Chapter 4

Consistent Measures of Distance for Multivariate POLSAR Data

In this chapter, some statiscally consistent measures of distance for multi-variate POLSAR data is presented. The first section introduces the basics of POLSAR statiscal analysis and the log-transformation proposed in this thesis. The second section presents conclusive evidence that POLSAR data is multiplicative and heteroskedastic in the original domain. Then log-transformation is shown converting these into an additive and homoskedastic model. Consequently a few consistent measures of distance, namely the dispersion and the contrast measurements, is presented in the third section. Section (TODO:REF) illustrates how a few state-of-the-art results in POLSAR processing might be explained using the present approach. From another perspective, section (TODO:REF) discusses how the time-tested results for univariate SAR can also be considered as a special case for the multi-variate POLSAR results. Finally the last section strengthens our model showing the matching between simulation and real-life captured data not only for a few dimensionally-collapsed single-dimensional measures but also for all elements of the multi-dimensional POLSAR covariance matrix.

4.1 POLSAR statistical analysis

In this section, 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:

$$s_{part} = \begin{bmatrix} S_h \\ S_v \end{bmatrix} \tag{4.1}$$

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

$$s_{full} = \begin{bmatrix} S_{hh} \\ \sqrt{2}S_{hv} \\ S_{vv} \end{bmatrix} \tag{4.2}$$

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 Σ, then the probably density function (PDF) of s can be written as:

$$pdf(s;\Sigma) = \frac{1}{\pi^d|\Sigma|}e^{-s^{*T}\Sigma^{-1}s} \tag{4.3}$$

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

$$C_v = \langle ss^{*T} \rangle = \frac{1}{L}\sum_{i=1}^{L} s_i s_i^{*T} \tag{4.4}$$

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:

$$pdf(Z;d,\Sigma,L) = \frac{|Z|^{L-d}}{|\Sigma^L|\Gamma_d(L)}e^{-tr(\Sigma^{-1}Z)} \tag{4.5}$$

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 deterimant of the covarinace matrix. In [96], 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

$$\chi_L^d = (2L)^d \frac{|C_v|}{|\Sigma_v|} \sim \prod_{i=0}^{d-1} \chi^2(2L-2i) \tag{4.6}$$

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

$$\Lambda_L^d = ln\left[(2L)^d \frac{|C_v|}{|\Sigma_v|}\right] \sim \sum_{i=0}^{d-1} \Lambda^\chi(2L-2i) \tag{4.7}$$

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

4.2 Hetoroskedastic POLSAR data 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 variance is fixed and be independent of the underlying mean, is demonstrated. To keep the section flowing, the mathematical derivation is only presented here in major sketches. For more detailed derivation, Appendix **??** is to be reviewed.

From Eqn. 4.6 and Eqn. 4.7 we have:

$$|C_v| \sim |\Sigma_v| \cdot \frac{1}{(2L)^d} \cdot \prod_{i=0}^{d-1} \chi^2(2L-2i) \tag{4.8}$$

$$\ln|C_v| \sim \ln|\Sigma_v| - d \cdot \ln(2L) + \sum_{i=0}^{d-1} \Lambda(2L-2i) \tag{4.9}$$

In a given homogeneous POLSAR area, the parameters Σ_v, d and L can be considered as constant. Thus Eqn. 4.8 gives the theoretical explaination that in the original POLSAR domain, a multiplicative speckle noise pattern is present. At the same time, Eqn. 4.9 shows that the logarithmic transformation converts this into a more familiar additive noise.

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

$$pdf(x;k) = \frac{x^{(k/2)-1}e^{-x/2}}{2^{k/2}\Gamma\left(\frac{k}{2}\right)} \tag{4.10}$$

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

$$pdf(x;2L=k) = \frac{e^{Lx-e^x/2}}{2^L\Gamma(L)} \tag{4.11}$$

Given the probability distribution function, the characteristic function of both the chi-squared and log-chi-squared random variables can be written as:

$$CF_\chi(t) = (1-2it)^L \tag{4.12}$$

$$CF_\Lambda(t) = 2^{it}\frac{\Gamma(L+it)}{\Gamma(L)} \tag{4.13}$$

Subsequently their means and variances can be computed from the given characteristic functions as:

$$avg\left[\chi(2L)\right] = 2L \tag{4.14}$$

$$var\left[\chi(2L)\right] = 4L \tag{4.15}$$

$$avg\left[\Lambda(2L)\right] = \psi^0(L) + \ln 2 \tag{4.16}$$

$$var\left[\Lambda(2L)\right] = \psi^1(L) \tag{4.17}$$

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. 4.8 and 4.9, we have:

