Generation of Sample Complex Wishart Distributed Matrices and Change Detection in Polarimetric SAR Data

Nielsen, Allan Aasbjerg; Skriver, Henning; Conradsen, Knut

Published in:

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Generation of Sample Complex Wishart Distributed Matrices and Change Detection in Polarimetric SAR Data

Allan A. Nielsen\(^a\), Henning Skriver\(^b\) and Knut Conradsen\(^a\)

Technical University of Denmark
\(^a\)DTU Compute – Applied Mathematics and Computer Science
\(^b\)DTU Space – National Space Institute
DK-2800 Kgs. Lyngby, Denmark

ABSTRACT
The complex Wishart distribution is used to describe synthetic aperture radar (SAR) data in the so-called covariance matrix representation. We give a test statistic for detection of differences between two instances in this distribution and an associated probability measure. Generated complex Wishart distributed covariance matrices are used to show that this test statistic and the probability measure in situations with no differences follow the expected distributions.

Keywords
Polarimetric SAR data, dual polarization SAR data, test statistic, change detection, target detection.

1. INTRODUCTION
In this paper we study the behaviour of the test statistic and the associated probability measure for detection of differences such as change over time in polarimetric synthetic aperture radar (SAR) data in the covariance (or equivalently the coherency) matrix formulation [1] described in [2, 3]. We study the test statistic and the probability as functions of the equivalent number of looks (ENL). To emulate polarimetric data (from for example Radarsat-2, ALOS or TerraSAR-X) and dual polarimetry data (from for example Sentinel-1 or COSMO-SkyMed), we generate many outcomes of both 3×3 and 2×2 Hermitian, positive definite matrices, as well as 2×2 diagonal-only matrices with the same parameters (only, we let the ENL vary). In a change detection setting, this emulates the situation of no change between two time points. We therefore know the distributions of the test statistic and the probability, see [2, 4] for details.

2. SYNTHETIC APERTURE RADAR
A fully polarimetric SAR measures the 2×2 complex so-called scattering matrix at each resolution cell on the ground

\[
S = \begin{bmatrix}
S_{hh} & S_{hv} \\
S_{vh} & S_{vv}
\end{bmatrix}
\]

The scattering matrix relates the incident and the scattered electric fields [1]. If \(S_{rt}\) denotes the complex scattering amplitude for receive and transmit polarizations \((r, t \in \{h, v\})\) for horizontal and vertical polarizations), then reciprocity, which normally applies to natural targets, gives \(S_{bh} = S_{hb}\) (in the backscattering direction using the backscattering alignment convention) [1]. Assuming reciprocity, the scattering matrix is represented by the three-component complex target vector \(s = [S_{hh} \ S_{hv} \ S_{vh}]^T\), where \(^T\) means transpose. The inherent speckle in the SAR data can be reduced by spatial averaging at the expense of spatial resolution. In this so-called multilook case a more appropriate representation of the backscattered signal is the covariance matrix in which the average properties of a group of resolution cells can be expressed in a single matrix formed by the outer products of the averaged target vectors. The average covariance matrix is defined as [1]

\[
\langle C \rangle = \begin{pmatrix}
\langle S_{hh}S_{hh}^* \rangle & \langle S_{hh}S_{hv}^* \rangle & \langle S_{hh}S_{vh}^* \rangle \\
\langle S_{hv}S_{hh}^* \rangle & \langle S_{hv}S_{hv}^* \rangle & \langle S_{hv}S_{vh}^* \rangle \\
\langle S_{vh}S_{hh}^* \rangle & \langle S_{vh}S_{hv}^* \rangle & \langle S_{vh}S_{vh}^* \rangle
\end{pmatrix}
\]

where \(\langle \cdot \rangle\) denotes ensemble averaging and \(^*\) denotes complex conjugation.

Spaceborne instruments often transmit only one polarization, say vertical, and receive both polarizations giving rise to dual polarization data. In the covariance matrix representation each pixel at each time point is a matrix

\[
\langle C \rangle_{dual} = \begin{pmatrix}
\langle S_{vh}S_{vh}^* \rangle & \langle S_{vh}S_{hv}^* \rangle \\
\langle S_{hv}S_{vh}^* \rangle & \langle S_{hv}S_{hv}^* \rangle
\end{pmatrix}
\]

Sometimes we have the diagonal elements only.

