Article

The Performance Analysis Based on SAR Sample Covariance Matrix

Esra Erten

Institute of Environmental Engineering, ETH Zurich, Schafmattstr. 6, CH-8093 Zurich, Switzerland;
E-Mail: erten@ifu.baug.ethz.ch; Tel.: +41-44-633-4059

Abstract: Multi-channel systems appear in several fields of application in science. In the Synthetic Aperture Radar (SAR) context, multi-channel systems may refer to different domains, as multi-polarization, multi-interferometric or multi-temporal data, or even a combination of them. Due to the inherent speckle phenomenon present in SAR images, the statistical description of the data is almost mandatory for its utilization. The complex images acquired over natural media present in general zero-mean circular Gaussian characteristics. In this case, second order statistics as the multi-channel covariance matrix fully describe the data. For practical situations however, the covariance matrix has to be estimated using a limited number of samples, and this sample covariance matrix follow the complex Wishart distribution. In this context, the eigendecomposition of the multi-channel covariance matrix has been shown in different areas of high relevance regarding the physical properties of the imaged scene. Specifically, the maximum eigenvalue of the covariance matrix has been frequently used in different applications as target or change detection, estimation of the dominant scattering mechanism in polarimetric data, moving target indication, *etc*. In this paper, the statistical behavior of the maximum eigenvalue derived from the eigendecomposition of the sample multi-channel covariance matrix in terms of multi-channel SAR images is simplified for SAR community. Validation is performed against simulated data and examples of estimation and detection problems using the analytical expressions are as well given.

1. Introduction

Multi-channel systems with random nature appear in a wide range of fields in the literature. For several of them, the central limit theorem applies and their random behavior can be modeled by Gaussian statistics, being thus statistically fully described by the first and second order moments. In the case of multi-channel SAR systems, the assumption of zero-mean multivariate complex Gaussian distribution is frequently valid for geophysical media, being thus fully described by the complex Hermitian covariance matrix. This is the statistical case treated in this work.

Since a SAR signal corresponds to the superposition of the scattered fields from all scatters inside a resolution cell, physical parameter estimation may be performed by exploiting the multi-channel covariance matrix. For that, eigendecomposition theorems have been widely used in the literature, decomposing the whole covariance matrix into elementary quantities, in order to provide a better physical interpretation of the data.

The eigendecomposition of the multi-channel SAR covariance matrix leads to useful information in a large range of SAR applications, including Ground Moving Target Indication (GMTI) [1], polarimetric SAR (PolSAR) [2], interferometric SAR (InSAR) [3], polarimetric interferometric SAR (PolInSAR) [4], change detection [5], target detection [6] and filtering [7,8]. Particularly, the covariance matrix maximum eigenvalue has been proved to be a key parameter in various areas. In [5] and [9], the change detection problem was elaborated using the maximum eigenvalue of the interferometric covariance matrix. The eigendecomposition of the SAR covariance matrix with application to SAR GMTI is used in and to calculate the probability of moving target detection in homogeneous and heterogeneous terrain. Regarding polarimetry, the eigendecomposition of the coherency (which is also a covariance) matrix is the most common tool among incoherent target decomposition theorems for the interpretation of the earth's surface. The maximum eigenvalue of the coherency matrix is also often used in polarimetry, due to the close relationship to the dominant scattering mechanism of the scattering process [2,10,11].

Although there is a wide area of applications regarding the maximum eigenvalue of the multi-channel covariance matrix, a certain lack of information in the literature concerning its statistical characterization in terms of multi-channel SAR may be verified. The statistical description can be useful to understand bias effects and to analyze estimation as well as detection problems. The estimated (sample) covariance matrix follows a Wishart distribution, which has been of interest along the years since its first derivation. In areas as communication systems, several works using such results have been published [12–16]. The application of the Wishart distribution in remote sensing was however considerably late introduced [17].

The cornerstone study considering the statistical description of the covariance matrix eigendecomposition in polarimetry has been carried out in [18]. The majority of the analysis in [18] was performed on the basis of numerical methods. In this paper the results of [18] is supported by addressing analytical solutions. Additionally, it is derived an exact closed form expressions for the Moment Generating Function (MGF) of the sample covariance matrix maximum eigenvalue.

The validation of the derived expressions is performed using simulated data, demonstrating its agreement with the theory. The effect of the number of samples and underlying correlation scenario of the sample covariance matrix on the bias in the estimation of the maximum eigenvalue is also

investigated. Finally, examples of applications are as well given, in the fields of estimation and detection theory.

The next section reminds the basics of the statistical description of multi-channel SAR systems and of the covariance matrix eigendecomposition. It also includes the derived theorems presenting the statistical founds. Section 3 includes numerical examples validating the theorems as well as an analysis of their behavior. Sections 4 and 5 show their usage for the analysis of the estimation bias and for the elaboration of detection problems, respectively. Section 6 concludes the paper with discussions and directions for future work. The detailed statistical derivations can be found in the Appendix.

2. Statistical Characteristics of the Maximum Eigenvalue of the Multi-Channel Sample Covariance Matrix

2.1. Preliminaries

The individual elements of a m-dimensional multi-channel system may be organized in a m-dimensional vector k. As the elements are assumed to follow zero mean complex Gaussian distributions, the vector k is said to follow a m-dimensional multivariate normal distribution, with zero mean and true covariance matrix Σ among the vector elements, and is represented by $k \sim \mathcal{N}_m^C(0, \Sigma)$ [19,20]. For zero mean Gaussian statistics the covariance matrix fully describes the data, playing a key role in several application fields. In practical situations however, the true covariance matrix Σ is unknown and has to be estimated by its maximum likelihood estimator (MLE), the sample covariance matrix $Z = (1/n) \sum_{j=1}^n k_j k_j^\dagger$, where n is the number of estimation samples and $\dagger$ is the transpose conjugate operator. In the SAR context, the number of independent samples is also called *looks*. The elements of Z follow a m dimensional complex Wishart distribution with n degrees of freedom and true covariance matrix Σ, represented by $Z \sim \mathcal{W}_m^C(n, \Sigma)$ and defined as [20]:

$$\mathrm{p}_Z(Z) = \frac{n^{mn}|Z|^{n-m} \mathrm{etr}\left(-n\Sigma^{-1}Z\right)}{|\Sigma|^n \tilde{\Gamma}_m(n)} \qquad \text{with} \qquad \tilde{\Gamma}_m(n) = \pi^{m(m-1)/2} \prod_{i=1}^m \Gamma(n-i+1) \qquad (1)$$

where $\Gamma(\cdot)$ is the gamma function and $\mathrm{etr}(\cdot)$ is the exponential trace of a matrix.

The spectral theorem from linear algebra allows the decomposition of the full rank m-dimensional covariance matrix in a set of m one-rank covariance matrices using its eigenvalues and eigenvectors. Accordingly, the decompositions of the true covariance matrix $\Sigma = \sum_{i=1}^m l_i(e_i e_i^\dagger)$ and its estimator $Z = \sum_{i=1}^m \lambda_i(e_i' e_i'^\dagger)$ are given by

$$\Sigma = Q \begin{bmatrix} l_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & l_m \end{bmatrix} Q^\dagger \text{ and } Z = Q' \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_m \end{bmatrix} Q'^\dagger, \quad \begin{aligned} Q &= [e_1, e_2, \cdots, e_m] \\ Q' &= [e_1', e_2', \cdots, e_m'] \end{aligned} \qquad (2)$$

with their real non-negative eigenvalues l_i and λ_i, and respective complex eigenvectors e_i, e_i', for $i = \{1, ...m\}$.

2.2. Sample Maximum Eigenvalue Statistical Description

The following Theorem III concerns the derived *Moment Generating Function* (MGF) of the maximum eigenvalue of the sample covariance matrix Z of a multi-channel statistical system, which is a critical step in removing the bias of the largest eigenvalue. The Theorem III, which is derived due to the previous works (Theorems I and II), is the main result of the paper. It is detailed in the appendix and will be used for the elaboration of illustrative estimation and detection problems in Sections 4 and 5.

Throughout the paper, $|\cdot|$ is the matrix determinant, $\langle X \rangle_n = (1/n) \sum_{i=1}^{n} X_i$ denotes the estimator of the random matrix X formed from a sample of size n.

Theorem I: Let $k \sim \mathcal{N}_m^C(0, \Sigma)$ be a m-dimensional complex vector whose elements follow zero mean Gaussian distributions with associated $m \times m$ covariance matrix Σ. Let Σ have $l_m \leq \dots \leq l_1$ eigenvalues. Then the *Cumulative Density Function* (CDF) of the maximum eigenvalue λ_{max} of the sample covariance matrix $\langle k k^\dagger \rangle_n$, with the assumption $m \leq n$, is given by

$$F_{\lambda_{max}}(x) = \mathcal{S} |\Psi(x)|, \quad \text{with constant term} \quad \mathcal{S} = \frac{\pi^{m(m-1)} n^{m(2n-m+1)/2}}{\tilde{\Gamma}_m(m) \tilde{\Gamma}_m(n)} \frac{\prod_{k=1}^{m-1} k^{m-k}}{\prod_{i=1}^{m} l_i^n \prod_{i<j}^{m} \left(\frac{1}{l_j} - \frac{1}{l_i} \right)} \quad (3)$$

where $\Psi(x)$ is a $m \times m$ matrix with its (i,j)th element $\Psi(x)_{i,j} = \frac{\gamma(n+1-j, x\frac{n}{l_i})}{(\frac{n}{l_i})^{(n+1-j)}}$, and γ is the incomplete Gamma function (Equation (2.42) in [21]).

Proof: [14] gave the closed expression of CDF of the largest eigenvalue of the complex Wishart matrices. Here, it is written in the form of SAR covariance matrix after small continuation, see also Appendix A.