$$avg\left[|C_v|\right] = \frac{|\Sigma_v|}{L^d}\prod_{i=0}^{d-1}(L-i) \tag{4.18}$$

$$var\left[|C_v|\right] = \frac{|\Sigma_v|^2\left[\prod_{i=0}^{d-1}(L-i)(L-i+1)-\prod_{i=0}^{d-1}(L-i)^2\right]}{L^{2d}} \tag{4.19}$$

$$avg\left[\ln|C_v|\right] = \ln|\Sigma_v|-d\cdot\ln L+\sum_{i=0}^{d-1}\psi^0(L-i) \tag{4.20}$$

$$var\left[\ln|C_v|\right] = \sum_{i=0}^{d-1}\psi^1(L-i) \tag{4.21}$$

Over a real and captured image, while the paramaters d and L do not change for the whole image, the underlying Σ_v is expected to differ from one region to the next. Thus over an heterogeneous scene, the stochastic process for $|C_v|$ and $\ln|C_v|$ varies depending on the underlying signal Σ_v. In such context, Eqn. 4.19 indicates that the variance of $|C_v|$ also differs depending on the underlying signal Σ_v, which indicates its heteroskedastic property. At the same time, in the log-transformed domain, Eqn. 4.21 shows that the variance of $\ln|C_v|$ is invariant and independent of Σ_v manifesting its homoskedastic nature.

4.3 Consistent Measures of Distance for POLSAR data

Similar to the way dispersion and contrast is defined in our previous work [97], this section introduces the consistent sense of distance in a few different perspective. First assumming that the true value of the underlying signal Σ_v is known *a priori*, the following random variables, namely ratio ($\mathbb{R}$) and log-distance ($\mathbb{L}$), are observable according to their definitions:

$$\mathbb{R} = \frac{|C_v|}{|\Sigma_v|} \tag{4.22}$$

$$\mathbb{L} = \ln|C_v| - \ln|\Sigma_v| \tag{4.23}$$

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

$$\mathbb{D} = \ln|C_v| - avg(\ln|C_v|) \tag{4.24}$$

$$\mathbb{C} = \ln(|C_{v1}|) - \ln(|C_{v2}|) \tag{4.25}$$

Taking the results from Eqns. 4.8, 4.9 and 4.20 we have

$$\mathbb{R} \sim \frac{1}{(2L)^d} \cdot \prod_{i=0}^{d-1} \chi^2(2L-2i) \tag{4.26}$$

$$\mathbb{L} \sim \sum_{i=0}^{d-1} \Lambda(2L-2i) - d \cdot \ln(2L) \tag{4.27}$$

$$\mathbb{D} \sim \sum_{i=0}^{d-1} \Lambda(2L-2i) - d \cdot \ln 2 + k \tag{4.28}$$

$$\mathbb{C} \sim \sum_{i=0}^{d-1} \Delta(2L-2i) \tag{4.29}$$

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. 4.13, Appendix **??** derives the characteristic functions for the summative

random variables as:

$$CF_{\Lambda_L^d}(t) = \frac{2^{idt}}{\Gamma(L)^d} \prod_{j=0}^{d-1} \Gamma(L-j+it) \tag{4.30}$$

$$CF_{\mathbb{L}}(t) = \frac{1}{L^{idt}\Gamma(L)^d} \prod_{j=0}^{d-1} \Gamma(L-j+it) \tag{4.31}$$

$$CF_{\mathbb{D}}(t) = \frac{e^{ikt}}{\Gamma(L)^d} \prod_{j=0}^{d-1} \Gamma(L-j+it) \tag{4.32}$$

$$CF_{\Delta(2L)} = \frac{\Gamma(2L)B(L-it, L+it)}{\Gamma(L)^2} \tag{4.33}$$

$$CF_{\mathbb{C}}(t) = \prod_{j=0}^{d-1} \frac{\Gamma(2L-2j)B(L-j-it, L-j+it)}{\Gamma(L-j)^2} \tag{4.34}$$

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 Σ_v. This serves as conclusive evidence that these random variables follows consistent and fixed distributions, regardless of the underlying signal Σ_v.

