v$  is known *a priori*, the following random variables, namely ratio ( $\mathbb{R}$ ) and log-distance ( $\mathbb{L}$ ), are observable according to their definitions:

$$\mathbb{R} = \frac{|C_v|}{|\Sigma_v|} \quad (22)$$

$$\mathbb{L} = \ln |C_v| - \ln |\Sigma_v| \quad (23)$$

From another perspective where the POLSAR is known coming from an homogeneous area but the true value of the underlying signal  $\Sigma_v$  is *unknown*, consider the dispersion ( $\mathbb{D}$ ) and contrast ( $\mathbb{C}$ ) random variables being defined as:

$$\mathbb{D} = \ln |C_v| - \text{avg}(\ln |C_v|) \quad (24)$$

$$\mathbb{C} = \ln(|C_{v1}|) - \ln(|C_{v2}|) \quad (25)$$

Taking the results from Eqns. 8, 9 and 20 we have

$$\mathbb{R} \sim \frac{1}{(2L)^d} \cdot \prod_{i=0}^{d-1} \chi^2(2L - 2i) \quad (26)$$

$$\mathbb{L} \sim \sum_{i=0}^{d-1} \Lambda(2L - 2i) - d \cdot \ln(2L) \quad (27)$$

$$\mathbb{D} \sim \sum_{i=0}^{d-1} \Lambda(2L - 2i) - d \cdot \ln 2 + k \quad (28)$$

$$\mathbb{C} \sim \sum_{i=0}^{d-1} \Delta(2L - 2i) \quad (29)$$

with  $\Delta(2L) \sim \Lambda(2L) - \Lambda(2L)$  and  $k = \sum_{i=0}^{d-1} \psi^0(L - i)$

Also given the characteristic functions (CF) for the elementary components  $\Lambda(2L)$  written in Eqn. 13, Appendix B derives the characteristic functions for the summative random variables as:

$$CF_{\Lambda_L^d}(t) = \frac{2^{idt}}{\Gamma(L)^d} \prod_{j=0}^{d-1} \Gamma(L - j + it) \quad (30)$$

$$CF_{\mathbb{L}}(t) = \frac{1}{L^{idt} \Gamma(L)^d} \prod_{j=0}^{d-1} \Gamma(L - j + it) \quad (31)$$

$$CF_{\mathbb{D}}(t) = \frac{e^{ikt}}{\Gamma(L)^d} \prod_{j=0}^{d-1} \Gamma(L - j + it) \quad (32)$$

$$CF_{\Delta(2L)} = \frac{\Gamma(2L)B(L - it, L + it)}{\Gamma(L)^2} \quad (33)$$

$$CF_{\mathbb{C}}(t) = \prod_{j=0}^{d-1} \frac{\Gamma(2L - 2j)B(L - j - it, L - j + it)}{\Gamma(L - j)^2} \quad (34)$$

Since each elementary simulation component follows fixed distributions (i.e.  $\chi^2(2L)$ ,  $\Lambda(2L)$ , ...), it is natural that these variables also follow fixed distributions. Moreover, they are independent to  $\Sigma_v$ . This serves as conclusive evidence that these random variables follows consistent and fixed distributions, regardless of the underlying signal  $\Sigma_v$ .

### IV. VALIDATING THE MODELS AGAINST REAL-LIFE DATA

This section describes an experiment to validate the models above against real-life captured data. The validation procedure is quite straightforward. Given that the stochastic models have been derived in the previous sections, they can be graphically visualized by the histogram plots of the simulated data. At the same time, the form of the real-life practical data is also observable via histogram plots of the data samples extracted from an homogeneous area. Therefore, the theoretical models can be validated if for the same parameters the two plots match each other reasonably well.

As a demonstration, a homogeneous area was chosen from the AIRSAR Flevoland POLSAR data as experimental data samples. Then theoretical models are employed in an attempt to explain the data. The validating models include: the determinant and its log-transformed models, together with the dissimilarity measures namely: the determinant ratio, the log-distance, the dispersion and the contrast measures of distance.

The models are closely related. Given the same parameter set, the determinant and determinant ratio are just scaled version of each other. Meanwhile, the log-determinant, the log-distance and the dispersion are also just shifted version of each other. Thus ideally speaking if one model is validated all the other models will also be, assuming that all the parameters of the image are known exactly. Still in this section, all the models will be made subject to investigation.

Among all these models, the least-assumed stochastic process for dispersion and contrast measures of distance are validated first. For each pixel in the region, the determinant of the covariance matrix is computed and the log-transformation is applied. Then the average of POLSAR covariance matrix determinant in the log-transformed domain, i.e.  $\text{avg}(\ln |C_v|)$ , is measured for dispersion. Subsequently the observable samples of dispersion and contrast are computed according to Eqns. 24 and 25. Then their histograms are traced out.

At the same time, theoretical simulations according to Eqns. 28 and 29 are carried out. The nominal value of 4 was taken as the dataset's number of look, while the dimensional number is set to 3 or 2 respectively depending on whether full or partial polarimetric SAR dataset is being investigated. All the histogram plots are presented in Fig. 1. Apparently, a visual match is observable which verify the applicability of the theoretical models for the dispersion and contrast measures of distance.

Apart from dispersion and contrast, the other four models to be investigated require an estimation of the "true" underlying signal  $|\Sigma_v|$ . There are two ways to estimate this quantity over an homogeneous area. The traditional way is to simply set the true signal equal to the average of the polsar covariance matrix in its original domain, i.e.  $\Sigma_v = \text{avg}(C_v)$ . Another approach is to estimate the true signal from the average of the log-determinant of the polsar covariance matrix (i.e.  $\text{avg}[\ln |C_v|]$ ) using Eqn. 20. Both approaches will be explored in this section. However, given that the log-determinant average has already been computed earlier, the second approach is used first for the validation of determinant-ratio and log-distance.

Fig. 2 validate the models of determinant-ratio and log-distance against real-life data. In this experiment, the theo-

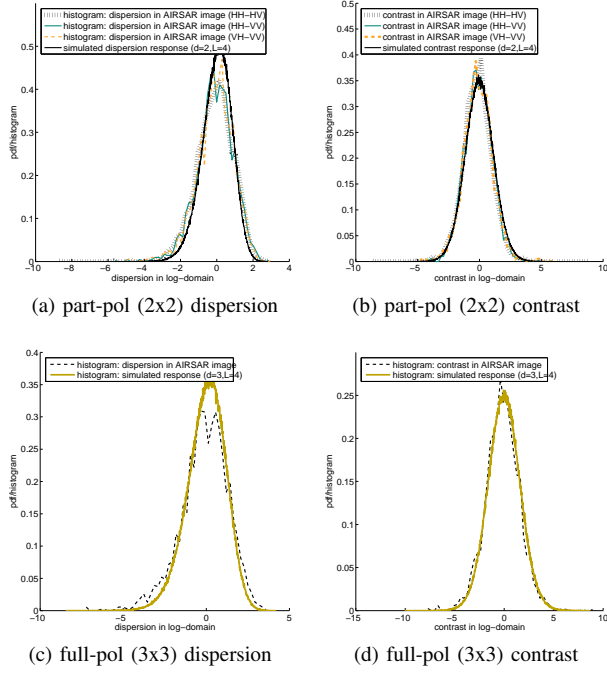


Fig. 1: Validating the dispersion and contrast models against both partial and full polarimetric AIRSAR Flevoland data.

retical models is simulated from Eqns 26 and 27, while the observable samples are computed using Eqns 22 and 23 with the true signal estimated from the log-determinant average, i.e.  $\text{avg}(\ln |C_v|)$ . Again a reasonable match is observed which validates the models for log-distance and determinant ratio.

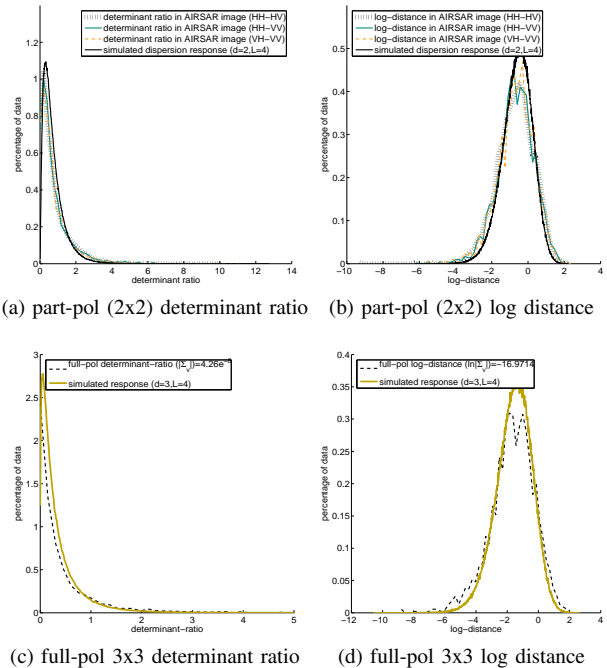


Fig. 2: Validating determinant-ratio and log-distance models with  $|\Sigma_v|$  is computed using  $\text{avg}(\ln |C_v|)$

Since the models for the determinant and log-determinant are just scaled or shifted version of the models for determinant-ratio and log-distance, similar validation results are to be expected. And if the true signal are computed in the same manner as described before then in fact similar match can be easily observed.

However, a more interesting phenomena is to be described. It happens in the validation process for determinant and its log-transformed model, where the theoretical response is taken from the simulated stochastic process described by Eqns. 8 and 9. The interesting phenomena happens when the true signal is estimated by the first approach i.e. equal to the average of the sample covariance matrix in its original domain. Interestingly, this approach results in a different estimation for the true signal with reference to the method described earlier.

Subsequently the validation plots, which are presented in Fig 3, exhibit some small discrepancies. The differences are easier to observe in the log-determinant plots, where except for some small shifted effects, the shapes of the models and the practical data exhibits certain resemblance.