Theorem II: Let $k \sim \mathcal{N}_m^C(0, \Sigma)$ be a m-dimensional complex vector whose elements follow zero mean Gaussian distributions with associated $m \times m$ covariance matrix Σ. Let Σ have $l_m \leq \dots \leq l_1$ eigenvalues. Then the *Probability Density Function* (PDF) of the maximum eigenvalue λ_{max} of the sample covariance matrix $\langle k k^\dagger \rangle_n$, with the assumption of $m \leq n$, is given by

$$p_{\lambda_{max}}(x) = \mathcal{S} |\Psi(x)| \operatorname{tr} \left[\Psi(x)^{-1} \Omega(x) \right] \quad (4)$$

where $\Omega(x)$ is an $m \times m$ matrix with its i,jth elements $\Omega(x)_{i,j} = \exp\left(-\frac{n}{l_i} x \right) x^{n-j}$, and $\Psi(x)$ and $\mathcal{S}$ are defined in Equation (3).

Proof: Equation (4) is obtained by differentiating Equation (3) with respect to x using (Equation (9) in [22])

$$\frac{d}{dt} |X(t)| = |X(t)| \operatorname{tr} \left(X(t)^{-1} \frac{d}{dt} X(t) \right). \quad (5)$$

Here it can be noted that when the true covariance matrix Σ is diagonal, the PDF of the largest eigenvalue reduces to Ermolaev and Rodyushkin result [23].

Theorem III: Let $k \sim \mathcal{N}_m^C(0, \Sigma)$ be a m-dimensional complex vector whose elements follow zero mean Gaussian distributions with associated $m \times m$ covariance matrix Σ. Let Σ have $l_m \leq \dots \leq l_1$

eigenvalues. Then for any positive integers s, the sth moment of the maximum eigenvalue λ_{max} of the sample covariance matrix $\langle kk^\dagger \rangle_n$, with the assumption of $m \leq n$, is given by

$$E(\lambda_{max}^s) = \mathcal{S} \sum_{i,j=1}^{m} (-1)^{i+j} \sum_{\pi_s \in \pi_{sub} \in S_m} \text{sgn}(\pi_s) \prod_{k=1}^{m-1} \frac{1}{n+1-\pi_s(k)} (\frac{n}{l_k})^{\frac{n-\pi_s(k)}{2}} a_k^{u_k} (b+A)^{-v-M} \Gamma(v+M) \times$$

$$F_A \left(v + M; u_1 - \lambda_1, \ldots, u_{m-1} - \lambda_{m-1}; 2u_1, \ldots, 2u_{m-1}; \frac{a_1}{b_1 + A}, \ldots, \frac{a_{m-1}}{b_{m-1} + A} \right)$$

$$\text{if } k = j \,\wedge\, \pi_s(j) = i, \; \pi_s \in \pi_{sub}$$

(6)

where $\mathcal{S}$ indicates the constant term as in *Theorem I* and $F_A(\cdots)$ is the hypergeometric function of several variables with $a_k = \frac{n}{l_k}$, $u_k = (2 + n - \pi_{sub}(k))/2$, $v = 1 + s + n - i + \sum_{k=1}^{m-1} \frac{n-\pi_{sub}(k)}{2}$, $b = \frac{n}{l_j} + \sum_{k=1}^{m-1} \frac{n}{2l_k}$, $\lambda_k = \frac{n-\pi_{sub}(k)}{2}$, $M = \sum_{k=1}^{m-1} u_k$, $A = \frac{1}{2} \sum_{k=1}^{m-1} a_k$. Here, the sum is computed over $(m-1)!$ permutations of π_s. S_m denotes the set of all $m!$ permutations of the set $S = \{1, 2, ..., m\}$, and $\text{sgn}(\pi_s)$ denotes the sign of the permutation π_s : $+1$ if π_s is an even permutation and -1 if it is odd.

Proof: See Appendix section B. In the appendix the moment of sample maximum eigenvalue is also given in more friendlier form for programming obtained by splitting the hypergeometric function of several variables into a sum of its variables.

2.3. Dependence of the Covariance Matrix Eigenvalues

Before starting the following analysis, it would be nice to illustrate the dependencies of the eigenvalues on the covariance matrix parameters, which is necessary for the further understanding of the addressed topics. For that, a two-dimensional covariance matrix is used, originated from the expected values of the outer product of the vector $k = [k_1 \quad k_2]^T$, *i.e.*,

$$\Sigma = E(kk^\dagger) = \begin{bmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \rho e^{j\phi} \\ \sigma_1 \sigma_2 \rho e^{-j\phi} & \sigma_2^2 \end{bmatrix} \tag{7}$$
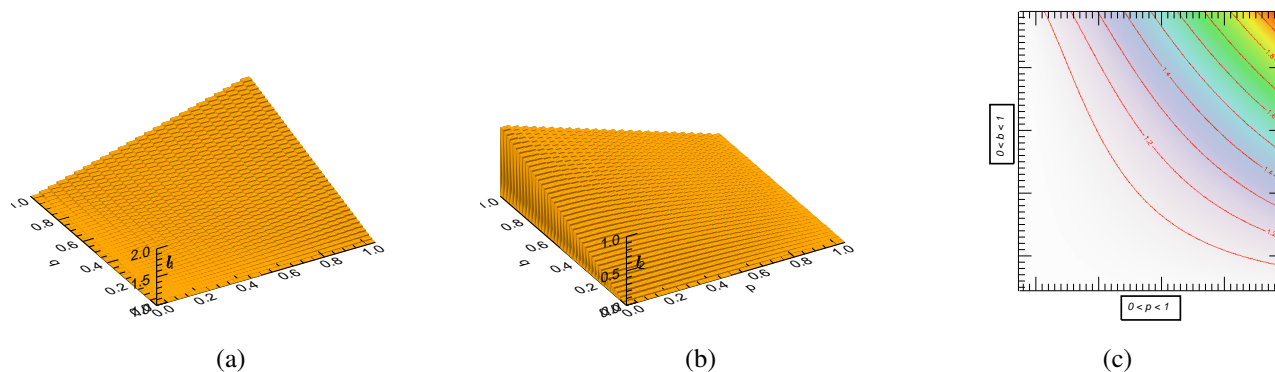
where σ_1^2, σ_2^2 are the variance or power of k_1 and k_2, respectively, ρ is the absolute and ϕ the phase of their complex correlation coefficient, and T means transpose.

The eigenvalues of the true covariance matrix can be shown to be given by [5]:

$$l_{1,2} = \frac{1}{2} \left[\frac{b+1}{\sqrt{b}} \pm \sqrt{\frac{b^2+1}{b} - 2 + 4\rho^2} \right], \quad \sigma_1^2 = b\sigma_2^2 . \tag{8}$$

The behavior of l_1 and l_2 is presented in Figure 1, as a function of a normalized power ratio b and a correlation ρ. Note their mutual dependence, making clear that b and ρ directly determine the behavior of both eigenvalues.

Figure 1. Two-dimensional system configuration. (**a**) True eigenvalue l_1 and (**b**) true eigenvalue l_2 as a function of correlation and normalized power ratio b. (**c**) Counter plots of l_1.

(a) (b) (c)

3. Validation and Analysis of the Theoretical Expressions

This section aims to validate the theorems mentioned in the previous section using simulated data. The simulated data have been generated using different multi-dimensional configurations, where the correlation between channels have been generated using the well known Mahalanobis transformation [24].

Figure 2. (**a**) Comparison between the theoretical distributions and the histograms obtained from simulated data of the maximum sample eigenvalue. (**b**) The distribution of the maximum sample eigenvalue as a function of the number of samples, for a three-dimensional system. In both case, the powers are given by $\sigma_{k_1} = \sigma_{k_2} = \sigma_{k_3} = 1$ and correlations by $\rho_{k_1 k_2} = 0$, $\rho_{k_1 k_3} = 0.8$ and $\rho_{k_2 k_3} = 0$. When $n \to \infty$, $\lambda_{max} = l_1 = 1.8$.
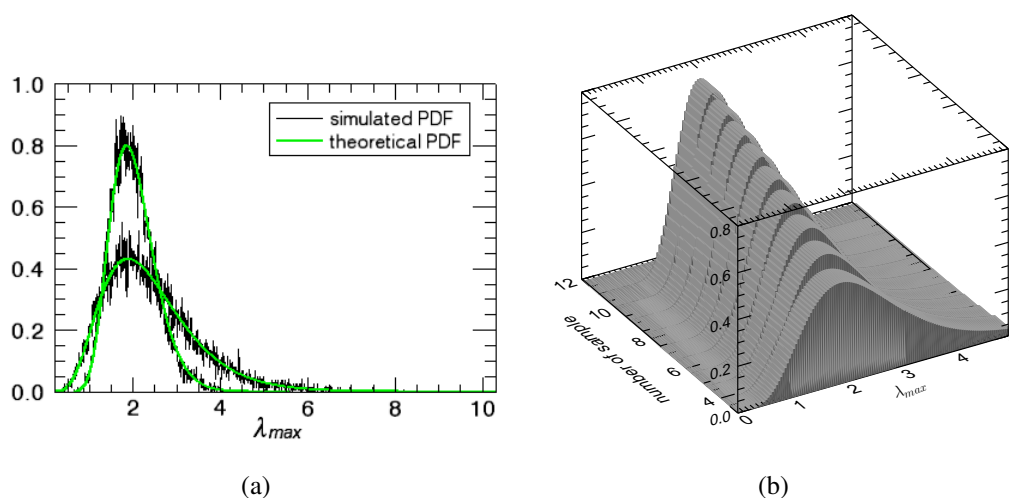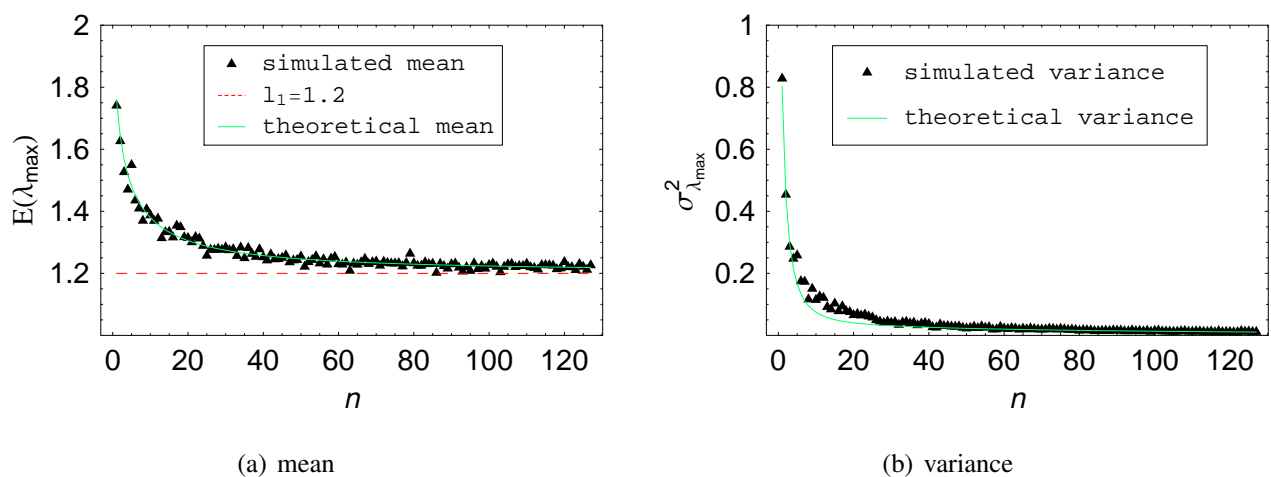