4.4 SAR as the Special Case of Polarimetric SAR

The prevous section has validated the use of our models for 3-dimensional ($d = 3$) full polarimetric and 2-dimensional ($d = 2$) partial polarimetric SAR cases. In this section, the focus is on the limit case where the dimensional number is reduced to 1, i.e. $d = 1$. Physically this means the multi-dimensional POLSAR dataset is collapsed into the one-dimensional and classical SAR data. Mathmatically, 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:

$$\mathbb{R} = \frac{|C_v|}{|\Sigma_v|} \sim \frac{1}{(2L)^d} \prod_{i=0}^{d-1} \chi^2(2L-2i) \tag{4.35}$$

$$\mathbb{L} = \ln|C_v| - \ln|\Sigma_v| \sim \sum_{i=0}^{d-1} \Lambda(2L-2i) - d \cdot \ln 2L \tag{4.36}$$

$$\mathbb{D} = \ln|C_v| - avg(\ln|C_v|) \sim \sum_{i=0}^{d-1} \Lambda(2L-2i) - d\ln 2 + k \tag{4.37}$$

$$\mathbb{C} = \ln|C_{1v}| - \ln|C_{2v}| \sim \sum_{i=0}^{d-1} \Delta(2L-2i) \tag{4.38}$$

$$\mathbb{A} = avg(\mathbb{L}) = \sum_{i=0}^{d-1} \psi^0(L-i) - d \cdot \ln L \tag{4.39}$$

$$\mathbb{V} = var(\mathbb{L}) = \sum_{i=0}^{d-1} \psi^1(L-i) \tag{4.40}$$

$$\mathbb{E} = mse(\mathbb{L}) = \left[\sum_{i=0}^{d-1} \psi^0(L-i) - d \cdot \ln L\right]^2 + \sum_{i=0}^{d-1} \psi^1(L-i) \tag{4.41}$$

Upon setting $d = 1$ into the above models, Section 4.6.4 shows that the reduced results are consistent with not only the following results from the our previous works on single-look SAR [97], i.e. $d = L = 1$,

$$\begin{aligned}
I &\sim \bar{I} \cdot pdf\left[e^{-R}\right] \\
\log_2 I &\sim \log_2 \bar{I} + pdf\left[2^x e^{-2^x} \ln 2\right] \\
\mathbb{R} &= \frac{I}{\bar{I}} \sim pdf\left[e^{-x}\right] \\
\mathbb{L} &= \log_2 I - \log_2 \bar{I} \sim pdf\left[2^x e^{-2^x} \ln 2\right] \\
\mathbb{D} &= \log_2 I - avg(\log_2 I) \sim pdf\left[e^{-(2^x e^{-\gamma})} 2^x e^{-\gamma} \ln 2\right] \\
\mathbb{C} &= \log_2 I_1 - \log_2 I_2 \sim pdf\left[\frac{2^x}{(1+2^x)^2} \ln 2\right] \\
\mathbb{A} &= avg(\mathbb{L}) = -\gamma/\ln 2 \\
\mathbb{V} &= var(\mathbb{L}) = \frac{\pi^2}{6}\frac{1}{\ln^2 2} \\
\mathbb{E} &= mse(\mathbb{L}) = \frac{1}{\ln^2 2}(\gamma^2 + \pi^2/6) = 4.1161
\end{aligned}$$

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

$$I \sim pdf\left[\frac{L^L x^{L-1} e^{-Lx/\bar{I}}}{\Gamma(L)\bar{I}^L}\right] \tag{4.42}$$

$$N = \ln I \sim pdf\left[\frac{L^L}{\Gamma(L)} e^{L(x-\bar{N})-Le^{x-\bar{N}}}\right] \tag{4.43}$$

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:

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

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

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

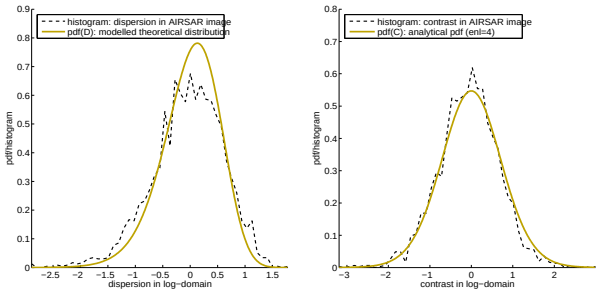
$$\mathbb{C} = \ln I_1 - \ln I_2 \sim pdf\left[\frac{e^x}{(1+e^x)^2}\right]$$

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

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

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

This newly derived models for multi-look SAR data can also be validated against real-life data. Fig. 4.1 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.



(a) verify multi-look SAR log-distance

(b) verify multi-look SAR contrast

Figure 4.1: Multi-Look SAR dispersion and contrast: modelled response matches very well with real-life captured data.