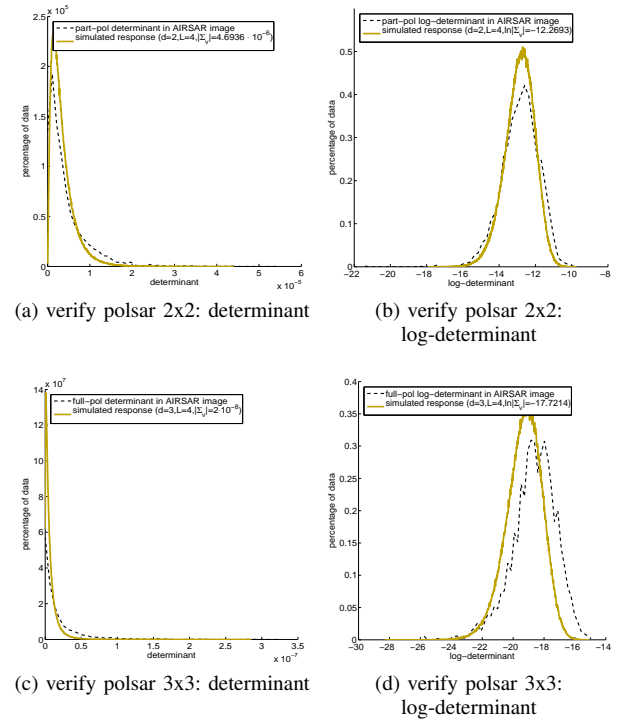


Fig. 3: Validating determinant and log-determinant models with  $\Sigma_v = \text{avg}(C_v)$

In short the least-assumed dispersion and contrast measures of distance are shown to match reasonably well with the practical data. The same can be stated for the other four models, namely: determinant, log-determinant, determinant ratio and log-distance, if the underlying parameters can be estimated reasonably well for the given image. However as described above a single “true signal”  $|\Sigma_v|$  can have two different estimated values, depending on which estimation method was being used. The discrepancy suggests that at least one parameter for the models was wrongly employed.

But what model parameter were used wrongly, and even if that can be corrected, would a better match become observable? The question is answered in Section VI, where not only the look number is shown to be misused but also the match of between the theoretical model and the practical data is shown to improve as well once a better look number (ENL) is estimated. For now, let us contend that: using appropriate estimation of the parameters, the proposed models match reasonably well with the practical data.

## V. SAR AS THE SPECIAL CASE OF POLARIMETRIC SAR

The previous section has validated the use of our models for 3-dimensional  $d = 3$  full polarimetric and two dimensional  $d = 2$  partial polarimetry cases. In this section, the focus is on the case where the dimensional number is reduced to 1  $d = 1$ . Physically this means the multi-dimensional POLSAR dataset is collapsed into the one-dimensional and classical SAR data. Mathematically, the sample covariance matrix is reduced to the sample variance and the determinant equates the scalar value. On another hand, it is well known that for SAR data, variance equals intensity. Thus the special case of our result is investigated carefully and is shown to be consistent with previous results for SAR intensity data. This can be thought of either as a cross-validation evidence for the proposed POLSAR models or alternatively as having SAR as the special case of POLSAR.

The results so far for our models can be summarized as:

$$\mathbb{R} = \frac{|C_v|}{|\Sigma_v|} \sim \frac{1}{(2L)^d} \prod_{i=0}^{d-1} \chi^2(2L - 2i) \quad (35)$$

$$\mathbb{L} = \ln |C_v| - \ln |\Sigma_v| \sim \sum_{i=0}^{d-1} \Lambda(2L - 2i) - d \cdot \ln 2L \quad (36)$$

$$\mathbb{D} = \ln |C_v| - \text{avg}(\ln |C_v|) \sim \sum_{i=0}^{d-1} \Lambda(2L - 2i) - d \ln 2 + k \quad (37)$$

$$\mathbb{C} = \ln |C_{1v}| - \ln |C_{2v}| \sim \sum_{i=0}^{d-1} \Delta(2L - 2i) \quad (38)$$

$$\mathbb{A} = \text{avg}(\mathbb{L}) = \sum_{i=0}^{d-1} \psi^0(L - i) - d \cdot \ln L \quad (39)$$

$$\mathbb{V} = \text{var}(\mathbb{L}) = \sum_{i=0}^{d-1} \psi^1(L - i) \quad (40)$$

$$\mathbb{E} = \text{mse}(\mathbb{L}) = \left[ \sum_{i=0}^{d-1} \psi^0(L - i) - d \cdot \ln L \right]^2 + \sum_{i=0}^{d-1} \psi^1(L - i) \quad (41)$$

Upon setting  $d = 1$  into the above models, Appendix C shows that the reduced results are consistent with not only the following results from the our previous works on single-look

SAR [2], i.e.  $d = L = 1$ ,

$$I \sim \bar{I} \cdot \text{pdf} [e^{-R}]$$

$$\log_2 I \sim \log_2 \bar{I} + \text{pdf} [2^x e^{-2^x} \ln 2]$$

$$\mathbb{R} = \frac{I}{\bar{I}} \sim \text{pdf} [e^{-x}]$$

$$\mathbb{L} = \log_2 I - \log_2 \bar{I} \sim \text{pdf} [2^x e^{-2^x} \ln 2]$$

$$\mathbb{D} = \log_2 I - \text{avg}(\log_2 I) \sim \text{pdf} [e^{-(2^x e^{-\gamma})} 2^x e^{-\gamma} \ln 2]$$

$$\mathbb{C} = \log_2 I_1 - \log_2 I_2 \sim \text{pdf} \left[ \frac{2^x}{(1 + 2^x)^2} \ln 2 \right]$$

$$\mathbb{A} = \text{avg}(\mathbb{L}) = -\gamma / \ln 2$$

$$\mathbb{V} = \text{var}(\mathbb{L}) = \frac{\pi^2}{6} \frac{1}{\ln^2 2}$$

$$\mathbb{E} = \text{mse}(\mathbb{L}) = \frac{1}{\ln^2 2} (\gamma^2 + \pi^2/6) = 4.1161$$

but also the following well-known results for multi-look SAR, i.e.  $d = 1, L > 1$ :

$$I \sim \text{pdf} \left[ \frac{L^L x^{L-1} e^{-Lx/\bar{I}}}{\Gamma(L) \bar{I}^L} \right] \quad (42)$$

$$N = \ln I \sim \text{pdf} \left[ \frac{L^L}{\Gamma(L)} e^{L(x-\bar{N})-Le^{x-\bar{N}}} \right] \quad (43)$$

Furthermore, the following results for multi-look SAR data, which can be thought of either as extensions of the corresponding single-look SAR results or as simple cases of the POLSAR results are also derived as:

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

$$\mathbb{L} = \ln I - \ln \bar{I} \sim \text{pdf} \left[ \frac{L^L e^{Lx-Le^x}}{\Gamma(L)} \right]$$

$$\mathbb{D} = \ln I - \text{avg}(\ln I) \sim \text{pdf} \left[ \frac{e^{L[x-\psi^0(L)]-e^{[x-\psi^0(L)]}}}{\Gamma(L)} \right]$$

$$\mathbb{C} = \ln I_1 - \ln I_2 \sim \text{pdf} \left[ \frac{e^x}{(1 + e^x)^2} \right]$$

$$\mathbb{A} = \text{avg}(\mathbb{L}) = \psi^0(L) - \ln L$$

$$\mathbb{V} = \text{var}(\mathbb{L}) = \psi^1(L)$$

$$\mathbb{E} = \text{mse}(\mathbb{L}) = [\psi^0(L) - \ln L]^2 + \psi^1(L)$$

This newly derived models for multi-look SAR data can also be validated against real-life data. Fig. 4 presents the the results of an experiment carried out for the stated purpose. In the experiment the intensity of a single-channel SAR data (HH) for a homogeneous area in the AIRSAR Flevoland dataset is extracted. Then the histograms for the log-distance and and contrast is plotted against the theoretical PDF given above. The ENL is set to the nominal number of 4. And good visual match is apparent in the final results.

## VI. HANDLING CERTAIN DISCREPANCY BETWEEN THE THEORETICAL MODEL AND PRATICAL DATA

Even though the initial assumptions for the proposed theory is intentionally kept to minimal, like all other similar models,

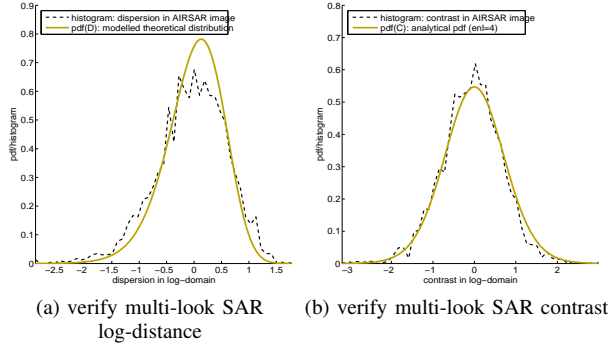


Fig. 4: Multi-Look SAR dispersion and contrast: modelled response matches very well with real-life captured data.

the proposed model in this paper is built upon certain presumptions. Practical conditions however may not always satisfy these prerequisites. In this section, certain gaps between the conditions found in practical real-life data and the theoretical assumptions are discussed. And it is shown that: the theoretical model proposed can handle the practical data, even when these “imperfections” are taken into account.