(a) (b)

Figure 2(a) shows the comparison of the Equation (4) with simulations. The theoretical PDF curves clearly agree with the histograms obtained from simulated data. As expected, the PDFs become narrower with increasing n, indicating less variance around the true value of the maximum eigenvalue $l_1 = 1.8$. This behavior can be better seen in Figure 2(b) where the distribution of λ_{max} as a function of the

number of estimation samples n is presented. Note also that the expected value of the distributions seems to change for different n.

Figure 3(a) shows the variation of the histogram mean as a function of n for a two-dimensional case with fix correlation $\rho = 0.2$. In Figure 3(a), the theoretical expected value of λ_{max} has been also over plotted for comparison. The curves match well, which validates *Theorem III*. The same has been carried out for the second order moment, *i.e.*, the variance of λ_{max}, and is presented in Figure 3(b). The agreement between the theoretical variance of λ_{max} and the variance of the simulated data can again be confirmed. Also, as already observed in Figure 2(a), the lower the number of samples n, the higher the variance of λ_{max}.

Figure 3. Theoretical *versus* simulation results (**a**) for the first order statistics (mean), and (**b**) for the second order statistics (variance), of the sample maximum eigenvalue of a two-dimensional system with powers $\sigma_{k_1} = \sigma_{k_2} = 1$ and correlation $\rho_{k_1 k_2} = 0.2$.

(a) mean

(b) variance

The impact of the correlation and the number of samples on the behavior of the third (skewness) and fourth (kurtosis) order moment of the maximum sample eigenvalue λ_{max} is presented in Figure 4. Skewness is a measure of how symmetrical the distribution is with respect to its mean. Figure 4(a,b) indicates that the skewness of λ_{max} converges to zero for increasing n and increasing ρ, expressing the tendency to a symmetrical distribution in that cases.

Kurtosis is a measure of the peakedness of the distribution. Figure 4(c,d) shows that for increasing n and ρ, the kurtosis of the λ_{max} distribution tends to three, which corresponds to the kurtosis of the normal Gaussian distribution.

4. Estimation Bias

Figure 5 shows the expected value (first order moment) of the maximum sample eigenvalue λ_{max} as a function of the true eigenvalues l_{max} for a fixed number of samples $n = 3$ and $k = 1$. The variation of the underlying correlation of the true covariance matrix ρ, which changes the values of l_{max}, has been as well indicated in a color code. Observe that the higher the correlation is, the higher l_{max} becomes.

Figure 4. Theoretical results for the third (skewness) and the fourth (kurtosis) order statistics of the sample maximum eigenvalue of a 2D system having powers $\sigma_{k_1} = \sigma_{k_2} = 1$ and correlations $\rho = \{0.2, 0.3, \ldots, 0.9\}$ versus the number of samples $n = \{2, 3, \ldots, 62\}$.
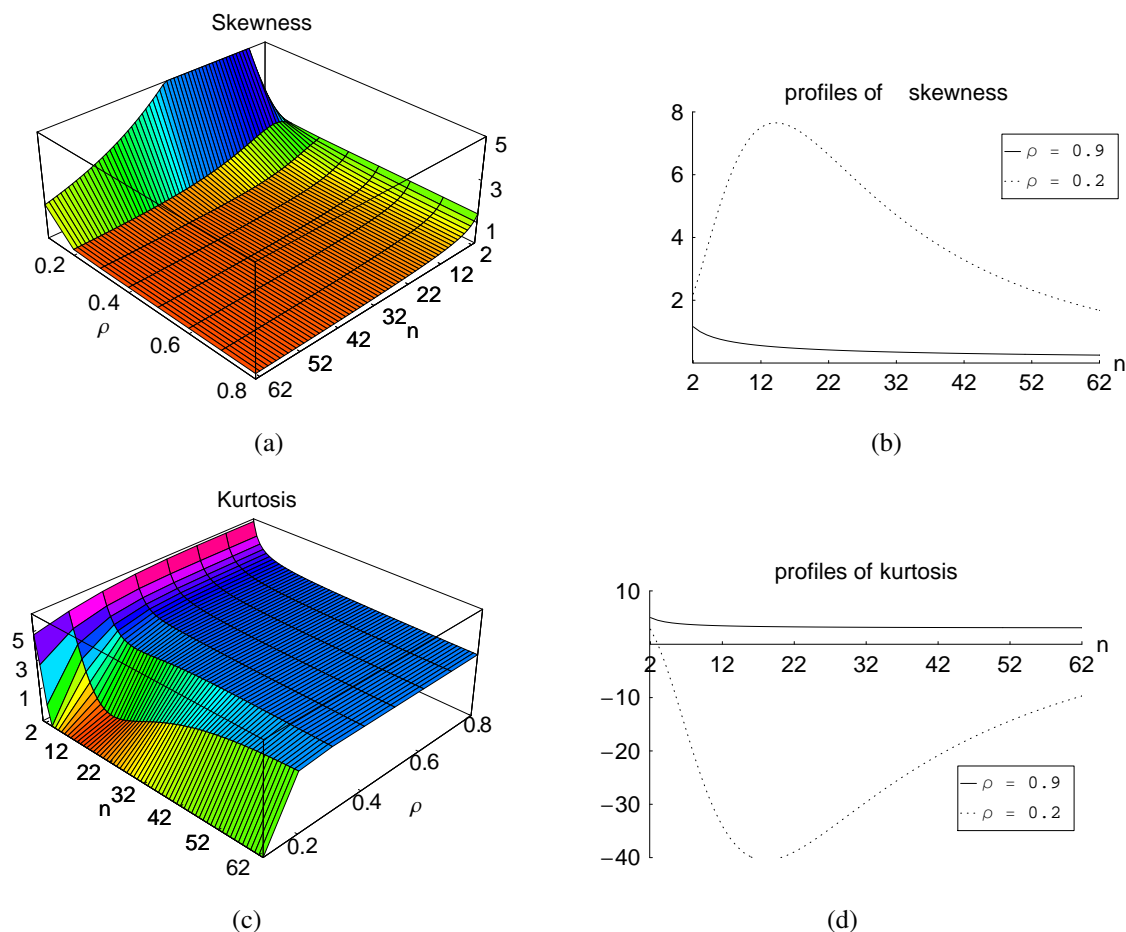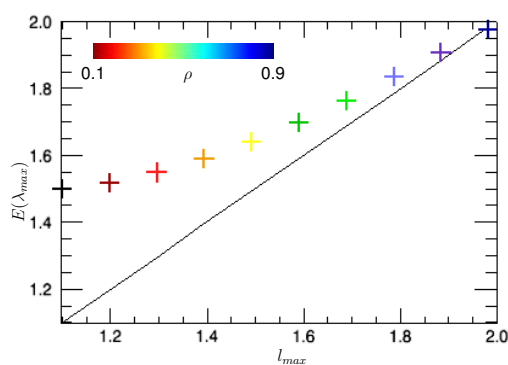

(a)

(b)

(c)

(d)

Figure 5. Effects of the correlation between channels on the expected value of the maximum sample eigenvalue keeping $n = 3$ and $k = 1$.
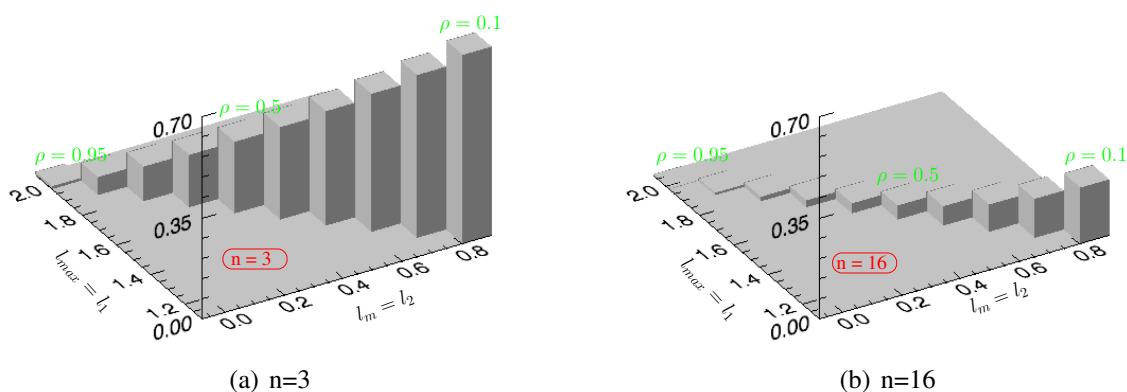

The *bias* of the estimator of a certain parameter is defined as the difference between the expected value of the estimator and the true value of the parameter. In the present case, the bias is hence given by

$E[\lambda_{max}] - l_{max}$. Notice thus from Figure 5 that the bias of λ_{max} becomes very strong for low values of the correlation l_{max} (or equivalently, low values of ρ).