4.5 Validating the models against real-life data

4.5.1 Explaining practical data with the given Number-of-Looks

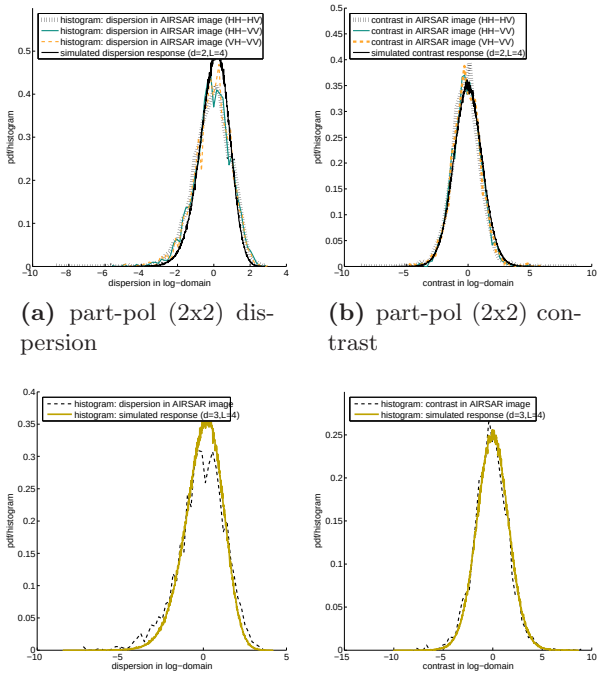
This section describes an experiment to validate the models above against real-life captured data. The validation procedure is quite straightforward. Given that the stochastic models have been derived in the previous sections, they can be graphically visualized by the histogram plots of the simulated 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 matrixs determinant in the log-transformed domain, i.e. $avg(ln|C_v|)$, is measured for dispersion. Subsequently the observable samples of dispersion and contrast are computed according to Eqns. 4.24 and 4.25. Then their histograms are traced out.

At the same time, theoretical simulations according to Eqns. 4.28 and 4.29 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. 4.2. Apparently, a visual match is observable which verify the applicability of the theoretical models for the dispersion and contrast measures of distance.



(a) part-pol (2x2) dispersion

(b) part-pol (2x2) contrast

(c) full-pol (3x3) dispersion

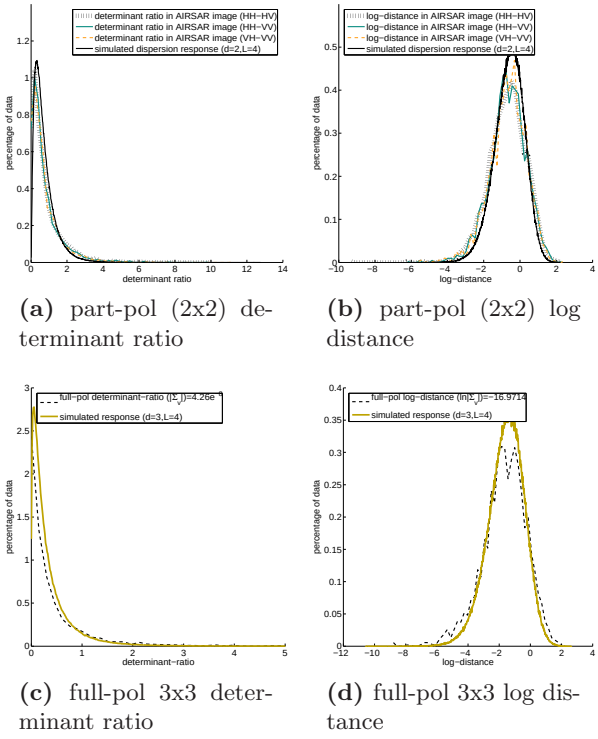
(d) full-pol (3x3) contrast

Figure 4.2: Validating the dispersion and contrast models against both partial and full polarimetric AIRSAR Flevoland data.

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. $\Sigma_v = avg(C_v)$. Another approach is to estimate the true signal from the average of the log-determinant of the polsar covariance matrix (i.e. $avg[ln|C_v|]$) using Eqn. 4.20. 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. 4.3 validate the models of determinant-ratio and log-distance against real-life data. In this experiment, the theoretical models is simulated from Eqns 4.26 and 4.27, while the observable samples are computed using Eqns 4.22 and 4.23 with the true signal estimated from the log-determinant average, i.e. $avg(\ln|C_v|)$. Again a reasonable match is observed which validates the models for log-distance and determinant ratio.