There are two main “imperfections” that are usually found in practical POLSAR data with reference to the theoretical model. The first is the mutually independent assumption for each component in the POLSAR target vector  $s$ . Practically however high correlation is routinely observable between the POLSAR data components, specifically between  $S_{hh}$  and  $S_{vv}$ . This phenomena also presents in our AIRSAR dataset, where  $\Sigma_v = \text{avg}(C_v) =$

$$\begin{bmatrix} 0.0084 & 1 \cdot 10^{-6} + 4 \cdot 10^{-4}i & 0.0071 - 0.0017i \\ 1 \cdot 10^{-6} - 4 \cdot 10^{-4}i & 0.0017 & -3 \cdot 10^{-4} - 2 \cdot 10^{-4}i \\ 0.0071 + 0.0017i & -3 \cdot 10^{-4} + 2 \cdot 10^{-4}i & 0.0122 \end{bmatrix}$$

Despite the mismatch, astute reader would probably notice that the proposed model apparently still valid under such condition as evidenced by the part-pol (HH-VV) and full-pol plots in Figs. 1 to 3. This suggests that the proposed model is also applicable on correlated POLSAR components, even though a full explanation for this, however, is outside the scope of this paper.

Another assumption of the model (for both SAR and POLSAR) is that the samples are statistically independent to each other. This is a reasonable assumption given that the transmission and receipt of analog signals are done independently for each radar pulse, i.e. for each resolution cell. Thus theoretically speaking, adjacent pixels in an image can be assumed to be statistically independent.

However, the actual imaging mechanism of a real-life (POL)SAR processor is that of digital nature, where the analog signal is to be converted into a digital data-set. Specifically, the analog signal in SAR which is characterized by the pulse bandwidth measurement, is fed into an analog-to-digital sampling and conversion process which is characterized by its sampling rate. Theoretically it is possible to define a sampling rate to ensure that each digital pixel correspond exactly to an analog physical cell resolution. Practically however, to ensure “perfect reconstruction” of the analog signal, the sampling rate

is normally set at a slightly higher value than the theoretical mark. This results in a higher number of samples / pixels than the number of physical cells available in the scene.

Stated differently, in practical data each physical radar cell is spread to more than one pixel and each pixel now contains less than one physical analog cell. This high sampling rate also results in a significantly higher correlation between pairs of pixels that maybe related within a single physical cell resolution than the correlation found between pairs of pixels that are further away and hence having less physical relation to each other. It also results in reduced effective number of look, in which say a window of 3x3 pixel actually contains less than 9 physical analog cells. The former phenomena is partially explained in [3] for SAR, while the later is experimental observed for POLSAR data in [4] and [5]. The oversampling practice is also documented by the producers of SAR processors. For AIRSAR, the sampling rate and pulse bandwidth combinations are either 90/40MHz or 45/20MHz [6]. While for RadarSat2, the pixel resolution and range - azimuth resolutions for SLC fine-quad mode is advertised as  $(4.7 \cdot 5.1)m^2 / (5.2 \cdot 7.7)m^2$  [7]

The proposed model can handle this imperfection that is available in practical data. The trick is that instead of using the nominal Number of Look given by the SAR processor, an ENL estimation procedure is employed. The first part of this section details a simple ENL estimation technique for POLSAR data. While the second part demonstrates how the practical imperfection manifest itself and how it can be handled in a RADARSAT2 dataset. It also illustrates how the match shown in Section IV for the AIRSAR Flevoland dataset can be improved, by using ENL estimation.

#### A. ENL Estimation

This sub-section describes a few technique to estimate the Effective Number of Look (ENL) for a given POLSAR dataset. The common approach in tackling this problem is by investigating the summary statistics of an known homogeneous area in the given data before making inferences about the inherent ENL. The summary statistics for  $|C_v|$  and  $\ln |C_v|$  has been derived in Section II. In fact Eqn. 20 indicates that there is a relationship among  $|\text{avg}(C_v)|$ ,  $\text{avg}(\ln |C_v|)$ ,  $d$ ,  $L$ . Recall that in carrying out validation process for AIRSAR Flevoland data using the nominal look number of 4, this relationship was broken. The reason for such broken relationship is believed to be an inexact look number was taken. In a given POLSAR dataset, since all values of  $|\text{avg}(C_v)|$ ,  $\text{avg}(\ln |C_v|)$ ,  $d$  is known, it is possible to estimate the “effective” number of look, by finding  $L$  that ensure the above relationship is valid.

In fact, this approach was taken in [5], where an equation of exactly the same form as Eqn. 20 was used to estimate the ENL. Unfortunately, the only known way to solve the equation for the unknown  $L$  requires the use of an “iterative numerical method”. Instead of relying on the equations for statistical mean to find ENL, our approach make used variance statistics in the homoskedastic log-domain to find ENL. Since the determinant of POLSAR covariance matrix can be considered as the equivalence of the intensity in SAR data, this approach

can be considered as the generic extension of our previous work on SAR ENL estimation [2] towards POLSAR data. Specifically, Eqn. 21 can be rewritten as:

$$\text{var}[\ln|C_v|] = f(L) = \sum_{i=0}^{d-1} \psi^1(L-i) \quad (44)$$

where  $\psi^1()$  denotes tri-gamma function.

Thus theoretically, given some measurable value for  $\text{var}[\ln|C_v|]$ , one could solve the above equation for the unknown  $L$ . Interestingly that would also require some iterative computations. Practically however, the shape of the right-hand-side can be pre-computed and for each computed value of  $\text{var}[\ln|C_v|]$ , a corresponding value for  $L$  can be found by referencing the variance value on the pre-computed graph. And if a graph is too tedious to be carried around an approximation can be made using a back-of-the-envelope calculation given as:

$$\hat{L} = d \left( \frac{1}{\text{var}(\ln|C_v|)} + 0.5 \right) \quad (45)$$

Fig. 5 not only shows the shapes of the function defined in Eqn. 44 for SAR and partial-POLSAR data  $f_{d=1}(L)$  and  $f_{d=2}(L)$  but also illustrates the approximation power of the simplified formula (Eqn. 45).

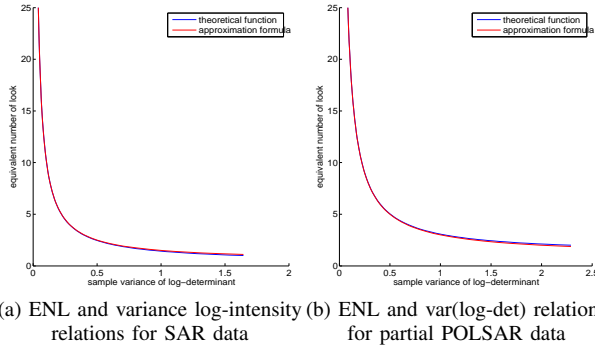


Fig. 5: The relations between ENL and sample variance of log-determinant/log-intensity

### B. Using estimated ENL to better explain practical data

For this experiment an example RADARSAT2 dataset was used. The dataset is in its Fine-Quad mode Single-Look format, and nine-look processing is applied before the dispersion histogram in log-transformed domain is computed for an homogeneous area. The histograms for both one-dimensional SAR and two-dimensional partial POLSAR data are plotted in Fig. ?? against the theoretical model for the nominal ENL value of 9. The match however is not very good.

A better match can be achieved by estimating the ENL from the observable variance of the log-determinant, and the theoretical model is simulated for the estimated ENL. Then the new sample histogram is plotted in the same figure, which indicated a much better consistency. This procedure can always be carried out for any given dataset, as long as an area of homogeneous area can be extracted.

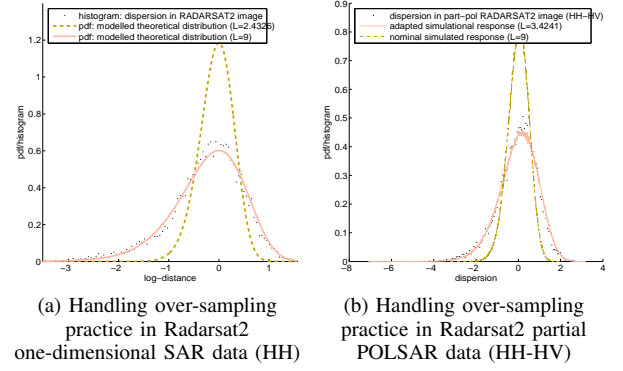


Fig. 6: 9-look processed Radarsat2 data do not exactly exhibit 9-look data characteristics. Homoskedastic model in log-transformed domain can successfully estimate the effective ENL and then explain the data reasonably well.

Fig. 7 shows that the over-sampling issue is also present in AIRSAR Flevoland dataset, even though to a much lesser extent. The “corrected” effective number-of-look offer observably better match between the model and real-life data. The mis-match problem appears to depend on how much over-sampling were used in the dataset as well as the dimension of the data.

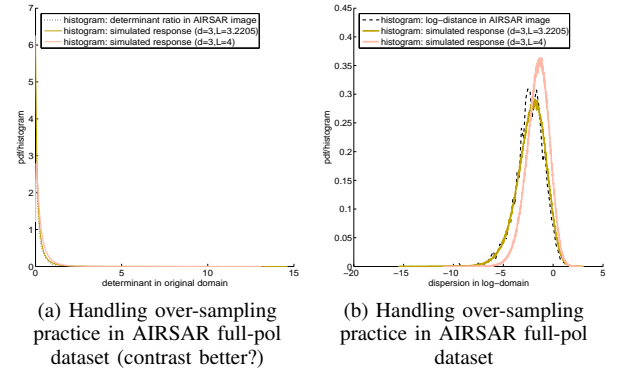


Fig. 7: AIRSAR Flevoland also exhibits phenomena of over-sampling practice, though at a lesser extent than the RADARSAT 2 data.

## VII. EVALUATING POLSAR SPECKLE FILTERS USING THE CONSISTENT MEASURES OF DISTANCE

Previous sections have developed a model for polsar, which is also shown to be applicable to sar, in this section, the use of consistent measures of distance in the context of polsar speckle filtering is briefly explored. It has been found [8] that for sar data set that in its original domain: ratio is a better evaluation than standard subtractive residual. However, ratio residual is argued as not being natural for digital display [9]. Our previous work in the context of sar speckle filtering found that in log-transformed domains, this ratio is transformed into a subtractive residual that is homoskedastic.