2.1. Test for Equality of Two Complex Covariance Matrices
To test whether two complex variance-covariance matrices \(\Sigma_x\) and \(\Sigma_y\) (both \(p \times p\)) are equal, i.e., to test the so-called null hypothesis

\[H_0 : \Sigma_x = \Sigma_y\]

against the alternative, \(H_1 : \Sigma_x \neq \Sigma_y\), the following test statistic applies (for the real case see [5])

\[
Q = \frac{(m+n)^{m+n}}{m^m n^n} \frac{|m\langle C \rangle_x|^n |n\langle C \rangle_y|^n}{|m\langle C \rangle_x + n\langle C \rangle_y|^{m+n}}.
\]

Here \(m\langle C \rangle_x = m\Sigma_x\) and \(n\langle C \rangle_y = n\Sigma_y\) follow the complex Wishart distribution, i.e., \(m\langle C \rangle_x \sim W_C(p, m, \Sigma_x)\), and \(n\langle C \rangle_y \sim W_C(p, n, \Sigma_y)\). \(Q \in [0, 1]\) with \(Q = 1\) for equality.
For the logarithm of the test statistic we get
\[ \ln Q = p \left\{ (m + n) \ln(m + n) - m \ln m - n \ln n \right\} \\
+ m \ln |m(C)_x| + n \ln |n(C)_y| \\
- (m + n) \ln |m(C)_x + n(C)_y| . \]

If
\[ \rho = 1 - \frac{2p^2 - 1}{6p} \left( \frac{1 + m}{n} + \frac{1 - m}{m + n} \right) \]
and
\[ \omega_2 = \frac{p^2}{4} \left( 1 - \frac{1}{p} \right)^2 + \frac{p^2(p^2 - 1)}{24p^2} \]
\[ \cdot \left( \frac{1}{m^2} + \frac{1}{n^2} - \frac{1}{(m + n)^2} \right) , \]

and if the observed value of \(-2 \rho \ln Q\) is \(z = -2 \rho \ln q_{obs}\), then the probability of finding a smaller value of \(-2 \rho \ln Q\) is
\[ P\{-2 \rho \ln Q \leq z\} \approx P\{\chi^2(f) \leq z\} \\
+ \omega_2[P\{\chi^2(f + 4) \leq z\} - P\{\chi^2(f) \leq z\}] . \]

3. GENERATION OF COMPLEX WISHART DISTRIBUTED MATRICES

We wish to look into the distributions of \(-2 \rho \ln Q\) and the associated \(P\) for no difference or no change situations. Therefore, in this section we describe the generation of samples of complex Wishart distributed Hermitian, positive definite matrices [6] where we can control the parameters \(p, m, (and \ n), \) and \(\Sigma\).

3.1. The complex multivariate normal distribution

A \(p\)-dimensional random vector \(Z \in \mathbb{C}^p\) follows a complex multivariate normal distribution with mean \(0\) and dispersion matrix (variance-covariance matrix) \(\Sigma\), i.e., \(Z \sim N_C(0, \Sigma)\) if the frequency function is
\[ f(z) = \frac{1}{\pi^p |\Sigma|} \exp(-z^H \Sigma^{-1} z) \]
where the superscript \(^H\) denotes the complex conjugate and transpose.

The dispersion matrix \(\Sigma\) is Hermitian and positive definite. We may write it on the form
\[ \Sigma = \begin{pmatrix} \sigma_1^2 & (\alpha_{12} + i\beta_{12}) & \sigma_1 \sigma_2 & \cdots \\ (\alpha_{12} - i\beta_{12}) & \sigma_2^2 & \cdots & \\ (\alpha_{13} - i\beta_{13}) & \sigma_1 \sigma_3 & \cdots & \\ (\alpha_{13} + i\beta_{13}) & \sigma_2 \sigma_3 & & \\ & \vdots & \ddots & \end{pmatrix} . \]

We can write this as \(\Sigma = Re(\Sigma) + i Im(\Sigma)\).

If \(Z = X + iY\) then the \(2p\)-dimensional, real valued vector \([X^T \ Y]^T\) (where the superscript \(^T\) denotes transpose) has the variance-covariance matrix
\[ D \begin{bmatrix} X \\ Y \end{bmatrix} = \frac{1}{2} \begin{bmatrix} Re(\Sigma) & -Im(\Sigma) \\ Im(\Sigma) & Re(\Sigma) \end{bmatrix} . \]

3.2. The complex Wishart distribution

A Hermitian, positive definite random matrix \(W\) follows a complex Wishart distribution, i.e., \(W \sim W_C(p, n, \Sigma)\) if the frequency function is
\[ f(w) = \frac{1}{\Gamma_p(n)} |\Sigma|^{n/2} |w|^n \exp(\text{tr}(\Sigma^{-1} w)) \]
where \text{tr} denotes the trace and
\[ \Gamma_p(n) = \pi^{p(p-1)/2} \prod_{j=1}^p \Gamma(n - j + 1) \]
where \(\Gamma\) is the usual gamma function.