(a) part-pol (2x2) determinant ratio

(b) part-pol (2x2) log distance

(c) full-pol 3x3 determinant ratio

(d) full-pol 3x3 log distance

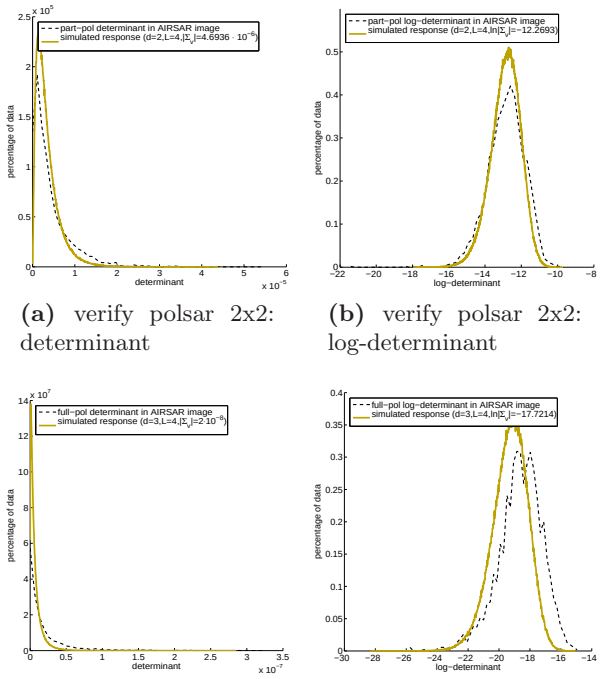
Figure 4.3: Validating determinant-ratio and log-distance models with $|\Sigma_v|$ is computed using $avg(\ln|C_v|)$

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. 4.8 and 4.9. 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 4.4, 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.



(a) verify polsar 2x2: determinant

(b) verify polsar 2x2: log-determinant

(c) verify polsar 3x3: determinant

(d) verify polsar 3x3: log-determinant

Figure 4.4: Validating determinant and log-determinant models with $\Sigma_v = avg(C_v)$

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 4.5.2, 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.

4.5.2 Comparing Theoretical Assumptions and Practical Implementations

Even though the initial assumptions for the proposed theory is intentionally kept to minimal, like all other similar models, 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 $\Sigma_v = avg(C_v) = \begin{vmatrix} 0.0084 & 1 \cdot 10^{-6} + 4 \cdot 10^{-4}i & 0.0071 - 0.0017i \\ 1 \cdot 10^{-6} - 4 \cdot 10^{-4}i & 0.0017 & -3 \cdot 10^{-4} - 2 \cdot 10^{-4}i \\ 0.0071 + 0.0017i & -3 \cdot 10^{-4} + 2 \cdot 10^{-4}i & 0.0122 \end{vmatrix}$. 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. 4.2 to 4.4. This suggests that the proposed model is also applicable on

correlated POLSAR components, eventhough a full explaination 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 [98] for SAR, while the later is experimental observed for POLSAR data in [99] and [100]. 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 [101]. While for RadarSat2, the pixel resolution and range - azimuth resolutions for SLC fine-quad mode is advertised as $(4.7 \cdot 5.1)m^2/(5.2 \cdot 7.7)m^2$ [102]

The proposed model can handle this imperfection that is available in pratical 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 4.5 for the AIRSAR Flevoland dataset can be improved, by using ENL estimation.

4.5.3 Estimating the Effect Number-of-Looks (ENL)

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 4.2. In fact Eqn. 4.20 indicates that there is a relationship among $|avg(C_v)|, avg(\ln|C_v|), d, L$. 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 datasat, since all values of $|avg(C_v)|, avg(\ln|C_v|), d$ 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 [100], where an equation of exactly the same form as Eqn. 4.20 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 [97] towards POLSAR data. Specifically, Eqn. 4.21 can be rewritten as:

$$var\left[ln|C_v|\right] = f(L) = \sum_{i=0}^{d-1} \psi^1(L-i) \tag{4.44}$$

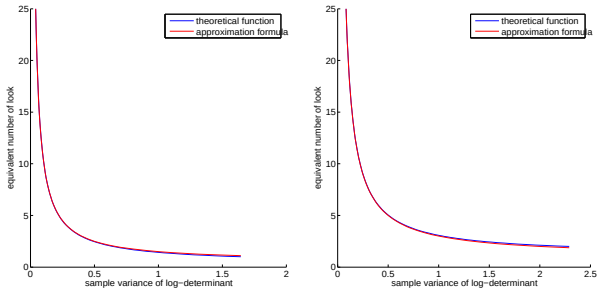
where $\psi^1()$ denotes tri-gamma function.

Thus theoretically, given some measurable value for $var\left[ln|C_v|\right]$, 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 $var\,[ln|C_v|]$, 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-envolope calculation given as:

$$\hat{L} = d\left(\frac{1}{var(\ln|C_v|)} + 0.5\right) \tag{4.45}$$

Fig. 4.5 not only shows the shapes of the function defined in Eqn. 4.44 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. 4.45).