The variance of sample log-determinant is shown linked to the ENL index. This form the basis for evaluating POLSAR speckle filters over homogeneous areas. The procedure is simple. To evaluate a given POLSAR speckle filter over homogeneous areas, the filter is applied over any known homogeneous areas and the sample variance of log-determinant is measured. The Equivant Number of Look (ENL) is then estimated either by referencing the prepared graphs given by Eqn. 44 or alternatively by setting the measured variance value into  $\text{var}[\ln |C_v|]$  in Eqn. 45.

In order for such a procedure to be generic enough, it is important that the given POLSAR speckle filter preserve the consistency property in the log-transformed domain. That can be tested by applying the POLSAR filter into different sets of homogeneous area and investigate the plots of the dis-similarity measures presented above. Fig. 8 presents two example plots to show that: the 3x3 POLSAR boxcar filter preserves the consistency property. The boxcar filter is applied into 2 sets of part-pol AIRSAR data over Flevoland (HH-HV and VH-VV). Log-determinant and the contrast measure is computed for the inputs and outputs filtered POLSAR data, and their plots are presented. In fact, the test procedure can be applied on any of the models presented above.

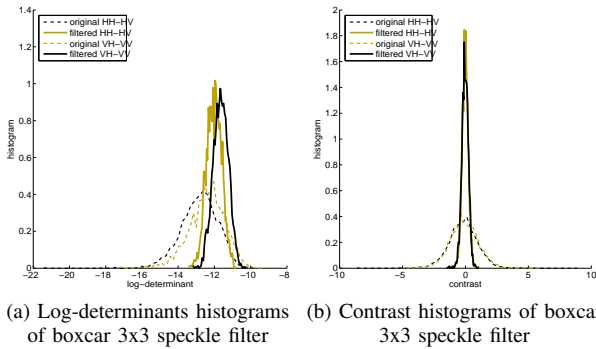


Fig. 8: POLSAR 3x3 boxcar filter preserves the consistency property. Consistency means: as long as the area is homogeneous, regardless of the underlying signal  $\Sigma_v$  the shapes of the histograms should be the same.

The consistency property of POLSAR speckle filter is important not only to make the estimation of ENL become general enough but also to ensure that any classification / detection algorithm which is based on the scalar and consistent measures of distance would work on both pre-filtered and post-filtered data. Otherwise if a polsar speckle filter gives different plots for different homogeneous areas, then not only its enl estimation will be dependent on the underlying signal, But its output also shall not follow the statistical distribution family that characterize multi-look polsar. Thus the preservation of this consistency is believed to be an important consideration for polsar speckle filter if he want a host of detection and classification algorithms to work on his filtered data output.

In evaluation over heterogeneous area, the consistent measures of distance may also be an invaluable tool. For a start, since the model for log-determinant is additive and homoskedastic, log-determinant images may be better naturally

suited for gray-level digital images. specifically for evaluation of statistical estimators, It is both important and convinient to investigate the estimators' error / residual image.

For future further analysis, the residual is defined here as the distance between the log-determinants of the filtered outputs and the original input. Ideally speaking, under the context of additive model, a perfect estimator's residual should be consists only of random noise. And under the assumption of homoskedasticity, the Gauss Markov theorem becomes applicable. Thus the optimal estimator is expected to exhibits minimal Mean Squared Error. In homogeneous scene, this is reflected in the expectation of minimal bias and variance (hence maximal ENL). Over heterogeneous scene, where the underlying signal is not known apriori, the second best gauge is possibly to have the residual mse being as close as possible to the MSE of the inherent noise.

To illustrate the above analysis, an experiment is carried out to evaluate the performance of boxcar 3x3 and boxcar 5x5 POLSAR filter on the AIRSAR Flevoland partial polarimetric data (HH-HV). A square 700x700 portion of the AIRSAR dataset is extracted, and the two POLSAR speckle filters is applied on to the patch. Then the log-determinant images of the filtered outputs are display in Fig. 9. At the same time the residual is computed for both cases filters and the images are also displayed in the same figure. Assumming the quantitative evaluation of SAR speckle filters can also be extended to POLSAR speckle filters, the Mean Squared Error (MSE) of the filters residuals are computed and compared with the "optimal" value. This value is computed by setting  $d = 2, L = 4$  into Eqn. C.40 making the expected MSE being  $\text{mse}(\mathbb{L}) = 1.0132$ .

Fig. 9 shows that not only the log-determinant image can offer a nice visualization of the scene, but also the distortion impact of the filter can also be made visible by the residual image. In visual evaluation, while it is quite hard to observe the worsening blurring-effects of the boxcar 5x5 speckle filter as compared to the 3x3 filter in the additive log-determinant image of the filtered output such a conclusion can be made relatively easier by visually investigating the residual image.

And when quantified evaluation is carried out where the residual mse is compared with the expected level of noise to be removed, the excessive blurring effects of the 5x5 filter become crystallly clear. In fact, even the 3x3 boxcar filter itself might be also a bit blurry. A conclusion which is hard to make just by looking at the filtered imagery. However, by investigating the residual between the unfiltered input and the filtered results in the additive and homoskedastic model, both visual and quantitative evaluations offer conclusive evidence.

Much more detailed discussion on this topic, under the context of evaluating sar speckle filter is published in our previous work [2]. Unfortunately, due to space restriction, only a brief and critical exploration is explored here for polsar speckle filters. This section serves more as an illustration of how the proposed theoretical model can be applied on practical scenario, as it is not as a full proposal for such an evaluation procedure for polsar speckle filters.



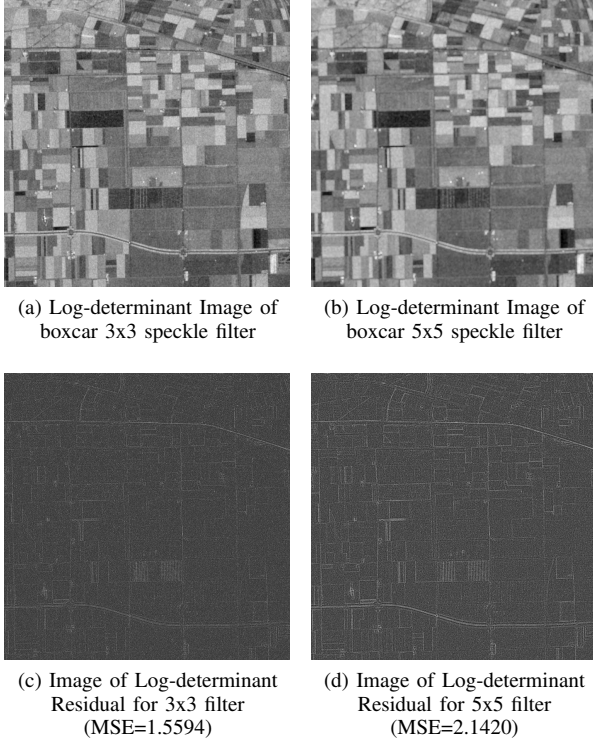


Fig. 9: Visually Evaluating POLSAR Boxcar 3x3 vs. 5x5 Speckle Filters on AIRSA Flevoland part-pol data (HH-HV) with expected MSE=1.0312 at ENL=4.

## VIII. DISCUSSION AND CONCLUSION

### A. Related Work in Literature

This sub-section reviews some relevant and published works on proposing different measures of distance for POLSAR data. Specifically a few different matrix distances have been proposed and evaluated in recent review papers [10] [11]

The common used measure of distance for matrices are either the Euclidean or the Manhattan distance. They are defined in the following equations, respectively The Manhattan distance is

$$d(C_x, C_y) = \sum_{i,j} |\Re(C_x - C_y)_{i,j}| + \sum_{i,j} |\Im(C_x - C_y)_{i,j}| \quad (46)$$

$$d(C_x, C_y) = \sqrt{\sum_{i,j} |C_x - C_y|_{i,j}^2} \quad (47)$$

where  $A_{i,j}$  denotes the  $(i,j)$  elements of the matrix  $A$ ,  $||$  denotes absolute values and  $\Re, \Im$  denote the real and imaginary parts respectively. In the context of polsar covariance matrix, however, they are not widely used probably because of the multiplicative nature of the data.

In the field of POLSAR, the Wishart Distance is probably the most widely used as part of the well-known Wishart Classifier [12]. The distance is defined as [13]

$$d(C_x, C_y) = \ln |C_y| + \text{tr}(C_x C_y^{-1}) \quad (48)$$

As a measure of distance, its main disadvantage is that  $d(C_y, C_x) = \ln |C_x| \neq 0$ .

Recent works have suggested a number of other dissimilarity measures. They are: the asymmetric and the symmetric refined Wishart distance [14],

$$d(C_x, C_y) = \frac{1}{2} \text{tr}(C_x^{-1} C_y + C_y^{-1} C_x) - d \quad (49)$$

$$d(C_x, C_y) = \ln |C_x| - \ln |C_y| + \text{tr}(C_x C_y^{-1}) - d \quad (50)$$

the Bartlett distance [11],

$$d(C_x, C_y) = 2 \ln |C_{x+y}| - \ln |C_x| - \ln |C_y| - 2d \ln 2 \quad (51)$$

the Bhattacharyya distance [15],

$$r(C_x, C_y) = \frac{|C_x|^{1/2} |C_y|^{1/2}}{|(C_x + C_y)/2|} \quad (52)$$

and the Wishart Statistical test distance [16]

$$d(C_x, C_y) = (L_x + L_y) \ln |C| - L_x \ln |C_x| - L_y \ln |C_y| \quad (53)$$

In comparison to these formulae, the distance measure for two covariance matrix proposed in this paper can be written as:

$$d(C_x, C_y) = \ln |C_x| - \ln |C_y| \quad (54)$$

At the first glance, while the proposed formula may be a lot simpler, it is still very similar to others in making use of the log-determinant.