The effect is emphasized in Figure 6(a,b), where the bias of λ_{max} is presented as a function of l_{max} and l_2, for $n = 3$ and $n = 16$, respectively. The values of the underlying correlation ρ have been also indicated. The bias is small either when the correlation is high or the number of samples is sufficiently large. When the correlation between channels is low, the bias becomes very significant and a large number of samples is necessary to decrease the estimation bias.

Figure 6. The bias, $E[\lambda_{max}] - l_{max}$, of the sample maximum eigenvalue with the number of samples 3 and 16 in various correlated channels having a standard deviation of $\sigma_{k_1} = \sigma_{k_2} = 1$.

(a) n=3   (b) n=16

5. Analysis of Detection Problems

Detection theory is a means to quantify the ability of a procedure to detect a parameter (or signal) immersed in a noise environment. A decision has to be taken, in order to say yes, there is a signal, or no, there is no signal. Such decision is usually taken under the application of a decision threshold. Noise contributes off course negatively inducing wrong decisions. For the quantification of detection accuracy, some quantities are usually defined as the probability of detection (PD) and probability of false alarm (PFA), allowing a detection problem analysis.

The maximum eigenvalue of covariance matrices has been frequently used in the literature for the elaboration of certain detection problems. In the Ground Moving Target Indication (GMTI) area, for instance, the signal to clutter plus noise ratio has been studied under the distribution of the sample eigenvalues, which allowed the implementation of a constant false alarm rate detector [9]. The eigenvalues of the covariance matrix of a SAR image pair have been also used in order to formulate a problem in the change detection area [5]. Another example is the determination of the existence of just a single dominant scattering mechanism in a SAR polarimetric acquisition, which also has been evaluated using a threshold in the maximum sample eigenvalue of the polarimetric covariance matrix [11]. Target detection and polarimetric filtering represent other fields of application in which the maximum eigenvalue can be used.

Having the closed form expressions of the PDF (Equation (4)) and/or CDF (Equation (3)) of the maximum sample eigenvalue of the covariance matrix, the PD and PFA when applying a threshold in λ_{max} can be analytically computed, allowing a complete detection problem analysis.

5.1. Detection of a Dominant Maximum Eigenvalue

Since several detection problems rely in fact on the choice of a dominant or non-dominant l_{max} by applying a threshold in λ_{max}, the following problem with two hypotheses is elaborated

$$H_0 \; (l_{max} \text{ is dominant}) : l_{max} = l_1 \quad \text{and} \quad l_i = 0$$
$$H_1 \; (l_{max} \text{ is not dominant}) : l_{max} = l_1 = l_i$$

for $i = \{2, 3, \ldots, m\}$.

In one application area H_0 may mean that just a single scattering mechanism is present inside a resolution cell of polarimetric SAR data, in other application H_0 can mean that a change happened or a moving target is present in the scene.

Since l_{max} is not achievable, a threshold $\mathcal{T}$ has to be used in λ_{max}, originating the following PD and PFA

$$
\begin{aligned}
PD &= \text{p[accepting } H_0 \mid H_0 \text{ is true]} \\
&= \text{p}[\lambda_{max} > \mathcal{T} \mid H_0 \text{ is true]} = \int_{\mathcal{T}}^{\infty} \text{p}_{\lambda_{max}}(x; H_0)\, \mathrm{d}x \\
PFA &= \text{p[accepting } H_0 \mid H_1 \text{ is true]} \\
&= \text{p}[\lambda_{max} > \mathcal{T} \mid H_1 \text{ is true]} = \int_{\mathcal{T}}^{\infty} \text{p}_{\lambda_{max}}(x, H_1)\, \mathrm{d}x
\end{aligned}
$$

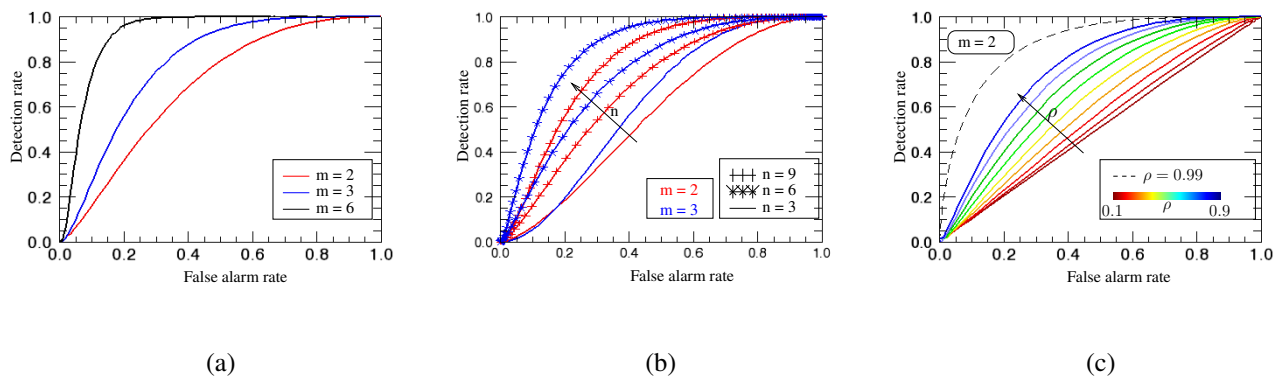as a function of the decision threshold $\mathcal{T}$.

Hence, for a two-dimensional system configuration, the distribution $\text{p}_{\lambda_{max}}(x; H_0)$ is the distribution of λ_{max} when $l_1 = 2$ and $l_2 = 0$. For a three-dimensional system, $\text{p}_{\lambda_{max}}(x; H_0)$ is determined evaluating the distribution of λ_{max} when $l_1 = 3$ and $l_2 = l_3 = 0$. Higher dimensional system configurations follow the same rule.

On the other hand, the distribution $\text{p}_{\lambda_{max}}(x; H_1)$ is given by $l_1 = l_2 = 1$ and $l_1 = l_2 = l_3 = 1$ for a two- and three-dimensional system, respectively.

For each value of $\mathcal{T}$, there exists a pair (PFA, PD). The curves of PD versus PFA are called Receiver Operating Characteristic (ROC) curves and express the detection performance. The more a ROC curve bends toward the upper left, the better is the detection performance since a higher PD and lower PFA is achieved.

Figure 7(a) shows the detection performance as a function of the dimension of the multichannel system, *i.e.*, $m = 2$, e.g., interferometric, $m = 3$, e.g., polarimetric and $m = 6$, e.g., polarimetric-interferometric system. For all cases, with fixed number of samples $n = 6$, one can realize that the performance of detection is significantly improved as the number of SAR images increases. Figure 7(b) shows the ROC curves for $m = \{2, 3\}$ and for different number of samples n. It can be seen that for both multidimensional system configurations the number of samples increases the detection performance.

Figure 7. Maximum sample eigenvalue detection performance as a function of the number or channels m, samples n and correlation ρ. (**a**) ROC curves for different multidimensional systems, $m = \{2, 3, 6\}$ and $n = 6$. (**b**) ROC curves for the two- and three-dimensional cases, and for different number of samples n. (**c**) ROC curves for the two-dimensional case, $n = 3$, and for different correlation.

(a) (b) (c)

In order to state how correlation effects the detection performance, the detection problem is reformulated as follows

$$H_0 \ (l_{max} \text{ is dominant}) : l_{max} = l_1 > l_2 \geq l_3, \ldots, \geq l_m \geq 0$$
$$H_1 \ (l_{max} \text{ is not dominant}) : l_{max} = l_1 = l_i \qquad i = \{2, 3, \ldots, m\}.$$

In this way, every time that correlation is greater than zero the eigenvalues have different values, and one is larger than the others. For the two-dimensional case, for instance, $p_{\lambda_{max}}(x; H_0)$ is the distribution of λ_{max} when $l_1 > l_2$, which changes for different values of the correlation ρ. Figure 7(c) presents the ROC curves for this case. For small correlation between channels, the detector suffers from a significant false-alarm rate. As expected, when correlation increases the ROC curve bends toward the upper left, indicating better detection performance. The limit is reached when $\rho = 1$, meaning that $l_1 = 2$ and $l_2 = 0$, which corresponds to the cases of the previous detection problem.

5.2. Target Detection Using Polarimetric Matched Filter

In this section the application of the expressions for the detection of specific targets using the Polarimetric Matched Filter (PMF) concept is illustrated. For that, a short review on PMF is required [6,7].

In a polarimetric acquisition, the dimension of the system corresponds to the number of channels, which are in general four, but three for the backscattering reciprocal case. The measured vector k is hence three- or four-dimensional. The elements of the vector k may be weighted in a given way, in order to satisfy a certain condition. The weighting can be accounted for by making $y = h^\dagger k$, where h is a complex vector with same dimension as k. A condition usually aimed to be satisfied in the literature is the maximization of the quadratic detector $|y|^2$ (Equation (59) in [6]). The optimal weighting vector h

in order to detect a distributed target with the target vector k is dependent on the target characteristics. Accordingly, the expected value of $|y|^2$ is given by

$$E\{|y|^2\} = E\{h^\dagger k (h^\dagger k)^\dagger\} = h^\dagger \Sigma_t h \tag{9}$$

where $\Sigma_t = E\{kk^\dagger\}$ is the target covariance matrix. Regarding Rayleigh quotient (Theorem 15.91 in [25]), for any m-dimensional complex vector x and a given $m \times m$ Hermitian matrix A, $x^\dagger A x \leq \|x\|^2 a_{max}$, where a_{max} is the maximum eigenvalue of A. The equality is valid if x is along the direction of the a_{max} eigenvector U_{max} ($\|U_{max}\| = 1$). Hence, for a given distributed target with covariance matrix Σ_t, the optimum weighting vector h is given by the eigenvector of the maximum eigenvalue of Σ_t.