3.3. The empirical variance-covariance matrix for normal samples

Let \(Z_1, \ldots, Z_n\) be independent \(N_C(0, \Sigma)\)-distributed random variables. Let \(\hat{\Sigma}\) be the maximum likelihood estimate of \(\Sigma\). Then
\[ n \hat{\Sigma} = \sum_{j=1}^n Z_j Z_j^H \sim W_C(p, n, \Sigma) . \]

3.4. Linear transformations

If \(Z \sim N_C(0, \Sigma)\) then the linearly transformed variable \(AZ\) where \(A\) is a matrix of constants will again be normal, and \(AZ \sim N_C(0, A \Sigma A^H)\). If \(W \sim W_C(p, n, \Sigma)\) then \(AW A^H \sim W_C(p, n, A \Sigma A^H)\).

3.5. Generation of \(W_C(p, n, I)\) and \(W_C(p, n, \Sigma)\) distributed random variables

We consider observations \(Z_j = X_j + iY_j, j = 1, \ldots, n\) and organize the real and the imaginary parts in the data matrices
\[ X = \begin{bmatrix} X_1^T \\ X_n^T \end{bmatrix} = \begin{bmatrix} X_{11} & \cdots & X_{1p} \\ \vdots & \ddots & \vdots \\ X_{n1} & \cdots & X_{np} \end{bmatrix} \]
\[ Y = \begin{bmatrix} Y_1^T \\ Y_n^T \end{bmatrix} = \begin{bmatrix} Y_{11} & \cdots & Y_{1p} \\ \vdots & \ddots & \vdots \\ Y_{n1} & \cdots & Y_{np} \end{bmatrix} \]
and we assume that all \(X\) and \(Y\) variables are independent and \(N(0,1)\)-distributed. Then
\[ S = \frac{1}{2} \{X^T X + Y^T Y - i(X^T Y - Y^T X)\} \sim W_C(p, n, I) . \]

If \(C\) is the Cholesky factorization of the Hermitian, positive definite matrix \(\Sigma\), i.e., \(\Sigma = CC^H\), then
\[ W = C S C^H \sim W_C(p, n, \Sigma) . \]
4. EXPERIMENTS

To emulate polarimetric and dual polarization data from real SAR systems such as Radarsat-2 (http://www.asc-csa.gc.ca/eng/satellites/radarsat2) and Sentinel-1 (https://sentinel.esa.int/web/sentinel/missions/sentinel-1), we use the above to generate two instances of samples of complex Wishart distributed $3 \times 3$, $2 \times 2$ as well as $2 \times 2$ diagonal-only matrices. In this latter case, the matrices are not complex (and they are not Wishart distributed). However, in [2] we show that also in this case we can use the tools developed there for detection of differences such as change over time in (polarimetric) SAR data, see also [4,7–9].

Two experiments are performed, one with very different equivalent numbers of looks, ENL, 10 and 100. The other experiment deals with combinations of ENL up to 20 in the two instances.

4.1. Very different ENL

In a first experiment, we generate 128k samples of independent complex Wishart distributed Hermitian, positive definite matrices for each of two instances with the same $\Sigma$, one with $L_1 = 100$ looks, one with $L_2 = 10$ looks. This corresponds to no change in a change detection setting or to no target in a target detection setting [10] where the number of looks for a centre pixel and its surrounding pixels are very different. With no change or target, we know the asymptotic distribution of the test statistic $-2\rho \ln Q$ and the associated $P$-value.

Figure 1 shows histograms for this no target situation for $3 \times 3$ matrices, $-2\rho \ln Q$ and the theoretical distribution in red (top), and $P$ (bottom).

Figure 2 shows histograms for the no target situation for $2 \times 2$ matrices, $-2\rho \ln Q$ and the theoretical distribution in red (top), and $P$ (bottom).

Figure 3 shows histograms for the no target situation for $2 \times 2$ diagonal-only matrices, $-2\rho \ln Q$ and the theoretical distribution in red (top), and $P$ (bottom).

In all three cases depicted in Figures 1, 2 and 3 the histograms very beautifully show the expected behaviour for no target: the combination of $\chi^2$ distributions of $-2\rho \ln Q$ and the uniform distribution of $P$.