Closer investigation of these dissimilarity measures reveals that most of them are related to each other. The Bhattacharyya distance is easily shown to be related to the Bartlett distance. At the same time the Bartlett distance can be considered as the special case of the Wishart Statistical Test distance, where the two data sets have the same number of looks  $L_x = L_y$ . And it is also intuitively trivial to arrive at the conclusion that the contrast measure of distance proposed in this paper has fixed statistical behaviour just by comparing the symmetric and asymmetric versions of the Refined Wishart distance, assuming that they are shown to follow fixed distributions.

The close relation among these is further supported by the fact that in fact, in their proposal, all of the work referenced the statistical model developed in this paper as their foundation. Interestingly, the change detection model also makes use of determinant ratio and log determinant in its derivation. The next sub-section will illustrate how the log-determinant model studied in this paper can be an alternative and simple derivation for the change-detection statistics.

### B. The Likelihood Statistics for POLSAR

To detect if the two scaled multi-look POLSAR covariance matrix  $Z_x$  and  $Z_y$ , which have  $L_x$  and  $L_y$  as the corresponding number of looks, come from the same underlying stochastic process, the likelihood ratio statistics for POLSAR covariance matrix is considered [17]

$$Q = \frac{(L_x + L_y)^{d \cdot (L_x + L_y)}}{L_x^{d \cdot L_x} L_y^{d \cdot L_y}} \frac{|Z_x|^{L_x} |Z_y|^{L_y}}{|Z_x + Z_y|^{(L_x + L_y)}} \quad (55)$$

Taking the log-transformation of the above statistics, and note that  $C_{vx} = Z_x/L_x$ ,  $C_{vy} = Z_y/L_y$  and  $C_{vxy} = (Z_x + Z_y)/(L_x + L_y)$  it becomes

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

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

To detect changes, a test statistics is developed based on this measure of distance. This means a distribution is to be derived for the dissimilarity measure. However, originally in the proposed work, only an asymptotic distribution is derived. In this sub-section, a statistical model is developed for this measure of distance.

Since both  $Z_x$  and  $Z_y$  follow complex wishart distribution with  $L_x$  and  $L_y$  degrees of freedom,  $Z_x + Z_y$  also follows the complex wishart distribution with  $L_x + L_y$  degrees of freedom. In view of the models denoted in Eqn. 9, it is evident that not only the bound for  $\ln Q$ , or equivalently  $Q$ , can be derived but the whole statistical distribution for it can be simulated as well:

$$\ln Q \sim k + L_x \Lambda_{L_x}^d + L_y \Lambda_{L_y}^d - (L_x + L_y) \Lambda_{(L_x+L_y)}^d \quad (56)$$

$$Q \sim e^k \frac{(\chi_{L_x}^d)^{L_x} \cdot (\chi_{L_y}^d)^{L_y}}{(\chi_{L_x+L_y}^d)^{L_x+L_y}} \quad (57)$$

where  $k = d[(L_x + L_y) \ln(L_x + L_y) - L_x \ln L_x - L_y \ln L_y]$ .

In the common case of  $L_x = L_y$ , the test statistics become  $\ln Q = \ln |C_{vx}| + \ln |C_{vy}| - 2 \ln |C_{vxy}|$ . This and the barlett distance introduced earlier are closely related, in fact they are just shifted versions of each other. under exactly the same assumptions, The contrast model which is written as  $\mathbb{C} = \ln C_{vx} - \ln C_{vy}$  and presented above should provides another consistent and equivalent approach, with probably simpler conceptual model and computational derivations. The details application of this in edge-detection however is outside the scope of this paper.

### C. Discussion

Let us start the discussion by noting a few theoretical properties of the proposed statistical model First, The use of covariance matrix log-determinant may be related to the standard eigen-decomposition method of the second order statistics POLSAR matrices. In fact, the log-determinant can also be computed as the sum of log-eigen values. Specifically  $\ln |M| = \sum \ln \lambda_M$  where  $\lambda_M$  denotes all the eigen-values of  $M$ . Thus similar to other eigen-value based approach (e.g. Entropy/Anisotropy, ...), the models presented here is invariant to polarization basis transformations.

Second, the model is developed for the polsar covariance matrix. However, since the polsar coherent matrix is related with the covariance matrix via an unitary transformation, which preserves the determinant as invariant, The model should also be applicable on the coherency matrix.

The model is far from complete. It calls for the reduction of the multi-dimensional POLSAR data into a scalar value. While this is probably desirable for a wide range of application where a one-dimensional number is required to represents the

complex multi-dimensional data, such a reduction is probably not lossless. Thus similar to the way the Wishart Classifier is employed, to better understand polsar data the use of this technique should be complemented with some high-dimensional POLSAR target-decomposition techniques (e.g. Freeman Durden or entropy/anisotropy TODO:CITE ...)

But at the same time, the model is also promising. The models presented in this paper are first developed for partial and monostatic POLSAR data, It is then shown to be also applicable to the traditional sar data. Since the model's assumptions are quite minimal, It may also work on bi-static and intererometric data, though that requires further investigations. Other interesting phenomena which may warrant more study include the applicability of the model on correlated polsar channels ( $S_{hh}, S_{vv}$ ) as well as a better explanation in the use of mse to evaluate polsar speckle filters.

Many practical applications may also be developed using the model and its derivatives. First, The model allow for the estimation and simulation of non-natural equivalent number of look. This come in especially handy when some real-life data is given, where the effective number of look may not be equals to the nominal stated number-of-look after multi-look-processing. In such case, better-match models can be derived by roughly estimating the effective ENL which usually results in a number with fractions.

Similar to the way other measures of distance can be used to derive POLSAR classifier [12], change detector [17], edge detector [18] or other clustering cite and speckle filtering techniques [19] [20] new detection / classification, clustering or speckle filtering algorithms can be derived using the models presented in this paper. And since extensive evidence have been shown supporting the use of MSE for SAR data, it is reasonable to expect the relevant of MSE for POLSAR data. If that is the case, a large number of existing algorithms can become applicable to the POLSAR model in log-transformed domain, which has been shown to be both additive and homoskedastic.

### D. Conclusion

In conclusion, an additive and homoskedastic model, which results in several scalar and consistent measures-of-distance for multi-variate POLSAR data were proposed in this paper. The theoretical model is shown to be powerful. Not only it can provide alternative and sometimes simpler explanations to a range of theoretical concepts: i.e. POLSAR test statistics or ENL estimation but it also puts several well-known models for the traditional SAR within its natural coverage. The statistical model is based on the determinant of the POLSAR covariance matrix, which when converted into one-dimensional data is gracefully transformed into the traditional SAR intensity. Consequently, the derived dis-similarity measures may be employed in a wide range of application where a scalar number is required to represent the complex multi-dimensional POLSAR data. The model is also shown to be practically versatile. It is capable of handling a few imperfect conditions found in practical data. As both an application of the model, as well as an extension of our previous work in evaluating SAR speckle

filters [19], the application of these additive and homoskedastic distances in the context of evaluating POLSAR speckle filters is briefly explored. And intially promising results are reported.

## APPENDIX A

### HOMOSKEDASTIC MODEL FOR THE LOG-DETERMINANT

#### A. Log-Chi-Square Distribution and its Derivatives

This section provides the mathematical derivations for the log-transformed version of chi-squared random variables.

Chi-squared random variables  $\chi \sim \chi^2(k)$  follows the pdf:

$$pdf(\chi; k) = \frac{\chi^{(k/2)-1} e^{-\chi/2}}{2^{k/2} \Gamma(\frac{k}{2})} \quad (\text{A.1})$$

Setting  $L=k/2$  into Eqn. A.1

$$pdf(\chi) = \frac{\chi^{L-1} e^{-\chi/2}}{2^L \Gamma(L)} \quad (\text{A.2})$$

Applying the variable change theorem, which states that: if  $y = \phi(x)$  with  $\phi(c) = a$  and  $\phi(d) = b$ , then:

$$\int_a^b f(y) dy = \int_c^d f[\phi(x)] \frac{d\phi}{dx} dx \quad (\text{A.3})$$

into the log-transformation, which changes the random variables  $\Lambda = \ln(\chi)$ , we have:

$$\begin{aligned} d\chi &= e^\Lambda d\Lambda \\ \frac{\chi^{L-1} e^{-\chi/2}}{2^L \Gamma(L)} d\chi &= \frac{(e^\Lambda)^{L-1} e^{-e^\Lambda/2}}{2^L \Gamma(L)} e^\Lambda d\Lambda \end{aligned}$$

In other words, we have:

$$pdf(\Lambda; L) = \frac{e^{L\Lambda - e^\Lambda/2}}{2^L \Gamma(L)} \quad (\text{A.4})$$

From the PDF given in Eqn. A.4, characteristics function can be computed. By definition, the characteristic function (CF)  $\varphi_X(t)$  for a random variable  $X$  is computed as:

$$\begin{aligned} \varphi_X(t) = E[e^{itX}] &= \int_{-\infty}^{\infty} e^{itx} dF_X(x) \\ &= \int_{-\infty}^{\infty} e^{itx} f_X(x) dx \end{aligned}$$

with  $\varphi_X(t)$  is the characteristic function,  $F_X(x)$  is the CDF function of  $X$  and  $f_X(x)$  is the PDF function of  $X$ . Thus the characteristic function for log-chi-squared distribution is defined as:

$$\varphi_\Lambda(t) = \int_0^\infty e^{itx} \frac{e^{Lx - e^x/2}}{2^L \Gamma(L)} dx \quad (\text{A.5})$$

Gamma function is defined over complex domain as:  $\Gamma(z) = \int_0^\infty e^{-x} x^{z-1} dx$ . Thus  $\Gamma(L + it) = \int_0^\infty e^{-x} x^{L+it-1} dx$ . Set  $x = e^z/2$  then  $dx = e^z/2 dz$ , we have  $\Gamma(L + it) = \int_0^\infty e^{itz} \frac{e^{Lz - e^z/2}}{2^{L+it}} dz$

That is:

$$\varphi_\Lambda(t) = 2^{it} \frac{\Gamma(L + it)}{\Gamma(L)} \quad (\text{A.6})$$

Consequently, the first and second derivative of log-chi-squared distribution can be computed. The first derivative is given as:

$$\frac{\partial \varphi_\Lambda(t)}{\partial t} = \frac{i2^{it} \Gamma(L + it)}{\Gamma(L)} [\ln 2 + \psi^0(L + it)] \quad (\text{A.7})$$

due to

$$\begin{aligned} \frac{\partial \Gamma(x)}{\partial x} &= \Gamma(x) \psi^0(x), \\ \frac{\partial \Gamma(L + it)}{\partial t} &= i \Gamma(L + it) \psi^0(L + it), \\ \frac{\partial 2^{it}}{\partial t} &= i2^{it} \ln(2), \\ \partial(u \cdot v)/\partial t &= u \cdot \partial v/\partial t + v \cdot \partial u/\partial t, \end{aligned}$$

where  $\psi^0()$  denotes di-gamma function.