(a) ENL and variance log-intensity relations for SAR data

(b) ENL and var(log-det) relations for partial POLSAR data

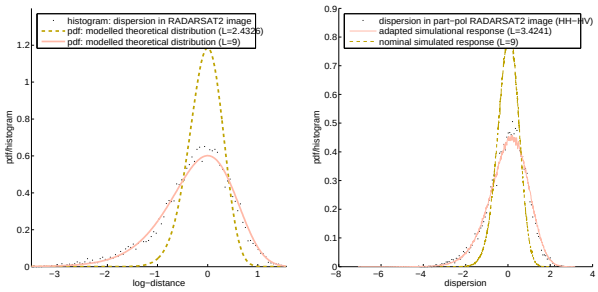
Figure 4.5: The relationships between ENL and the sample variance of log-determinant and log-intensity

4.5.4 Using the estimated ENL to better explain pratical 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.

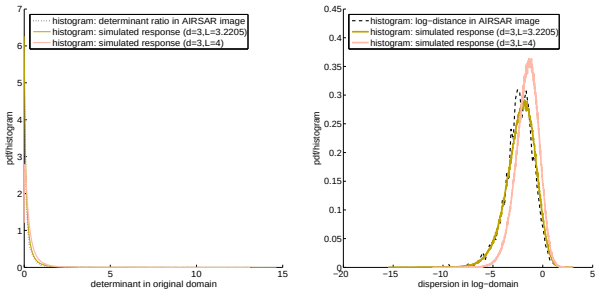


(a) Handling over-sampling practice in Radarsat2 one-dimensional SAR data (HH)

(b) Handling over-sampling practice in Radarsat2 partial POLSAR data (HH-HV)

Figure 4.6: 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 offer better explaination for the data.

Fig. 4.7 shows that the over-sampling issue is also present in AIRSAR Flevoland dataset, eventhough 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.



(a) Handling over-sampling practice in AIRSAR full-pol dataset (contrast better?)

(b) Handling over-sampling practice in AIRSAR full-pol dataset

Figure 4.7: AIRSAR Flevoland also exhibits the over-sampling phenomena, though at a much smaller extend than the RADARSAT 2 data.

4.6 Details of Mathematical Derivations

4.6.1 The Log-Chi-Square Distribution

This appendix provides the mathematical proofs for the log-transformed version of chi-squared random variables.

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

$$pdf(\chi;\, k) = \frac{\chi^{(k/2)-1} e^{-\chi/2}}{2^{k/2}\Gamma\left(\frac{k}{2}\right)} \tag{4.46}$$

Setting L=k/2 into Eqn. 4.46

$$pdf(\chi) = \frac{\chi^{L-1} e^{-\chi/2}}{2^{L}\Gamma(L)} \tag{4.47}$$

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

$$\int_a^b f(y)\, dy = \int_c^d f[\phi(x)] \frac{d\phi}{dx} dx \tag{4.48}$$

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:

$$pdf(\Lambda; L) = \frac{e^{L\Lambda - e^{\Lambda}/2}}{2^{L}\Gamma(L)} \tag{4.49}$$

From the PDF given in Eqn. 4.49, characteristics function can be computed. By definition, the characteristic function (CF) $\varphi_X(t)$ for a random variable X is computed as:

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

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:

$$\varphi_\Lambda(t) = \int_0^\infty e^{itx} \frac{e^{Lx-e^x/2}}{2^L \Gamma(L)}\, dx \tag{4.50}$$

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 $dx = e^z/2dz$, we have $\Gamma(L+it) = \int_0^\infty e^{itz} \frac{e^{Lz-e^z/2}}{2^{L+it}} dz$

That is:

$$\varphi_\Lambda(t) = 2^{it} \frac{\Gamma(L+it)}{\Gamma(L)} \tag{4.51}$$

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

$$\frac{\partial \varphi_\Lambda(t)}{\partial t} = \frac{i 2^{it} \Gamma(L+it)}{\Gamma(L)} \left[\ln 2 + \psi^0(L+it)\right] \tag{4.52}$$

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:

$$\frac{\partial^2 \varphi_\Lambda(t)}{\partial t^2} = \frac{i^2 2^{it} \Gamma(L+it)}{\Gamma(L)} \left(\left[\ln 2 + \psi^0(L+it)\right]^2 + \psi^1(L+it) \right) \tag{4.53}$$

due to:

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

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:

$$\mathrm{E}\left(\Lambda^n\right)=i^{-n}\,\varphi_\Lambda^{(n)}(0)=i^{-n}\,\left[\frac{d^n}{dt^n}\varphi_\Lambda(t)\right]_{t=0} \tag{4.54}$$