4.2. ENL up to 20

In a second experiment, we generate 128k samples of independent complex Wishart distributed data for each of two instances with the same $\Sigma$. Results are described in detail in the subsections below, Again, this experiment emulates the situation with no change between two time points in a change detection setting. Ideally, of course, there is no change and we therefore know the asymptotic distributions of the test statistic $-2\rho \ln Q$ (and the simpler $-2 \ln Q$) and the associated $P$-value. The calculations are carried out for a number of combinations of ENL for both “time points”. In the subsections below, we show plots of the mean value of the 128k samples of $-2\rho \ln Q$ (and $-2 \ln Q$) and the associated $P$-values as functions of ENL. Again, ENL at “time point one” is called $L_1$ and at “time point two” $L_2$. Ideal surfaces and plots are flat and horizontal.

A few runs with different realizations of $\Sigma$ (not shown) indicate that the resulting figures and plots vary only little with $\Sigma$. 

![Fig. 1. Histograms for the no target situation with $L_1 = 100$ and $L_2 = 10$ for $3 \times 3$ matrices, $-2\rho \ln Q$ and the theoretical distribution in red (top), and $P$ (bottom).](image1)

![Fig. 2. Histograms for the no target situation with $L_1 = 100$ and $L_2 = 10$ for $2 \times 2$ matrices, $-2\rho \ln Q$ and the theoretical distribution in red (top), and $P$ (bottom).](image2)

![Fig. 3. Histograms for the no target situation with $L_1 = 100$ and $L_2 = 10$ for $2 \times 2$ diagonal-only matrices, $-2\rho \ln Q$ and the theoretical distribution in red (top), and $P$ (bottom).](image3)
4.3. **3×3 matrices**

Figures 4, 5, 6 and 7 show mean values of the test statistics $-2\rho \ln Q$, $-2 \ln Q$ and associated $P$-values for 128k samples generated for two instances of matrices with the same $\Sigma$ corresponding to no-change in a change detection setting.

$-2\rho \ln Q$ and its associated $P$-value clearly outperform $-2 \ln Q$ and its $P$-value. For same ENL, $-2\rho \ln Q$ and its $P$-value stop varying as functions of ENL at a value of around 9.

Ideal surfaces and plots are flat and horizontal, for $-2\rho \ln Q$ ideal values are just over 9 (depending on $\rho$ and therefore on ENL), for $-2 \ln Q$ ideal values are 9, and for $P$ ideal values are 1/2.

---

**Fig. 4.** Mean values of the test statistic $-2\rho \ln Q$ (top) for 128k samples generated for two instances of $3 \times 3$ complex Wishart distributed matrices with the same $\Sigma$, i.e., a no-change situation. The test statistic is shown as a function of the number of looks at the two instances. The associated $P$-value is shown at the bottom.

**Fig. 5.** Mean values of the test statistic $-2\rho \ln Q$ (top) for 128k samples generated for two instances of $3 \times 3$ complex Wishart distributed matrices with the same $\Sigma$, i.e., a no-change situation. The test statistic is shown as a function of the same number of looks at the two instances. The associated $P$-value is shown at the bottom.

**Fig. 6.** Mean values of the test statistic $-2 \ln Q$ (top) for 128k samples generated for two instances of $3 \times 3$ complex Wishart distributed matrices with the same $\Sigma$, i.e., a no-change situation. The test statistic is shown as a function of the number of looks at the two instances. The associated $P$-value is shown at the bottom.

**Fig. 7.** Mean values of the test statistic $-2 \ln Q$ (top) for 128k samples generated for two instances of $3 \times 3$ complex Wishart distributed matrices with the same $\Sigma$, i.e., a no-change situation. The test statistic is shown as a function of the same number of looks at the two instances. The associated $P$-value is shown at the bottom.

Note the different scales and the different orientations of the $L_1$ and $L_2$ axes in Figures 4 to 7.
4.4 $2 \times 2$ matrices

Figures 8, 9, 10 and 11 show mean values of the test statistics $-2\rho \ln Q$, $-2 \ln Q$ and associated $P$-values for 128k samples generated for two instances of matrices with the same $\Sigma$ corresponding to no-change in a change detection setting.

$-2\rho \ln Q$ and its associated $P$-value clearly outperform $-2 \ln Q$ and its $P$-value. For same ENL, $-2\rho \ln Q$ and its $P$-value clearly outperform $-2 \ln Q$ and its $P$-value. For same ENL, $-2\rho \ln Q$ and its $P$-value stop varying as functions of ENL at a value of around 4.

Ideal surfaces and plots are flat and horizontal, for $-2\rho \ln Q$ ideal values are just over 4 (depending on $\rho$ and therefore on ENL), for $-2 \ln Q$ ideal values are 4, and for $P$ ideal values are 1/2.