Meanwhile, the second derivative can be written as:

$$\frac{\partial^2 \varphi_\Lambda(t)}{\partial t^2} = \frac{i^2 2^{2it} \Gamma(L + it)}{\Gamma(L)} \left( [\ln 2 + \psi^0(L + it)]^2 + \psi^1(L + it) \right) \quad (\text{A.8})$$

due to:

$$\begin{aligned} \frac{d2^{it} \Gamma(L + it)}{dt} &= i2^{it} \Gamma(L + it) [\ln 2 + \psi^0(L + it)], \\ \frac{d\psi^0(t)}{dt} &= \psi^1(t), \\ \frac{d\psi^0(L + it)}{dt} &= i\psi^1(L + it), \\ \partial(u \cdot v)/\partial t &= u \cdot \partial v/\partial t + v \cdot \partial u/\partial t, \end{aligned}$$

with  $\psi^1()$  denotes tri-gamma function.

The  $n^{th}$  moments of random variable  $X$  can be computed from the derivatives of its characteristic function as:

$$E(\Lambda^n) = i^{-n} \varphi_\Lambda^{(n)}(0) = i^{-n} \left[ \frac{d^n}{dt^n} \varphi_\Lambda(t) \right]_{t=0} \quad (\text{A.9})$$

Thus

$$\begin{aligned} E(\Lambda) &= i^{-1} \left[ \frac{d\varphi_\Lambda(t)}{dt} \right]_{t=0} \\ &= i^{-1} \left[ \frac{i2^{it} \Gamma(L + it)}{\Gamma(L)} [\ln 2 + \psi^0(L + it)] \right]_{t=0} \end{aligned}$$

That is

$$avg(\Lambda) = \psi^0(L) + \ln(2) \quad (\text{A.10})$$

Similarly,

$$\begin{aligned} E(\Lambda^2) &= i^{-2} \left[ \frac{d^2 \varphi_\Lambda(t)}{dt^2} \right]_{t=0} \\ &= \left[ \frac{2^{2it} \Gamma(L + it)}{\Gamma(L)} \left( [\ln 2 + \psi^0(L + it)]^2 + \psi^1(L + it) \right) \right]_{t=0} \end{aligned}$$

That is:

$$E(\Lambda^2) = [\psi^0(L) + \ln(2)]^2 + \psi^1(L) \quad (\text{A.11})$$

Thus

$$var(\Lambda) = E(\Lambda^2) - E^2(\Lambda) = \psi^1(L) \quad (\text{A.12})$$

### B. Averages and Variances of POLSAR Covariance Matrix Determinant and Log-Determinant

In this section, the expected value and variance value of these mixture of random variables is derived

$$\chi_L^d \sim \prod_{i=0}^{d-1} \chi(2L - 2i) \quad (\text{A.13})$$

$$\Lambda_L^d \sim \sum_{i=0}^{d-1} \Lambda(2L - 2i) \quad (\text{A.14})$$

given the averages and variances of individual components.

$$\text{avg} [\chi(2L)] = 2L \quad (\text{A.15})$$

$$\text{var} [\chi(2L)] = 4L \quad (\text{A.16})$$

$$\text{avg} [\Lambda(2L)] = \psi^0(L) + \ln 2 \quad (\text{A.17})$$

$$\text{var} [\Lambda(2L)] = \psi^1(L) \quad (\text{A.18})$$

Making use of the mutual indepent property of each component  $X_i$ , the variance and expectation of the sumation and product of random variables can be written as:

$$\begin{aligned} \text{avg} \left( \sum_{i=1}^n X_i \right) &= \sum_{i=1}^n \text{avg}(X_i), \\ \text{var} \left( \sum_{i=1}^n X_i \right) &= \sum_{i=1}^n \text{var}(X_i), \\ \text{avg} \left( \prod_{i=1}^n X_i \right) &= \prod_{i=1}^n \text{avg}(X_i), \\ \text{var} \left( \prod_{i=1}^n X_i \right) &= \prod_{i=1}^n [\text{avg}^2(X_i) + \text{var}(X_i)] - \prod_{i=1}^n \text{avg}^2(X_i). \end{aligned}$$

Thus they can be computed as:

$$\begin{aligned} \text{avg} [\chi_L^d] &= 2^d \cdot \prod_{i=0}^{d-1} (L - i), \\ \text{var} [\chi_L^d] &= \prod_{i=0}^{d-1} 4(L - i)(L - i + 1) - \prod_{i=0}^{d-1} 4(L - i)^2, \\ \text{avg} [\Lambda_L^d] &= d \cdot \ln 2 + \sum_{i=0}^{d-1} \psi^0(L - i), \\ \text{var} [\Lambda_L^d] &= \sum_{i=0}^{d-1} \psi^1(L - i) \end{aligned}$$

## APPENDIX B

### DERIVING THE CHARACTERISTIC FUNCTIONS FOR THE CONSISTENT MEASURES OF DISTANCE

Given characteristic function (CF) of the elementary log-chi square distributions can be written as:

$$CF_{\Lambda(2L)}(t) = 2^{it} \Gamma(L + it) / \Gamma(L)$$

the CF for the following random variables, which are combinations of the above elementary random variables, are to be

derived

$$\begin{aligned} \Lambda_L^d &\sim \sum_{i=0}^{d-1} \Lambda(2L - 2i) \\ \mathbb{L} &\sim \Lambda_L^d - d \cdot \ln(2L) \\ \mathbb{D} &\sim \mathbb{L} - d \cdot \ln L + \sum_{i=0}^{d-1} \psi^0(L - i) \\ \mathbb{C} &\sim \sum_{i=0}^{d-1} [\Lambda(2L - 2i) - \Lambda(2L - 2i)] \end{aligned}$$

Since

$$\begin{aligned} CF_{\sum X_i}(t) &= \prod CF_{X_i}(t) \\ CF_{x+k}(t) &= e^{itk} CF_x(t) \end{aligned}$$

we have:

$$CF_{\Lambda_L^d}(t) = \frac{2^{idt}}{\Gamma(L)^d} \prod_{j=0}^{d-1} \Gamma(L - j + it) \quad (\text{B.19})$$

$$CF_{\mathbb{L}} = \frac{1}{L^{idt} \Gamma(L)^d} \prod_{j=0}^{d-1} \Gamma(L - j + it) \quad (\text{B.20})$$

$$CF_{\mathbb{D}} = \frac{1}{\Gamma(L)^d} \prod_{j=0}^{d-1} e^{idt\psi^0(L-j)} \Gamma(L - j + it) \quad (\text{B.21})$$

Also due to

$$\begin{aligned} CF_{-\Lambda(2L)}(t) &= 2^{-it} \frac{\Gamma(L - it)}{\Gamma(L)} \\ \Delta(2L) &\sim \Lambda(2L) - \Lambda(2L) \\ \Gamma(L - it) \Gamma(L + it) &= \Gamma(2L) B(L - it, L + it) \\ CF_{\Delta(2L)}(t) &= \frac{\Gamma(2L) B(L - it, L + it)}{\Gamma^2(L)} \end{aligned}$$

we arrive at:

$$CF_{\mathbb{C}} = \prod_{j=0}^{d-1} \frac{\Gamma(2L - 2j) B(L - j - it, L - j + it)}{\Gamma^2(L - j)} \quad (\text{B.22})$$

with  $\Gamma()$  and  $B()$  denotes Gamma and Beta functions respectively.

## APPENDIX C

### SAR INTENSITY AS THE SPECIAL CASE OF POLSAR COVARIANCE MATRIX DETERMINANT

In this appendix, the following results for SAR intensity  $I$  is shown to be special cases of the results given in this paper for the determinant of POLSAR's covariance matrix  $\det[C_v]$ . Specifically, not only the following results from the our previous works on single-look SAR (TODO:CITE), i.e.