Thus

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

That is

$$avg(\Lambda)=\psi^0(L)+ln(2) \tag{4.55}$$

Similarly,

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

That is:

$$E(\Lambda^2)=\left[\psi^0(L)+\ln(2)\right]^2+\psi^1(L) \tag{4.56}$$

Thus

$$var(\Lambda)=E(\Lambda^2)-E^2(\Lambda)=\psi^1(L) \tag{4.57}$$

4.6.2 Summary Statistics of Some POLSAR Observables

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

$$\chi_L^d \sim \prod_{i=0}^{d-1} \chi(2L-2i) \tag{4.58}$$

$$\Lambda_L^d \sim \sum_{i=0}^{d-1} \Lambda(2L-2i) \tag{4.59}$$

given the averages and variances of individual components.

$$avg\left[\chi(2L)\right] = 2L \tag{4.60}$$

$$var\left[\chi(2L)\right] = 4L \tag{4.61}$$

$$avg\left[\Lambda(2L)\right] = \psi^0(L) + \ln 2 \tag{4.62}$$

$$var\left[\Lambda(2L)\right] = \psi^1(L) \tag{4.63}$$

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:

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

Thus they can be computed as:

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

4.6.3 The characteristic function of some dissimilarity measures for polsar

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, can 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:

$$CF_{\chi_L^d}(t) = \frac{2^{idt}}{\Gamma(L)^d}\prod_{j=0}^{d-1}\Gamma(L-j+it) \tag{4.64}$$

$$CF_{\mathbb{L}} = \frac{1}{L^{idt}\Gamma(L)^d}\prod_{j=0}^{d-1}\Gamma(L-j+it) \tag{4.65}$$

$$CF_{\mathbb{D}} = \frac{1}{\Gamma(L)^d}\prod_{j=0}^{d-1}e^{idt\psi^0(L-j)}\Gamma(L-j+it) \tag{4.66}$$

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:

$$CF_{\mathbb{C}} = \prod_{j=0}^{d-1}\frac{\Gamma(2L-2j)B(L-j-it,L-j+it)}{\Gamma^2(L-j)} \tag{4.67}$$

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

4.6.4 SAR intensity as the special case of the POLSAR cov-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:

$$
\begin{aligned}
I &\sim \bar{I} \cdot pdf\left[e^{-R}\right] && (4.68)\\
\log_2 I &\sim \log_2 \bar{I} + pdf\left[2^{D-2^D}\right] && (4.69)\\
\frac{I}{\bar{I}} = \mathbb{R} &\sim pdf\left[e^{-R}\right] && (4.70)\\
\log_2 I - \log_2 \bar{I} = \mathbb{D} &\sim pdf\left[2^D e^{-2^D} \ln 2\right] && (4.71)\\
\log_2 I_1 - \log_2 I_2 = \mathbb{C} &\sim pdf\left[\frac{2^c}{(1+2^c)^2} \ln 2\right] && (4.72)\\
avg(\mathbb{D}) &= -\gamma / \ln 2 && (4.73)\\
var(\mathbb{D}) &= \frac{\pi^2}{6}\frac{1}{\ln^2 2} && (4.74)\\
mse(\mathbb{D}) &= \frac{1}{\ln^2 2}(\gamma^2 + \pi^2/6) = 4.1161 && (4.75)
\end{aligned}
$$

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

$$
\begin{aligned}
I &\sim pdf\left[\frac{L^L I^{L-1} e^{-LI/\bar{I}}}{\Gamma(L)\bar{I}^L}\right] && (4.76)\\
N = \ln I &\sim pdf\left[\frac{L^L}{\Gamma(L)} e^{L(N-\bar{N}) - Le^{N-\bar{N}}}\right] && (4.77)
\end{aligned}
$$

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

$$
\begin{aligned}
|C_v| &\sim \frac{|\Sigma_v|}{(2L)^d} \prod_{i=0}^{d-1} \chi^2(2L-2i) && (4.78)\\
\ln|C_v| &\sim \ln|\Sigma_v| + \sum_{i=0}^{d-1} \Lambda(2L-2i) - d \cdot \ln 2L && (4.79)
\end{aligned}
$$

$$\frac{|C_v|}{|\Sigma_v|} = \mathbb{R} \sim \frac{1}{(2L)^d} \prod_{i=0}^{d-1} \chi^2(2L-2i) \tag{4.80}$$

$$\ln|C_v| - \ln|\Sigma_v| = \mathbb{D} \sim \sum_{i=0}^{d-1} \Lambda(2L-2i) - d \cdot \ln 2L \tag{4.81}$$

$$\ln|C_{1v}| - \ln|C_{2v}| = \mathbb{C} \sim \sum_{i=0}^{d-1} \Delta(2L-2i) \tag{4.82}$$

$$avg(\mathbb{D}) = \sum_{i=0}^{d-1} \psi^0(L-i) - d \cdot \ln L \tag{4.83}$$

$$var(\mathbb{D}) = \sum_{i=0}^{d-1} \psi^1(L-i) \tag{4.84}$$

$$mse(\mathbb{D}) = \left[\sum_{i=0}^{d-1} \psi^0(L-i) - d \cdot \ln L\right]^2 + \sum_{i=0}^{d-1} \psi^1(L-i) \tag{4.85}$$