The target vector k is a random sample and has alone no physical meaning. Therefore, usually multi-look processing is performed making, for n samples (or looks)

$$\overline{y}^2 = \sum_{i=1}^{n} |h^\dagger k_i|^2. \tag{10}$$

The target is assumed to be immersed in polarization independent clutter. In this way, the detection procedure is thus evaluated by choosing h as the eigenvector corresponding to the maximum eigenvalue of the covariance matrix Σ_t of the target to be detected, and applying a detection threshold $\overline{y}^2 > \mathcal{T}$.

In the presence of target $\overline{y}^2$ is maximum, given by $\overline{y}^2 = \|\overline{h}\|^2 \lambda_{max}$, where $\overline{h}$ is the eigenvector of λ_{max}. Hence, assuming without loss of generality that $\overline{h}$ has unitary length ($\|\overline{h}\|^2 = 1$), detection performances can be made using the derivations given in this work, as done in the previous section, where a threshold has been used in order to detect a dominant maximum sample eigenvalue ($\lambda_{max} = \overline{y}^2 > \mathcal{T}$).

Note that the probabilities of detection and false alarm defined in last section can be more simply evaluated by

$$PD = \int_{\mathcal{T}}^{\infty} p_{\lambda_{max}}(x; H_0) \, dx = 1 - \int_{-\infty}^{\mathcal{T}} p_{\lambda_{max}}(x; H_0) \, dx = 1 - F_{\lambda_{max}}(\mathcal{T}; H_0)$$
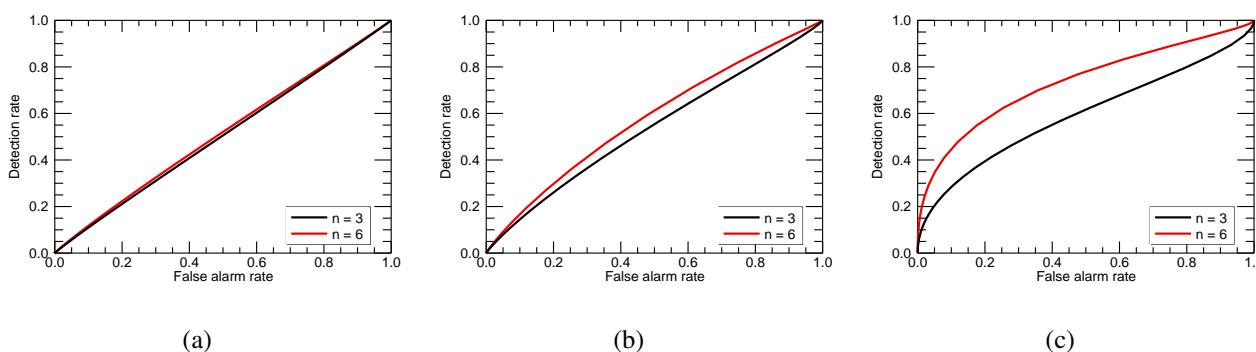$$PFA = 1 - F_{\lambda_{max}}(\mathcal{T}; H_1). \tag{11}$$

A three-dimensional polarimetric target detection problem was formulated making $k = [S_{HH} S_{HV} S_{VV}]^T$, where S_i is the polarimetric scattering matrix element in channel i, and using the target covariance matrix structure

$$\Sigma_t = S_{HH} \begin{bmatrix} 1 & 0 & \rho\sqrt{\gamma} \\ 0 & 2\epsilon & 0 \\ \rho^*\sqrt{\gamma} & 0 & \gamma \end{bmatrix}. \tag{12}$$

where $\epsilon = \frac{E\{|S_{HV}|^2\}}{E\{|S_{HH}|^2\}}$, $\gamma = \frac{E\{|S_{VV}|^2\}}{E\{|S_{HH}|^2\}}$ and ρ is the complex correlation coefficient between S_{HH} and S_{VV} [6,26]. Figure 8 shows the ROC curves for different number of samples for three different kinds of targets. Figure 8(a) corresponds to an azimuthal symmetric target having covariance matrix with parameters $\epsilon = 1$, $\gamma = 0.5$ and $\rho = 0$. Figure 8(c) corresponds to a reflection symmetric target having covariance matrix with parameters $\epsilon = 1$, $\gamma = 0.8$ and $\rho = 0.8$, while Figure 8(b) to a reflection symmetric target with parameters $\epsilon = 0$, $\gamma = 0.8$ and $\rho = 0.8$ in its covariance matrix. For all three cases, $S_{HH} = 1$ and the PFA was evaluated using the polarization independent clutter having $\epsilon = 1$, $\gamma = 1$ and $\rho = 0$. Note that the curves vary not just with the number of used samples n but are also different

for different targets having different Σ_t. This means that some types of targets are easier to detect than others, when using the quadratic detector described here and when the clutter is polarization independent. When the target covariance matrix is similar to the one of the clutter, the detection performance weakens (Figure 8(a)). On the other hand, when the covariance matrix of the target is significantly different from the clutter one, the detection performance improves (Figure 8(c)).

Figure 8. Detection performance of three different distributed scatterers using the PMF concept for different number of samples. (**a**) Azimuthal symmetric scatterer with $\epsilon = 1$, $\gamma = 0.5$ and $\rho = 0$. (**b**) Reflection symmetric scatterer with $\epsilon = 1$, $\gamma = 0.8$ and $\rho = 0.8$. (**c**) Reflection symmetric scatterer with $\epsilon = 0$, $\gamma = 0.8$ and $\rho = 0.8$. The PFA is evaluated using the polarization independent clutter with $\epsilon = 1$, $\gamma = 1$ and $\rho = 0$.

(a)   (b)   (c)

6. Conclusions and Discussion

In this paper a depth statistical analysis of the maximum eigenvalue of the eigendecomposition of the sample covariance matrix in terms of SAR applications is presented. The proposed analysis are supported by simulation results via several examples. The results are based on a exact closed-form expressions of PDF, CDF and MGF. In this study, existing density functions of the sample maximum eigenvalue were extended and/or implemented into multi-channel SAR system in order to obtain a simple expression of the sample eigenvalues giving a way to fruitful applications. From these closed-form expressions, it has been possible to develop new algorithms for unbiased calculations of parameters extracted from multi-channel SAR covariance matrix. In addition to these implementations, closed-form expressions were developed for the MGF of the sample maximum eigenvalue, which can be critical in the area of bias removal and detection performance analysis. This new closed-form expressions of the MGF can be also interesting for other application areas like MIMO systems (Multiple-Input Multiple-Output). Apart from estimation theory analysis including the MGF, the detection problem of the sample maximum eigenvalue has been also discussed.

References

1. Sikaneta, I.; Gierull, C.; Chouinard, J.Y. Metrics for SAR-GMTI Based on Eigen-Decomposition of the Sample Covariance Matrix. In *Proceedings of the 2003 International Radar Conference*, Adelaide, Australia, 3–5 September 2003; pp. 442-447.

2. Hajnsek, I.; Pottier, E.; Cloude, S.R. Inversion of surface parameters from polarimetric SAR. *IEEE Trans. Geosci. Remote Sens.* **2003**, *4*, 727-744.

3. Li, Z.; Bao, Z.; Li, H.; Liao, G. Image autocoregistration and InSAR interferogram estimation using joint subspace projection. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 288-297.

4. Erten, E.; Reigber, A.; Ferro-Famil, L.; Hellwich, O. A new coherent similarity measure for temporal multichannel scene characterization. *IEEE Trans. Geosci. Remote Sens.* **2011**, *50*, 1-13.

5. Zandoná-Schneider R.; Fernandes, D. Entropy Concept for Change Detection in Multitemporal SAR Images. In *Proceedings of EUSAR*, Köln, Germany, 4–6 June 2002.

6. Novak, L.M.; Sechtin, M.B.; Cardullo, M.J. Studies of target detection algorithms that use polarimetric radar data. *IEEE Trans. Aerosp. Electon. Syst.* **1989**, *25*, 150-165.

7. Liu, G.; Huang, S.; Torre, A.; Rubertone, F. The multilook polarimetric whitening filter (MPWF) for intensity speckle reduction in polarimetric SAR images. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1016-1020.

8. Li, Z.; Bao, Z.; Suo, Z.A. Joint image coregistration, phase noise suppression, and phase unwrapping method based on subspace projection for multibaseline InSAR systems. *IEEE Trans. Geosci. Remote Sens.* **2007**, *54*, 584-591.

9. Nadarajah, S. Comments on Eigendecomposition of the multi-channel covariance matrix with applications to SAR-GMTI. *Signal Process.* **2007**, *87*, 1534-1536.

10. Lee, J.S.; Grunes, M.R.; Schuler, D.L.; Pottier, E.; Ferro-Famil, L. Scattering-model-based speckle filtering of polarimetric SAR data. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 176-187.

11. Touzi, R. Target scattering decomposition in terms of roll-invariant target parameters. *IEEE Geosci. Remote Sens. Lett.* **2007**, *45*, 73-84.

12. Khatri, C.G. Non-central distributions of ith largest characteristic roots of three matrices concerning complex multivariate normal populations. *Ann. Inst. Stat. Math.* **1969**, *21*, 23-32.

13. Hirakawa, F. Some distributions of the latent roots of a complex Wishart matrix variate. *Ann. Inst. Stat. Math.* **1975**, *27*, 357-363.

14. Kang, M.; Alouini, M.S. Largest eigenvalue of complex Wishart matrices and performance analysis of MIMO MRC systems. *IEEE J. Sel. Areas Commun.* **2003**, *21*, 418-426.

15. McKay, M.R.; Grant, A.J.; Collings, I.B. Performance analysis of MIMO-MRC in double-correlated Rayleigh environments. *IEEE Trans. Commun.* **2007**, *55*, 497-507.

16. McKay, M.R. Random Matrix Theory Analysis of Multiple Antenna Communication Systems. Ph.D. Thesis, The University of Sydney, Sydney, Australia, 2006.

17. Lee, J.S.; Hoppel, K.W.; Mango, S.A.; Miller, A.R. Intensity and Phase Statistics of Multilook Polarimetric and Interferometric SAR Imagery. *IEEE Trans. Geosci. Remote Sens.* **1994**, *32*, 1117-1028.