$d = L = 1$ , is reviewed:

$$I \sim \bar{I} \cdot pdf[e^{-R}] \quad (C.23)$$

$$\log_2 I \sim \log_2 \bar{I} + pdf[2^{D-2^D}] \quad (C.24)$$

$$\frac{I}{\bar{I}} = \mathbb{R} \sim pdf[e^{-R}] \quad (C.25)$$

$$\log_2 I - \log_2 \bar{I} = \mathbb{D} \sim pdf[2^D e^{-2^D} \ln 2] \quad (C.26)$$

$$\log_2 I_1 - \log_2 I_2 = \mathbb{C} \sim pdf\left[\frac{2^c}{(1+2^c)^2} \ln 2\right] \quad (C.27)$$

$$avg(\mathbb{D}) = -\gamma / \ln 2 \quad (C.28)$$

$$var(\mathbb{D}) = \frac{\pi^2}{6} \frac{1}{\ln^2 2} \quad (C.29)$$

$$mse(\mathbb{D}) = \frac{1}{\ln^2 2} (\gamma^2 + \pi^2/6) = 4.10630 \quad (C.30)$$

but also the following well-known results for multi-look SAR, i.e.  $d = 1, L > 1$  is also considered:

$$I \sim pdf\left[\frac{L^L I^{L-1} e^{-LI/\bar{I}}}{\Gamma(L) \bar{I}^L}\right] \quad (C.31)$$

$$N = \ln I \sim pdf\left[\frac{L^L}{\Gamma(L)} e^{L(N-\bar{N})-L e^{N-\bar{N}}}\right] \quad (C.32)$$

It will be shown that all of these results are special cases of the result derived previously and rewritten below:

$$|C_v| \sim \frac{|\Sigma_v|}{(2L)^d} \prod_{i=0}^{d-1} \chi^2(2L-2i) \quad (C.33)$$

$$\ln |C_v| \sim \ln |\Sigma_v| + \sum_{i=0}^{d-1} \Lambda(2L-2i) - d \cdot \ln 2 \quad (C.34)$$

$$\frac{|C_v|}{|\Sigma_v|} = \mathbb{R} \sim \frac{1}{(2L)^d} \prod_{i=0}^{d-1} \chi^2(2L-2i) \quad (C.35)$$

$$\ln |C_v| - \ln |\Sigma_v| = \mathbb{D} \sim \sum_{i=0}^{d-1} \Lambda(2L-2i) - d \cdot \ln 2 \quad (C.36)$$

$$\ln |C_{1v}| - \ln |C_{2v}| = \mathbb{C} \sim \sum_{i=0}^{d-1} \Delta(2L-2i) \quad (C.37)$$

$$avg(\mathbb{D}) = \sum_{i=0}^{d-1} \psi^0(L-i) - d \cdot \ln L \quad (C.38)$$

$$var(\mathbb{D}) = \sum_{i=0}^{d-1} \psi^1(L-i) \quad (C.39)$$

$$mse(\mathbb{D}) = \left[ \sum_{i=0}^{d-1} \psi^0(L-i) - d \cdot \ln L \right]^2 + \sum_{i=0}^{d-1} \psi^1(L-i) \quad (C.40)$$

This appendix also derives new results for multi-look SAR data, which can be thought of either as extensions of the corresponding single-look SAR results or as simple cases of

the POLSAR results presented above. They are:

$$\frac{I}{\bar{I}} = \mathbb{R} \sim \frac{1}{2L} \chi^2(2L) \quad (C.41)$$

$$\ln I - \ln \bar{I} = \mathbb{D} \sim \Lambda(2L) - \ln 2L \quad (C.42)$$

$$\ln I_1 - \ln I_2 = \mathbb{C} \sim \Delta(2L) \quad (C.43)$$

$$avg(\mathbb{D}) = \psi^0(L) - \ln L \quad (C.44)$$

$$var(\mathbb{D}) = \psi^1(L) \quad (C.45)$$

$$mse(\mathbb{D}) = [\psi^0(L) - \ln L]^2 + \psi^1(L) \quad (C.46)$$

The derivation process detailed below consists of two-phases. The first phase collapse the generic multi-dimensional POLSAR results into the classical one-dimensional SAR domain. Mathematically this means setting the dimensional number in POLSAR to  $d = 1$  and collapsing the POLSAR covariance matrix into the variance measure in SAR, which also equals the SAR intensity i.e.  $|C_v| = I, |\Sigma_v| = \bar{I}$ .

The output of the first phase, in the general case is applicable to multi-look SAR data, where  $d = 1$  but  $L > 1$ . The second phase simplify the multi-look results into single-look results, presented in our previous work (TODO:CITE). Mathematically, that means setting  $L = 1$  in the multi-look result and converting from natural logarithmic domain used in this paper to the base-2 logarithm used in (TODO:CITE).

#### A. Original Domain: SAR Intensity and its ratio

Setting  $d = 1, |C_v| = I$  and  $|\Sigma_v| = \bar{I}$  into Eqns. C.33 and C.35 we have:

$$I \sim \frac{\bar{I}}{2L} \chi^2(2L)$$

$$\frac{I}{\bar{I}} = \mathbb{R} \sim \frac{1}{2L} \chi^2(2L)$$

Or in PDF forms, and applying variable change theorem we have:

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

$$\frac{I}{\bar{I}} \sim pdf\left[\frac{x^{L-1} e^{-x/2}}{2^L \Gamma(L)} \cdot dx/dt\right]_{x=2L \cdot t}$$

$$\sim pdf\left[\frac{L^L t^{L-1} e^{-Lt}}{\Gamma(L)}\right]$$

$$I \sim pdf\left[\frac{L^L t^{L-1} e^{-Lt}}{\Gamma(L)} \cdot dt/dx\right]_{t=x/\bar{I}}$$

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

Thus we have the following results for multi-look SAR

$$I \sim pdf\left[\frac{L^L x^{L-1} e^{-Lx/\bar{I}}}{\bar{I}^L \Gamma(L)}\right] \quad (C.47)$$

$$\frac{I}{\bar{I}} = \mathbb{R} \sim pdf\left[\frac{L^L x^{L-1} e^{-Lx}}{\Gamma(L)}\right] \quad (C.48)$$

Now setting  $L = 1$ , these results become:

$$I \sim pdf\left[\frac{e^{x/\bar{I}}}{\bar{I}}\right] \quad (C.49)$$

$$\frac{I}{\bar{I}} = \mathbb{R} \sim pdf[e^{-x}] \quad (C.50)$$

which is the same as in (TODO:CITE).

### B. Log-transformed domain: SAR log-intensity and the log-distance

The result for multi-look SAR data written in log-transformed domain can be derived from two different approaches. The first is to follow the simplification method, where the results for log-transformed POLSAR data is simplified into log-transformed multi-look SAR result.

The second approach is to apply log-transformation into the results derived in the previous section. In this section, it is shown that both approaches would results into identical results.

Setting  $d = 1$ ,  $|C_v| = I$  and  $|\Sigma_v| = \bar{I}$  into Eqns. C.34 and C.36 we have

$$\begin{aligned}\ln I &\sim \ln \bar{I} + \Lambda(2L) - \ln 2L \\ \ln I - \ln \bar{I} &\sim \Lambda(2L) - \ln 2L\end{aligned}$$

Or in PDF form, and applying variable change theorem we have:

$$\begin{aligned}\ln I - \ln \bar{I} + \ln 2L &\sim \text{pdf} \left[ \frac{e^{Lx - e^x/2}}{2^L \Gamma(L)} \right] \\ \ln I - \ln \bar{I} &\sim \text{pdf} \left[ \frac{e^{Lx - e^x/2}}{2^L \Gamma(L)} \cdot dx/dt \right]_{x=t+\ln 2L} \\ &\sim \text{pdf} \left[ \frac{L^L e^{Lt - Le^t}}{\Gamma(L)} \right] \\ \ln I &\sim \text{pdf} \left[ \frac{L^L e^{Lt - Le^t}}{\Gamma(L)} \cdot dt/dx \right]_{t=x-\ln \bar{I}} \\ &\sim \text{pdf} \left[ \frac{L^L e^{L(x-\bar{N}) - Le^{x-\bar{N}}}}{\Gamma(L)} \right]\end{aligned}$$

with  $\bar{N} = \ln \bar{I}$ .

Thus the first approach arrives at

$$\ln I = \mathbb{N} \sim \text{pdf} \left[ \frac{L^L e^{L(x-\bar{N}) - Le^{x-\bar{N}}}}{\Gamma(L)} \right] \quad (\text{C.51})$$

$$\ln I - \ln \bar{I} = \mathbb{L} \sim \text{pdf} \left[ \frac{L^L e^{Lt - Le^t}}{\Gamma(L)} \right] \quad (\text{C.52})$$

In the second approach, log-transformation is applied on previous result for multi-look SAR intensity and its ratio in the original domain (Eqns. C.48 and C.47). The also arrives at the same results as above, the details working however is omitted here for brevity.

To compute summary statistics for the multi-look SAR dispersion, set  $d = 1$  into the Eqns. C.40, C.38 and C.39 we have:

$$\begin{aligned}\text{avg}(\mathbb{L}) &= \psi^0(L) - \ln L \\ \text{var}(\mathbb{L}) &= \psi^1(L) \\ \text{mse}(\mathbb{L}) &= [\psi^0(L) - \ln L]^2 + \psi^1(L)\end{aligned}$$