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:

$$\frac{I}{\bar{I}} = \mathbb{R} \sim \frac{1}{2L}\chi^2(2L) \tag{4.86}$$

$$\ln I - \ln \bar{I} = \mathbb{D} \sim \Lambda(2L) - \ln 2L \tag{4.87}$$

$$\ln I_1 - \ln I_2 = \mathbb{C} \sim \Delta(2L) \tag{4.88}$$

$$avg(\mathbb{D}) = \psi^0(L) - \ln L \tag{4.89}$$

$$var(\mathbb{D}) = \psi^1(L) \tag{4.90}$$

$$mse(\mathbb{D}) = \left[\psi^0(L) - \ln L\right]^2 + \psi^1(L) \tag{4.91}$$

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).

4.6.4.1 Original Domain: SAR Intensity and its ratio

Setting $d = 1$, $|C_v| = I$ and $|\Sigma_v| = \bar{I}$ into Eqns. 4.78 and 4.80 we have:

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

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

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

Thus we have the following results for multi-look SAR

$$I \sim pdf\left[\frac{L^L x^{L-1}e^{-Lx/\bar{x}}}{\bar{I}^L\Gamma(L)}\right] \tag{4.92}$$

$$\frac{I}{\bar{I}} = \mathbb{R} \sim pdf\left[\frac{L^L x^{L-1}e^{-Lx}}{\Gamma(L)}\right] \tag{4.93}$$

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

$$I \sim pdf\left[\frac{e^{x/\bar{I}}}{\bar{I}}\right] \tag{4.94}$$

$$\frac{I}{\bar{I}} = \mathbb{R} \sim pdf\left[e^{-x}\right] \tag{4.95}$$

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

4.6.4.2 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. 4.79 and 4.81 we have

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

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

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

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

Thus the first approach arrives at

$$\ln I = \mathbb{N} \sim pdf\left[\frac{L^L e^{L(x-\bar{N})-Le^{x-\bar{N}}}}{\Gamma(L)}\right] \tag{4.96}$$

$$\ln I - \ln \bar{I} = \mathbb{L} \sim pdf\left[\frac{L^L e^{Lt-Le^t}}{\Gamma(L)}\right] \tag{4.97}$$

In the second approach, log-transformation is applied on previous result for multi-look SAR intensity and its ratio in the original domain (Eqns. 4.93 and 4.92). 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. 4.85, 4.83 and 4.84 we have:

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

This completes the first phase of the derivation process. The second phase of simpli-

fication 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

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

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

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

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

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

4.6.4.3 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:

$$\ln I - avg(\ln I) = \mathbb{D} \sim pdf\left[\frac{e^{L[x+\psi^0(L)]-Le^{x+\psi^0(L)-\ln L}}}{\Gamma(L)}\right] \tag{4.98}$$

due to $d = 1$ and

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

.

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

$$\mathbb{D} \sim pdf\left[e^{x-\gamma-e^{x-\gamma}}\right] \tag{4.99}$$

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

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

Thus we have

$$\mathbb{D}_2 \sim pdf\left[e^{-(2^x e^{-\gamma})}(2^x e^{-\gamma})\ln 2\right] \tag{4.100}$$

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

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

$$\ln I_1 - \ln I_2 = \mathbb{C} \sim \Delta(2L) \tag{4.101}$$

The characteristic function would then be

$$CF_{\mathbb{C}} = \frac{\Gamma(2L)B(L - it, L + it)}{\Gamma(L)^2} \tag{4.102}$$

Thus PDF can be written as

$$\mathbb{C} \sim pdf\left[\frac{\Gamma(2L)}{\Gamma(L)^2}\frac{e^{Lx}}{(1 + e^x)^{2L}}\right] \tag{4.103}$$

due to

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

Setting $L = 1$ into Eqn. 4.103 we have the PDF for contrast of single-look SAR:

$$\mathbb{C} \sim pdf \left[\frac{e^x}{(1+e^x)^2} \right] \tag{4.104}$$

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).