18. Lopez-Martinez, C.; Pottier, E.; Cloude, S.R. Statistical assessment of eigenvector-based target decomposition theorems in radar polarimetry. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 2058-2074.

19. Muirhead, R.J. *Aspects of Multivariate Statistic Theory*; John Wiley&Sons: New York, NY, USA, 1982.

20. Ferro-Famil, L. Télédetection multi-fréquentielle et multi-temporelle d'environnements naturels à partir de données SAR polarimétriques. Ph.D. Thesis, Université de Nantes, Nantes, France, 2000.
21. Andrews, L.C. *Special Functions of Mathematics for Engineers*; McGraw-Hill Publishing Co.: New York, NY, USA, 1992.
22. Golberg, M.A. The derivative of a determinant. *Am. Math. Month.* **1972**, *78*, 1124-1126.
23. Ermolaev, V.T.; Rodyushkin, K.V. The distribution function of the maximum eigenvalue of a sample correlation matrix of internal noise of antenna-array elements. *Radiophys. Quantum El.* **1999**, *42*, 439-444.
24. Peebles, P.Z. *Probability, Random Variables, and Random Signal Principles*; McGraw-Hill College: 1993.
25. Gradshteyn, I.S.; Ryzhik, I.M. *Table of Integrals, Series, and Products*; Academic Press: Burlington, MA, USA, 2007.
26. Li, J.-S.; Pottier, E. *Polarimetric Radar Imaging from Basics to Applications*; CRS Press: Boca Raton, FL, USA, 2009.
27. Chiani, M.; Win, M.Z.; Zanella, A. On the capacity of spatially correlated MIMO Rayleigh-fading channels. *IEEE Trans. Inform. Theory* **2003**, *49*, 2363-2371.
28. Mestre, X.; Lagunas, M.A. Finite sample size effect on minimum variance beamformers: Optimum diagonal loading factor for large arrays. *IEEE Trans. Signal Process.* **2003**, *54*, 69-82.

Appendix

This Appendix starts with the reminder of some basic results from linear algebra and statistical theory to prove the Theorem III proposed in this manuscript.

Definition 1 (Theorem A3 in [19]): The determinant of a square $m \times m$ matrix A is

$$|A| = \sum_{\pi} \text{sgn}(\pi) a_{1k_1} a_{2k_2} ... a_{mk_m} = \sum_{\pi} \text{sgn}(\pi) \prod_{i=1}^{m} a_{ik_i} \qquad \text{(A1)}$$

where $\sum_{\pi}$ denotes the summation over all $m!$ permutations, and $\text{sgn}(\pi)$ denotes the sign of the permutations determined by $(-1)^{\pi}$, where π is the permutation symbol. Notice that the determinant is a combination of the determinants of its sub-matrices, and permuting the columns or rows of a matrix changes the sign of the determinant related to the permutation.

Definition 2 (Corollary 2 in [27]): Given two arbitrary $m \times m$ matrices $\Phi(x)$ and $\Psi(x)$ with ijth elements $\Phi_i(x_j)$ and $\Psi_i(x_j)$, and an arbitrary function $\xi(\cdot)$, the following identity holds:

$$\int \dots \int_{\mathcal{D}} |\Phi(x)| |\Psi(x)| \prod_{k=1}^{m} \xi(x_k) dx_1 dx_2 \dots dx_m = \left| \left\{ \int_a^b \Phi_i(x) \Psi_j(x) \xi(x) dx \right\} \right\|_{i,j=1}^{m} \qquad \text{(A2)}$$

where the multiple integral is over the domain $\mathcal{D} = \{a \leq x_m \leq \dots \leq x_2 \leq x_1 \leq b\}$.

Proof of Theorem I

In the case $m \leq n$, the sample covariance matrix has full rank and follows a central Wishart distribution. It can be noted that [28] is an interesting work concerning the case $n \leq m$. In Wishart

history, the joint PDF of the real ordered sample eigenvalues $\infty \geq \lambda_1 \geq \ldots \geq \lambda_m \geq 0$ of Z is given by Equation (36) in [18]

$$p_\Xi(\Xi) = \frac{\pi^{m(m-1)} n^{m(2n-m+1)/2}}{\tilde{\Gamma}_m(m)\tilde{\Gamma}_m(n)} \left\| \exp\left(-n\frac{\lambda_j}{l_i}\right) \right\|_{i,j=1}^m \prod_{k=1}^{m-1} k^{m-k} \frac{\prod_{i=1}^m \lambda_i^{n-m} \prod_{i<j}^m (\lambda_i - \lambda_j)}{\prod_{i=1}^m l_i^n \prod_{i<j}^m \left(\frac{1}{l_j} - \frac{1}{l_i}\right)}, \qquad (A3)$$

$$\text{with} \quad \Xi = Q^H \langle Z \rangle_n Q$$

where $\Xi = \text{diag}\{\lambda_1, \ldots, \lambda_m\}$, $p_\Xi(\Xi)$ is the joint PDF of the eigenvalues of the sample covariance matrix $\infty \geq \lambda_1 \geq \ldots \geq \lambda_{m-1} \geq \lambda_m \geq 0$ and $\infty \geq l_1 \geq \ldots \geq l_{m-1} \geq l_m \geq 0$ are the eigenvalues of the true covariance matrix.

The CDF of the maximum eigenvalue $\lambda_{max} = \lambda_1$ is obtained using

$$\begin{aligned} F_{\lambda_{max}}(x) &= p(\lambda_{max} \leq x) \\ &= p(\lambda_m \leq x, \ldots, \lambda_1 \leq x) \\ &= \int_\mathcal{D} p(\lambda_1, \lambda_2, \ldots, \lambda_m) \mathrm{d}\lambda_1 \ldots \mathrm{d}\lambda_m \end{aligned} \qquad (A4)$$

where $p(\lambda_1, \lambda_2, \ldots, \lambda_m)$ is the joint PDF of the eigenvalues as defined in Equation (A3), and $\mathcal{D} = \{0 \leq \lambda_m \leq \ldots \leq \lambda_1 \leq x\}$. Before substituting Equation (A3) in Equation (A4) it is desirable to write a friendlier expression for the joint PDF Equation (A3). If the constant is denoted by $\mathcal{S}$ as in Equation (3), then Equation (A3) can be written as

$$p_\Xi(\Xi) = \mathcal{S} \prod_{i=1}^m \lambda_i^{n-m} \prod_{i<j}^m (\lambda_i - \lambda_j) \left\| \exp\left(-n\frac{\lambda_j}{l_i}\right) \right\|_{i,j=1}^m. \qquad (A5)$$

Due to the definition of the Vandermonde determinant [25] the multiplication of $\prod_{i=1}^m \lambda_i^{n-m}$ and $\prod_{i<j}^m (\lambda_i - \lambda_j)$ is:

$$\begin{bmatrix} \lambda_1^{n-m} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_m^{n-m} \end{bmatrix} \cdot \begin{bmatrix} \lambda_1^{m-1} & \cdots & \lambda_1^{m-m} \\ \vdots & \ddots & \cdots \\ \lambda_m^{m-1} & \cdots & \lambda_m^{m-m} \end{bmatrix} = \begin{bmatrix} \lambda_1^{n-1} & \cdots & \lambda_1^{n-m} \\ \vdots & \ddots & \cdots \\ \lambda_m^{n-1} & \cdots & \lambda_m^{n-m} \end{bmatrix} \rightarrow \left| \lambda_i^{n-j} \right\|_{i,j=1}^m \qquad (A6)$$

Then the joint PDF of the sample eigenvalues can be written in the form of

$$p_\Xi(\Xi) = \mathcal{S} \left| \lambda_i^{n-j} \right\|_{i,j=1}^m \cdot \left| \exp\left(-n\frac{\lambda_j}{l_i}\right) \right\|_{i,j=1}^m \qquad (A7)$$

By applying the result of *Definition 2* and *Definition 1* into (A4) with $\Psi_i(\lambda_j) = \lambda_i^{n-j}$ and $\Phi_i(\lambda_j) = \exp\left(-n\frac{\lambda_j}{l_i}\right)$, it yields

$$F_{\lambda_{max}}(x) = \mathcal{S} \left| \int_0^x \lambda^{n-j} \exp(-n\frac{\lambda}{l_i}) \mathrm{d}\lambda \right|. \qquad (A8)$$

Finally, solving the remaining integral using Equation (351) in [25] yields the desired closed-form expression of the sample maximum eigenvalue CDF of multi-channel covariance matrix

$$F_{\lambda_{max}}(x) = \mathcal{S} \left| \Psi(x)_{i,j} \right| \qquad \text{with} \quad \Psi(x)_{i,j} = \left(\frac{n}{l_i}\right)^{(j-n-1)} \gamma(n+1-j, x\frac{n}{l_i}) \qquad (A9)$$

where $\gamma(a, x) = \int_0^x x^{a-1} e^{-x} \mathrm{d}x$ denotes the lower incomplete gamma function.

Proof of Moments of the Sample Maximum Eigenvalue

For any positive integer s, the sth moment of the maximum eigenvalue is obtained using

$$E(\lambda_{max}^s) = \int_0^\infty x^s p_{\lambda_{max}}(x)dx \qquad (A10)$$

where $p_{\lambda_{max}}(x)$ is the PDF of λ_{max}. It has been shown in Equation (4) that

$$p_{\lambda_{max}}(x) = \mathcal{S}|\Psi(x)|\text{tr}\left[\Psi(x)^{-1}\Omega(x)\right] \qquad (A11)$$

where $\Omega(x)$ and $\Psi(x)$ are $m \times m$ matrices with ijth element $\Omega(x)_{i,j} = \exp\left(-\frac{n}{l_i}x\right)x^{n-j}$ and $\Psi(x)_{i,j} = \left(\frac{n}{l_i}\right)^{-(n+1-j)}\gamma(n+1-j, x\frac{n}{l_i})$, and $\mathcal{S}$ denotes the constant terms given in Equation (3).