Fig. 8. Mean values of the test statistic $-2\rho \ln Q$ (top) for 128k samples generated for two instances of $2 \times 2$ complex Wishart distributed matrices with the same $\Sigma$, i.e., a no-change situation. The test statistic is shown as a function of the number of looks at the two instances. The associated $P$-value is shown at the bottom.

Fig. 9. Mean values of the test statistic $-2\rho \ln Q$ (top) for 128k samples generated for two instances of $2 \times 2$ complex Wishart distributed matrices with the same $\Sigma$, i.e., a no-change situation. The test statistic is shown as a function of the same number of looks at the two instances. The associated $P$-value is shown at the bottom.

Fig. 10. Mean values of the test statistic $-2 \ln Q$ (top) for 128k samples generated for two instances of $2 \times 2$ complex Wishart distributed matrices with the same $\Sigma$, i.e., a no-change situation. The test statistic is shown as a function of the number of looks at the two instances. The associated $P$-value is shown at the bottom.

Note the different scales and the different orientations of the $L_1$ and $L_2$ axes in Figures 8 to 11.
4.5. 2×2 matrices, diagonal only

Figures 12, 13, 14 and 15 show mean values of the test statistics $-2\rho \ln Q$, $-2 \ln Q$ and associated $P$-values for 128k samples generated for two instances of matrices with the same $\Sigma$ corresponding to no-change in a change detection setting.

$-2\rho \ln Q$ and its associated $P$-value clearly outperform $-2 \ln Q$ and its $P$-value. In this case, $-2\rho \ln Q$ and its $P$ seem not to vary as functions of ENL.

Ideal surfaces and plots are flat and horizontal, for $-2\rho \ln Q$ ideal values are just over 2 (depending on $\rho$ and therefore on ENL), for $-2 \ln Q$ ideal values are 2, and for $P$ ideal values are 1/2.

**Fig. 12.** Mean values of the test statistic $-2\rho \ln Q$ (top) for 128k samples generated for two instances of $2 \times 2$ diagonal-only matrices with the same $\Sigma$, i.e., a no-change situation. The test statistic is shown as a function of the number of looks at the two instances. The associated $P$-value is shown at the bottom.

**Fig. 13.** Mean values of the test statistic $-2\rho \ln Q$ (top) for 128k samples generated for two instances of $2 \times 2$ diagonal-only matrices with the same $\Sigma$, i.e., a no-change situation. The test statistic is shown as a function of the same number of looks at the two instances. The associated $P$-value is shown at the bottom.

**Fig. 14.** Mean values of the test statistic $-2 \ln Q$ (top) for 128k samples generated for two instances of $2 \times 2$ diagonal-only matrices with the same $\Sigma$, i.e., a no-change situation. The test statistic is shown as a function of the number of looks at the two instances. The associated $P$-value is shown at the bottom.

**Fig. 15.** Mean values of the test statistic $-2 \ln Q$ (top) for 128k samples generated for two instances of $2 \times 2$ diagonal-only matrices with the same $\Sigma$, i.e., a no-change situation. The test statistic is shown as a function of the same number of looks at the two instances. The associated $P$-value is shown at the bottom.

Note the different scales and the different orientations of the $L_1$ and $L_2$ axes in Figures 12 to 15.
5. CONCLUSIONS
We have used generated data to simulate no target in a target detection setting and no change in a change detection setting in (polarimetric) bi-temporal SAR data. We have generated 3×3 Hermitian, positive definite matrices, 2×2 Hermitian, positive definite matrices, and 2×2 diagonal-only matrices to emulate polarimetric data (from for example Radarsat-2) and dual polariometry data (from for example Sentinel-1).

The first experiment in a target detection setting with very different number of looks at the two instances (one instance is a (few) centre pixel(s), the other instance represents the surrounding background pixels) gives histograms which very beautifully show the expected behaviour for no target including the uniform distribution of $P$.

The second experiment shows that $-2\rho \ln Q$ and its associated $P$-value clearly outperform $-2 \ln Q$ and its $P$-value. For same ENL, $-2\rho \ln Q$ and the associated $P$-value stop varying as functions of ENL at a value corresponding to the number of degrees of freedom.

A few runs with different realizations of $\Sigma$ (not shown) indicate that the resulting figures and plots vary only little with $\Sigma$.

Matlab code wc_rand.m to generate complex Wishart distributed, Hermitian, positive definite covariance matrix samples and to be used together with already published software to do the Wishart based change detection [2,4,8,9], is available on Allan Nielsen’s homepage.

6. REFERENCES