This completes the first phase of the derivation process. The second phase of simplification involves setting  $L = 1$  into the above results for multi-look SAR data, and converting natural

logarithm into base-2 logarithm. First, setting  $L = 1$  makes the above results become

$$\begin{aligned}\ln I = \mathbb{N} &\sim \text{pdf} \left[ e^{(x-\bar{N}) - e^{x-\bar{N}}} \right] \\ \ln I - \ln \bar{I} = \mathbb{L} &\sim \text{pdf} \left[ e^{x - e^x} \right] \\ \text{avg}(\mathbb{L}) &= \psi^0(1) = -\gamma \\ \text{var}(\mathbb{L}) &= \psi^1(1) = \pi^2/6 \\ \text{mse}(\mathbb{L}) &= [\psi^0(1)]^2 + \psi^1(1) = \gamma^2 + \pi^2/6\end{aligned}$$

with  $\gamma$  denotes the Euler-Mascharoni constant. Then to convert to base-2 logarithm from natural logarithmic transformation, variable change theorem is invoked. That is:

$$\begin{aligned}\log_2 I = \mathbb{N}_2 &\sim \text{pdf} \left[ e^{(x-\bar{N}) - e^{x-\bar{N}}} \cdot dx/dt \right]_{x=t \cdot \ln 2} \\ \mathbb{N}/\ln 2 = \mathbb{N}_2 &\sim \text{pdf} \left[ e^{(t \cdot \ln 2 - \bar{N}) - e^{t \cdot \ln 2 - \bar{N}}} \ln 2 \right]_{\bar{N}_2 = \bar{N} \cdot \ln 2} \\ &\sim \text{pdf} \left[ 2^{t - \bar{N}_2} e^{2^{t - \bar{N}_2}} \ln 2 \right]\end{aligned}$$

$$\begin{aligned}\log_2 I - \log_2 \bar{I} = \mathbb{L}/\ln 2 = \mathbb{L}_2 &\sim \text{pdf} \left[ e^{x - e^x} \right]_{x=t \cdot \ln 2} \\ &\sim \text{pdf} \left[ 2^t e^{2^t} \ln 2 \right]\end{aligned}$$

$$\begin{aligned}\text{avg}(\mathbb{L}_2) &= \text{avg}(\mathbb{L})/\ln 2 = -\gamma/\ln 2 \\ \text{var}(\mathbb{L}_2) &= \text{var}(\mathbb{L})/\ln^2 2 = \frac{\pi^2}{6} \frac{1}{\ln^2 2} \\ \text{mse}(\mathbb{L}_2) &= \text{mse}(\mathbb{L})/\ln^2 2 = \frac{1}{\ln^2 2} (\gamma^2 + \pi^2/6) = 4.1161\end{aligned}$$

### C. Deriving the PDF for SAR dispersion and contrast

The PDF for SAR dispersion can be easily derived from the PDF for the Log-distance given above as:

$$\ln I - \text{avg}(\ln I) = \mathbb{D} \sim \text{pdf} \left[ \frac{e^{L[x + \psi^0(L)] - Le^{x + \psi^0(L) - \ln L}}}{\Gamma(L)} \right] \quad (\text{C.53})$$

due to  $d = 1$  and

$$\begin{aligned}\mathbb{D} &\sim \mathbb{L} - \text{avg}(\mathbb{L}) \\ \text{avg}(\mathbb{L}) &= \psi^0(L) - \ln L \\ \mathbb{L} &\sim \text{pdf} \left[ \frac{L^L e^{Lt - Le^t}}{\Gamma(L)} \right]\end{aligned}$$

Setting  $L = 1$  for Single-Look SAR we have

$$\mathbb{D} \sim \text{pdf} \left[ e^{x - \gamma - e^{x - \gamma}} \right] \quad (\text{C.54})$$

due to:  $\psi^0(1) = -\gamma$  and  $\Gamma(1) = 1$  with  $\gamma$  being the Euler Mascheroni Constant which equals 0.5772. In base-2 logarithm, variable change theorem is invoked

$$\begin{aligned}\mathbb{D}_2 &= \log_2 I - \text{avg}(\log_2 I) = \mathbb{D}/\ln 2 \\ \mathbb{D}_2 &\sim \text{pdf} \left[ e^{x - \gamma - e^{x - \gamma}} \cdot \frac{dx}{dt} \right]_{x=t \cdot \ln 2}\end{aligned}$$

Thus we have

$$\mathbb{D}_2 \sim \text{pdf} \left[ e^{-(2^x e^{-\gamma})} (2^x e^{-\gamma}) \ln 2 \right] \quad (\text{C.55})$$



which is consistent to the result in our previous work (TODO:CITE).

Setting  $d = 1$  into Eqn. for contrast result in

$$\ln I_1 - \ln I_2 = \mathbb{C} \sim \Delta(2L) \quad (\text{C.56})$$

The characteristic function would then be

$$CF_{\mathbb{C}} = \frac{\Gamma(2L)B(L - it, L + it)}{\Gamma(L)^2} \quad (\text{C.57})$$

Thus PDF can be written as

$$\mathbb{C} \sim pdf \left[ \frac{\Gamma(2L)}{\Gamma(L)^2} \frac{e^{Lx}}{(1 + e^x)^{2L}} \right] \quad (\text{C.58})$$

due to

$$\begin{aligned} CF_{\mathbb{C}}(x) &= \frac{\Gamma(2L)}{\Gamma(L)^2} B(1/(1 + e^x), L - it, L + it) \\ &= \frac{\Gamma(2L)}{\Gamma(L)^2} \int_0^{1/(1+e^x)} z^{L-it-1} (1-z)^{L+it-1} dz \\ \frac{\partial}{\partial x} CF_{\mathbb{C}}(x) &= \frac{\partial CF_{\mathbb{C}}(x)}{\partial 1/(1+e^x)} \cdot \frac{\partial 1/(1+e^x)}{\partial x} \\ &= e^{ix} \frac{\Gamma(2L)}{\Gamma(L)^2} \frac{e^{Lx}}{(1 + e^x)^{2L}} \end{aligned}$$

Setting  $L = 1$  into Eqn. C.58 we have the PDF for contrast of single-look SAR:

$$\mathbb{C} \sim pdf \left[ \frac{e^x}{(1 + e^x)^2} \right] \quad (\text{C.59})$$

Converting to base-2 logarithm we have

$$\begin{aligned} \mathbb{C}/\ln 2 = \mathbb{C}_2 &\sim pdf \left[ \frac{e^x}{(1 + e^x)^2} \cdot dx/dt \right]_{x=t \cdot \ln 2} \\ &\sim pdf \left[ \ln 2 \frac{2^t}{(1 + 2^t)^2} \right] \end{aligned}$$

which is also consistent to the result in our previous work (TODO:CITE).

## REFERENCES

- [1] N. R. Goodman, "The distribution of the determinant of a complex wishart distributed matrix," *Annals of Mathematical Statistics*, vol. 34, no. 1, pp. 178–180, 1963.
- [2] T. H. Le, I. V. McLoughlin, and C. H. Vun, "Using mse to evaluate sar speckle filters," *IEEE Transactions on Geoscience and Remote Sensing*, vol. ??, no. ??, pp. ???–???, ??? 2013.
- [3] R. Raney and G. Wessels, "Spatial considerations in sar speckle consideration," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 26, no. 5, pp. 666–672, sep 1988.
- [4] J.-S. Lee, K. Hoppel, S. Mango, and A. Miller, "Intensity and phase statistics of multilook polarimetric and interferometric sar imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 32, no. 5, pp. 1017–1028, sep 1994.
- [5] S. Anfinsen, A. Doulgeris, and T. Eltoft, "Estimation of the equivalent number of looks in polarimetric synthetic aperture radar imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 47, no. 11, pp. 3795–3809, nov. 2009.
- [6] "Airsar implementation," [Accessed Feb 2013]. [Online]. Available: <http://airsar.jpl.nasa.gov/documents/genairsar/chapter3.pdf>
- [7] "Radarsat2 product description," [Accessed Feb 2013]. [Online]. Available: [http://gs.mdacorporation.com/products/sensor/radarsat2/RS2\\_Product\\_Description.pdf](http://gs.mdacorporation.com/products/sensor/radarsat2/RS2_Product_Description.pdf)
- [8] E. Rignot and J. van Zyl, "Change detection techniques for ers-1 sar data," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 31, no. 4, pp. 896–906, jul 1993.
- [9] F. Medeiros, N. Mascarenhas, and L. Costa, "Evaluation of speckle noise MAP filtering algorithms applied to SAR images," *International Journal of Remote Sensing*, vol. 24, no. 24, pp. 5197–5218, Dec 2003.
- [10] M. Dabboor, J. Yackel, M. Hossain, and A. Braun, "Comparing matrix distance measures for unsupervised POLSAR data classification of sea ice based on agglomerative clustering," *INTERNATIONAL JOURNAL OF REMOTE SENSING*, vol. 34, no. 4, pp. 1492–1505, 2013.
- [11] P. Kersten, J.-S. Lee, and T. Ainsworth, "Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and em clustering," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 43, no. 3, pp. 519–527, march 2005.
- [12] J.-S. Lee, M. Grunes, and G. de Grandi, "Polarimetric sar speckle filtering and its implication for classification," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 37, no. 5, pp. 2363–2373, sep 1999.
- [13] J. S. LEE, M. R. GRUNES, and R. KWOK, "Classification of multi-look polarimetric sar imagery based on complex wishart distribution," *International Journal of Remote Sensing*, vol. 15, no. 11, pp. 2299–2311, 1994.
- [14] S. N. Anfinsen, R. Jenssen, and T. Eltoft, "Spectral clustering of polarimetric sar data with wishart-derived distance measures," January 2007.
- [15] K. Y. Lee and T. Bretschneider, "Derivation of separability measures based on central complex gaussian and wishart distributions," in *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*, july 2011, pp. 3740–3743.
- [16] F. Cao, W. Hong, Y. Wu, and E. Pottier, "An unsupervised segmentation with an adaptive number of clusters using the span/h/α/a space and the complex wishart clustering for fully polarimetric sar data analysis," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, no. 11, pp. 3454–3467, nov. 2007.
- [17] K. Conradsen, A. Nielsen, J. Schou, and H. Skriver, "A test statistic in the complex wishart distribution and its application to change detection in polarimetric sar data," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 41, no. 1, pp. 4–19, jan 2003.
- [18] J. Schou, H. Skriver, A. Nielsen, and K. Conradsen, "Cfar edge detector for polarimetric sar images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 41, no. 1, pp. 20–32, jan 2003.
- [19] T. H. Le, I. V. McLoughlin, K. Y. Lee, and T. Brestchneider, "SLC SAR speckle filtering using homoskedastic features of logarithmic transformation," in *Proceedings of the 31th Asian Conference on Remote Sensing (ACRS)*, Hanoi, Vietnam, Nov 2010.
- [20] T. H. Le and I. V. McLoughlin, "SAR Fuzzy-MLE speckle filter using the distance consistency property in homoskedastic log-transformed domain," in *Proceedings of the 32th Asian Conference on Remote Sensing (ACRS)*, Taipei, Taiwan, Nov 2011.