It is desirable to write a friendlier expression for Equation (A11) to be useful for deriving the moments of the sample maximum eigenvalue. Due to the definition of the inverse of the $m \times m$ matrix $A^{-1} = C_{ij}^T/|A|$, Equation (A11) can be written in the following form

$$p_{\lambda_{max}}(x) = \mathcal{S}\,\text{tr}\left[C^T(x)\Omega(x)\right] \qquad (A12)$$

where $C^T(x)$ is the transpose of the cofactor matrix $\Psi(x)$. Since the trace of a square matrix is the sum of the diagonal elements of the matrix, Equation (A11) can be written as the sum of the diagonal elements of $D(x) = (C^T(x)\Omega(x))$.

$$E(\lambda_{max}^s) = \mathcal{S}\int_0^\infty x^s \text{tr}\left[C^T(x)\Omega(x)\right] dx \qquad (A13)$$

$$= \mathcal{S}\int_0^\infty x^s \text{tr}\left[D\right] dx \qquad (A14)$$

Now the problem in Equation (A11) is the evaluation of the diagonal elements of the $D(x)$. The matrix $D(x)$ is the product of the two square matrices and the trace of the product of two matrices is $\text{tr}(C^T\Omega) = \sum_{i,j=1}^m C_{ji}\Omega_{ji} = C_{11}\Omega_{11} + \cdots + C_{m1}\Omega_{m1} + \cdots + C_{1m}\Omega_{1m} + \cdots + C_{mm}\Omega_{mm}$ shown as

$$tr\left(\begin{bmatrix} C_{11} & C_{21} & \cdots & C_{m1} \\ C_{12} & C_{22} & \cdots & C_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1m} & C_{2m} & \cdots & C_{mm} \end{bmatrix} \cdot \begin{bmatrix} e^{-x\frac{n}{l_1}}x^{n-1} & e^{-x\frac{n}{l_1}}x^{n-2} & \cdots & e^{-x\frac{n}{l_1}}x^{n-1} \\ e^{-x\frac{n}{l_2}}x^{n-1} & e^{-x\frac{n}{l_2}}x^{n-2} & \cdots & e^{-x\frac{n}{l_2}}x^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-x\frac{n}{l_m}}x^{n-1} & e^{-x\frac{n}{l_m}}x^{n-2} & \cdots & e^{-x\frac{n}{l_m}}x^{n-m} \end{bmatrix}\right) \qquad (A15)$$

where C_{ji} is the determinant of the sub-matrix M_{ji} obtained from $\Psi(x)$ by removing its jth row and ith column. Then, the j, i th cofactor of Ψ_{ji}, denoted by C_{ji}, is $C_{ji} = (-1)^{i+j}|M_{ji}|$. Now, Equation (A14) is:

$$E(\lambda_{max}^s) = \mathcal{S}\sum_{j,i=1}^m \int_0^\infty x^s C_{ji}\Omega_{ji}dx \qquad (A16)$$

where C_{ji} can be extracted from $\Psi(x)$ by *Definition I* as follows,

$$C_{ji} = \frac{(-1)^{i+j}\sum_{\pi_{sub}\in S_m}\text{sgn}(\pi_{sub})\prod_{k=1}^m \Psi_{k,\pi_{sub}(k)}}{\Psi_{j,i}}, \quad \text{if } k = j \wedge \pi_s(j) = i,\ \pi_s \in \pi_{sub}\ . \qquad (A17)$$

Here, the sum is computed over $(m-1)!$ permutations of π_s. S_m denotes the set of all $m!$ permutations of the set $S = \{1, 2, ..., m\}$, and $\mathrm{sgn}(\pi_{sub})$ denotes the sign of the permutation π_s : $+1$ if π_s is an even permutation and -1 if it is odd. After interchanging the order of the summations and integral in Equation (A16), the integration reduces to the following form

$$E(\lambda_{max}^s) = \mathcal{S} \sum_{i,j=1}^{m} (-1)^{i+j} \int_0^\infty x^s \exp\left(-\frac{n}{l_j}x\right) x^{n-i}$$

$$\times \left(\frac{\displaystyle\sum_{\pi_{sub} \in S_m} \mathrm{sgn}(\pi_{sub}) \prod_{k=1}^{m} \gamma(n+1-\pi_{sub}(k), x\frac{n}{l_k})(\frac{n}{l_k})^{\pi_{sub}(k)-n-1}}{\gamma(n+1-i, x\frac{n}{l_j})(\frac{n}{l_j})^{i-1-n}} \right) \mathrm{d}x$$

$$\Rightarrow \mathcal{S} \sum_{i,j=1}^{m} (-1)^{i+j} \sum_{\pi_{sub} \in S_m} \mathrm{sgn}(\pi_{sub}) \int_0^\infty x^s \exp\left(-\frac{n}{l_j}x\right) x^{n-i}$$

$$\times \left(\frac{\prod_{k=1}^{m} \gamma(n+1-\pi_{sub}(k), x\frac{n}{l_k})(\frac{n}{l_k})^{\pi_{sub}(k)-n-1}}{\gamma(n+1-i, x\frac{n}{l_j})(\frac{n}{l_j})^{i-1-n}} \right) \mathrm{d}x. \tag{A18}$$

It can be seen from Equation (A18) that the integration includes the multiplication of $(m-1)$ times incomplete gamma functions with exponential and powers. After applying identities (ex. 10.3, Equations 10.15 and 10.65 in [21]),

$$\gamma(n, ax) = \frac{1}{n}(ax)^{\frac{n-1}{2}} e^{\frac{-ax}{2}} \mathcal{M}_{\frac{n-1}{2}, \frac{n}{2}}(ax), \quad \frac{n}{2} \neq -1, -2, -3, ..., \tag{A19}$$

$$\gamma\left(n+1-\pi_{sub}(k), \frac{n}{l_k}x\right) = \frac{1}{n+1-\pi_{sub}(k)} \left(\frac{n}{l_k}x\right)^{\frac{n-\pi_{sub}(k)}{2}} e^{\frac{-nx}{2l_k}} \mathcal{M}_{\frac{n-\pi_{sub}(k)}{2}, \frac{n+1-\pi_{sub}(k)}{2}}\left(\frac{n}{l_k}x\right), \tag{A20}$$

a solution of integration in Equation (A18) may be obtained by using the following formula

$$\int_0^\infty x^{v-1} e^{-bx} \mathcal{M}_{\lambda_1, u_1 - \frac{1}{2}}(xa_1) \ldots \mathcal{M}_{\lambda_m, u_m - \frac{1}{2}}(xa_m) \mathrm{d}x = a_1^{u_1} \ldots a_m^{u_m} (b+A)^{-v-M} \Gamma(v+M) \times$$

$$F_A\left(v+M; u_1 - \lambda_1, ..., u_m - \lambda_m; 2u_1, ..., 2u_m; \frac{a_1}{b_1+A}, ..., \frac{a_m}{b_m+A}\right), \tag{A21}$$

where $M = u_1 + \cdots + u_m$ and $A = (1/2)(a_1 + \cdots + a_m)$ are valid for $\Re(M+v) > 0$, and $\Re(b \pm \frac{1}{2}a_1 \pm \cdots \pm \frac{1}{2}a_n) > 0$. Here $\mathcal{M}_{k,m}(x)$ is a Whittaker function of the first kind Equation (7.622) in [25] and $F_A(\alpha : \beta_1, \ldots, \beta_m; \gamma_1, \ldots, \gamma_m : z_1, \ldots, z_m)$ is a hypergeometric function of several variables Equation (9.19) in [25]. After replacing the integration part of Equation (A18) by Equation (A21), it can easily be shown that for m dimensional SAR system with $n \geq m$, the moments of the sample maximum eigenvalue is

$$E(\lambda_{max}^s) = \mathcal{S} \sum_{i,j=1}^{m} (-1)^{i+j} \sum_{\substack{\pi_{sub} \in S_m \\ k \neq i \wedge \pi_{sub}(k) \neq j}} \mathrm{sgn}(\pi_{sub})$$

$$\int_0^\infty \prod_{k=1}^{m-1} \frac{1}{n+1-\pi_{sub}(k)} \left(\frac{n}{l_k}\right)^{\frac{n-\pi_{sub}(k)}{2}} x^{s+n-i+\frac{n-\pi_{sub}(k)}{2}} \exp\left(-x\left(\frac{n}{l_j}+\frac{n}{2l_k}\right)\right) \mathcal{M}_{\frac{n-\pi_{sub}(k)}{2}, \frac{n+1-\pi_{sub}(k)}{2}}\left(x\frac{n}{l_k}\right) \mathrm{d}x$$

$$E(\lambda_{max}^s) = \mathcal{S} \sum_{i,j=1}^{m} (-1)^{i+j} \sum_{\pi_{sub} \in S_m} \text{sgn}(\pi_{sub}) \prod_{k=1}^{m-1} \frac{1}{n+1-\pi_{sub}(k)} \left(\frac{n}{l_k}\right)^{\frac{n-\pi_{sub}(k)}{2}} a_k^{u_k} \times$$

$$(b+A)^{-v-M} \Gamma(v+M) \times F_A\left(v+M; u_1 - \lambda_1, \ldots, u_{m-1} - \lambda_{m-1}; 2u_1, \ldots, 2u_{m-1}; \frac{a_1}{b_1+A},\right.$$

$$\left. \cdots, \frac{a_{m-1}}{b_{m-1}+A} \right)$$

(A22)

where $a_k = \frac{n}{l_k}$, $u_k = (2 + n - \pi_{sub}(k))/2$, $v = 1 + s + n - i + \sum_{k=1}^{m-1} \frac{n-\pi_{sub}(k)}{2}$, $b = \frac{n}{l_j} + \sum_{k=1}^{m-1} \frac{n}{2l_k}$, $\lambda_k = \frac{n-\pi_{sub}(k)}{2}$, $M = \sum_{k=1}^{m-1} u_k$ and $A = \frac{1}{2} \sum_{k=1}^{m-1} a_k$. Notice that k and $\pi_{sub}(k)$ takes values from $(k, \pi_{sub}(k) = \{k, \pi_{sub}(k) \in \{1, 2, \ldots, m\}, \ k \neq j \ \wedge \ \pi_{sub}(j) \neq i\})$ related to given i, j.

Although Equation (A22) is an exact closed-form expression for the MGF of the sample maximum eigenvalue, the hypergeometric function of several variables is in general very difficult to evaluate, and it is better to express Equation (A22) in a more tractable form for future analysis. Regarding Equation (2.53) in [21], the incomplete gamma function can be written in the closed form

$$\gamma\left(n+1-\pi_{sub}(k), x\frac{n}{l_k}\right) = (n-\pi_{sub}(k))! \left(1 - \exp\left(-x\frac{n}{l_k}\right) \left(\sum_{t=0}^{n-\pi_{sub}(k)} \frac{\left(x\frac{n}{l_k}\right)^t}{t!}\right)\right)$$

(A23)

for $n \in \mathbb{Z}^+$, which is a constraint fulfilled by our problem. For the sake of simplicity, $f(k, \pi_{sub(k)}) = \exp\left(-x\frac{n}{l_k}\right) \left(\sum_{t=0}^{n-\pi_{sub}(k)} \frac{\left(x\frac{n}{l_k}\right)^t}{t!}\right)$, and $\prod_{k=1}^{m} (n - \pi_{sub}(k))! \left(1 - f(k, \pi_{sub(k)})\right) = \prod_{k=1}^{m} a_k! \left(1 - f_k\right)$ indicate m times multiplications of incomplete gamma function. Then if this multiplication is extended as in the form

$$\prod_{k=1}^{m} a_k! \left(1 + \sum_{k_1=1}^{\binom{m}{1}} (-1)^1 f_{k_1} + \sum_{k_2=2>k_1=1}^{\binom{m}{2}} (-1)^2 f_{k_1} f_{k_2} + \cdots + \sum_{k_m=m>\cdots>k_2=2>k_1=1}^{\binom{m}{m}} (-1)^m f_{k_1} f_{k_2} \cdots f_{k_m}\right)$$

(A24)

and the integration can be easily implemented into Equation (A18) as follows

$$E(\lambda_{max}^s) = \mathcal{S} \sum_{i,j=1}^{m} (-1)^{i+j} \sum_{\pi_{sub} \in S_m} \text{sgn}(\pi_{sub}) \int_0^{\infty} x^{s+n-i} \exp\left(-\frac{n}{l_j}x\right) \left(\prod_{k=1}^{m} (n-\pi_{sub}(k))! \right.$$

$$\left. \left(1 - f(k, \pi_{sub(k)})\right)\right) / \Psi_{ji} \mathrm{d}x$$

(A25)

By analyzing that the term Ψ_{ji} in the determinant of Equation (A25) reduces to $(m-1)$ multiplication of incomplete gamma functions:

$$E(\lambda_{max}^s) = \mathcal{S} \sum_{i,j=1}^{m} (-1)^{i+j} \sum_{\pi_{sub} \in S_m} \text{sgn}(\pi_{sub}) \prod_{k=1}^{m} a_k! \left(\int_0^\infty x^s \exp\left(-\frac{n}{l_j}x\right) x^{n-i} \mathrm{d}x + \sum_{k_1=1}^{\binom{m-1}{1}} \int_0^\infty f_{k_1} x^s \right.$$

$$\left. \exp\left(-\frac{n}{l_j}x\right) x^{n-i} \mathrm{d}x + \cdots + \sum_{k_{m-1}=m-1>\cdots>k_2=2>k_1=1}^{\binom{m-1}{m-1}} \int_0^\infty f_{k_1} f_{k_2} \cdots f_{k_{m-1}} x^s \exp\left(-\frac{n}{l_j}x\right) x^{n-i} \mathrm{d}x \right).$$

(A26)

Applying the rule $\int_0^\infty x^n e^{-ux} \mathrm{d}x = n! u^{-n-1}$ in Equation (A26) it yields:

$$E(\lambda_{max}^s) = \mathcal{S} \sum_{i,j=1}^{m} (-1)^{i+j} \sum_{\pi_{sub} \in S_m} \text{sgn}(\pi_{sub}) \prod_{k=1}^{m-1} a_k!$$

$$\left(\frac{(n-i+s)!}{\left(\frac{n}{l_j}\right)^{n-i+s+1}} + \sum_{k_1=1}^{\binom{m-1}{1}} \sum_{t_{k_1}=0}^{n-\pi_{sub}(k_1)} \frac{\left(\frac{n}{l_{k_1}}\right)^{t_{k_1}}}{t_{k_1}!} \frac{(s+n-i+t_{k_1})!}{\left(\frac{n}{l_j}+\frac{n}{l_{k_1}}\right)^{s+n-i+t_{k_1}+1}} + \cdots + \right.$$

$$\left. \sum_{t_{k_{m-1}}=0}^{n-\pi_{sub}(k_{m-1})} \frac{\left(\frac{n}{l_{k_{m-1}}}\right)^{t_{k_{m-1}}}}{t_{k_{m-1}}!} \cdots \sum_{t_{k_1}=0}^{n-\pi_{sub}(k_1)} \frac{\left(\frac{n}{l_{k_1}}\right)^{t_{k_1}}}{t_{k_1}!} \frac{(s+n-i+t_{k_1}+\cdots+t_{k_{m-1}})!}{\left(\frac{n}{l_j}+\frac{n}{l_{k_1}}+\cdots+\frac{n}{l_{k_{m-1}}}\right)^{s+n-i+t_{k_1}+\cdots+t_{k_{m-1}}+1}} \right).$$

(A27)

The sth moment of sample maximum eigenvalue can finally be written in a closed-form expression as

$$E(\lambda_{max}^s) = \mathcal{S} \sum_{i,j=1}^{m} \sum_{\pi_{sub} \in S_m} \text{sgn}(\pi_{sub}) \mathcal{G}(k, \pi_{sub}(k)), \ \ \text{if } k=j \ \wedge \ \pi_k(j)=i, \ \pi_k \in \pi_{sub} \quad (A28)$$

$$\mathcal{G}(k, \pi_{sub}(k)) = \prod_{k=1}^{m-1} (n-\pi_{sub}(k))! \times \left(\frac{(n-i+s)!}{\left(\frac{n}{l_j}\right)^{n-i+s+1}} \right. \tag{A29}$$

$$+ \sum_{k_1=1}^{\binom{m-1}{1}} \sum_{t_{k_1}=0}^{n-\pi_{sub}(k_1)} \frac{\left(\frac{n}{l_{k_1}}\right)^{t_{k_1}}}{t_{k_1}!} \frac{(s+n-i+t_{k_1})!}{\left(\frac{n}{l_j}+\frac{n}{l_{k_1}}\right)^{s+n-i+t_{k_1}+1}}$$

$$+ \sum_{k_2=2>k_1=1}^{\binom{m-1}{2}} \sum_{t_{k_1}=0}^{n-\pi_{sub}(k_1)} \frac{\left(\frac{n}{l_{k_1}}\right)^{t_{k_1}}}{t_{k_1}!} \sum_{t_{k_2}=0}^{n-\pi_{sub}(k_2)} \frac{\left(\frac{n}{l_{k_2}}\right)^{t_{k_2}}}{t_{k_2}!} \frac{(s+n-i+t_{k_1}+t_{k_2})!}{\left(\frac{n}{l_j}+\frac{n}{l_{k_1}}+\frac{n}{l_{k_2}}\right)^{s+n-i+t_{k_1}+t_{k_2}+1}}$$

$$+ \cdots$$

$$+ \sum_{t_{k_{m-1}}=0}^{n-\pi_{sub}(k_{m-1})} \frac{\left(\frac{n}{l_{k_{m-1}}}\right)^{t_{k_{m-1}}}}{t_{k_{m-1}}!} \cdots \sum_{t_{k_1}=0}^{n-\pi_{sub}(k_1)} \frac{\left(\frac{n}{l_{k_1}}\right)^{t_{k_1}}}{t_{k_1}!}$$

$$\left. \frac{(s+n-i+t_{k_1}+\cdots+t_{k_{m-1}})!}{\left(\frac{n}{l_j}+\frac{n}{l_{k_1}}+\cdots+\frac{n}{l_{k_{m-1}}}\right)^{s+n-i+t_{k_1}+\cdots+t_{k_{m-1}}+1}} \right)$$

(A30)

The moments of the sample maximum eigenvalue regarding its PDF is solved. For larger m dimensional systems, such an approach may become complicated due to the need for a large number of calculation of cofactors. However, for applications where only a few eigenvalues, like polarimetry ($m = 3$) and interferometry ($m = 2$), are of interest, the numerical calculations are quite rapid and stable. For example, in the simplest case of $m = 2$, $n \geq 2$, the MGF is given by

$$
E(\lambda_{max}^s) = \mathcal{S} \left(\sum_{j,i=1}^{2} \int_0^\infty x^s x^{n-i} \exp(-x\frac{n}{l_j}) \sum_{\pi_{sub} \in S_m} \mathrm{sgn}(\pi_{sub}) \gamma(n+1-k, x\frac{n}{l_{\pi_{sub}(k)}}) (\frac{n}{l_{\pi_{sub}(k)}})^{k-n-1} \right)
$$

$$
\text{if } \pi_k(j) = i, \pi_k \in \pi_{sub}, \text{ and if } k \neq j \text{ and } l_{\pi_{sub}(k)} \neq i, \ k = 1, 2
$$

$$
= \mathcal{S}\Gamma(2n)\Gamma(n-1) \left(\frac{l_1 l_2}{nl_1 + nl_2} \right)^{2n}
$$

$$
\times \left(\sum_{j,i=1}^{2} \sum_{\pi_{sub} \in S_m} \mathrm{sgn}(\pi_{sub})(n+1-k) {}_2F_1(1, 2n; n+2-k; \frac{l_i}{l_i + l_{\pi_{sub}(k)}}) \right).
$$

(A31)

where the constant $\mathcal{S} = \frac{(l_1 l_2)^{1-n} n^{2n}}{\Gamma(n-1)\Gamma(n+1)(l_1-l_2)}$ for $m = 2$ and ${}_2F_1(a, b; c; x)$ is a Hypergeometric function Equation (9.8) in [21].