Complex singular Wishart matrices and applications*

T. Ratnarajah[†‡] R. Vaillancourt[†§]

CRM-3185

May 2005

*Thanks are due to the Natural Sciences and Engineering Council of Canada and the Centre de recherches mathématiques of the Université de Montréal for their support.

†Department of Mathematics and Statistics, University of Ottawa, 585 King Edward Ave., Ottawa ON K1N 6N5 Canada

‡tharm@mathstat.uottawa.ca

§remi@uottawa.ca

Abstract

In this paper, complex singular Wishart matrices and their applications are investigated. In particular, a volume element on the space of positive semidefinite $m \times m$ complex matrices of rank $n < m$ is introduced and some transformation properties are established. The Jacobian for the change of variables in the singular value decomposition of general $m \times n$ complex matrices is derived. Then the density functions are formulated for all rank n complex singular Wishart distributions. From this, the joint eigenvalue density of low rank complex Wishart matrices are derived. Finally, application of these densities in information theory is given.

Keywords and Phrases. Complex singular matrix distribution, complex Wishart distribution, singular value decomposition, channel capacity, Rayleigh distributed MIMO channel.

AMS(MOS) subject classification. 62H10, 60E05, 94A15, 94A05.

Résumé

On étudie les matrices de Wishart complexes singulières. On introduit un élément volumétrique sur l'espace des matrices $m \times m$ complexes semi-définies de rang $n < m$ et l'on en établit certaines propriétés, telles la décomposition selon les valeurs singulières et leurs fonctions de densité, d'où l'on obtient la densité conjointe des valeurs propres de ces matrices de bas rang. On appliques ces densités à la théorie de l'information.

1 Introduction

The present communictions revolution is due, in part, by advances in wireless communications, such as wireless internet and multimedia communications. As the wireless industry becomes ubiquitous and polular, the need for a high-data rate and large user capacity on a wireless platform is the key driver in developing robust communications techniques, which offer a substantially increased information capacity. Recently, the theory of complex random matrices has attracted a lot of attention as a powerful mathematical tool that enables the computation of MIMO capacities for correlated channels. Multiple input, multiple output (MIMO) wireless communication systems, where the number of inputs (or transmitters) n_t is greater than the number of outputs (or receivers) n_r $(n_t > n_r)$ and the channel is correlated at the transmitter end.

The capacity of a communication channel expresses the maximum rate at which information can be reliably conveyed by the channel [1]. A MIMO channel can be represented by an $n_r \times n_t$ complex random matrix $\mathrm{H} \sim \mathcal{C}N(0, \Sigma_r \otimes \Sigma_t)$, where Σ_t and Σ_r represent the channel correlation at the transmitter and receiver ends, respectively. If $\Sigma_r = \sigma^2 I_{n_r}$ (or I_{n_r}) and $\Sigma_t = I_{n_t}$ (or $\sigma^2 I_{n_t}$) then the channel is said to be an *uncorrelated Rayleigh distributed channel*, and the complex nonsingular Wishart matrix theory can be used to compute the capacities for both cases $n_r \geq n_t$ and $n_t > n_r$, see [2]. However, if $\Sigma_r = I_{n_r}$ and $\Sigma_t = \Sigma$ then the complex nonsingular and singular Wishart matrix theories are needed to compute the capacities for the cases $n_r \geq n_t$ and $n_t > n_r$, respectively. See [3] for the case $n_r \geq n_t$. In this paper we assume that $\Sigma_r = I_{n_r}$, $\Sigma_t = \Sigma$ and $n_t > n_r$, which leads us to represent the channel capacity in the form of a complex singular Wishart matrix. This is the motivation behind this study.

Let an $n \times m$ complex Gaussian (or normal) random matrix A be distributed as $\mathrm{A} \sim \mathcal{C}N(M, I_n \otimes \Sigma)$ with mean $\mathcal{E}\{\mathrm{A}\} = M$ and covariance $\mathrm{cov}\{\mathrm{A}\} = I_n \otimes \Sigma$. Here we read the symbol "$\sim$" as "is distributed as", $\mathcal{C}N$ denotes the complex normal distribution and $\otimes$ denotes the Kronecker product. Then the matrix $\mathrm{W} = \mathrm{A}^H\mathrm{A}$ is called a complex noncentral Wishart matrix. If $M = 0$, then W is called a complex central Wishart matrix. The complex central and noncentral Wishart distributions are denoted by $\mathcal{C}W_m(n, \Sigma)$ and $\mathcal{C}W_m(n, \Sigma, \Omega)$, respectively, where $\Omega = \Sigma^{-1}M^HM$. The complex Wishart matrices are well studied in the literature only for $n \geq m$, for example, see [4], [5] and [6].

In this paper, we extend the study of complex central Wishart distributions to the singular case, where $0 < n < m$ and $n, m \in \mathbb{Z}$. Thus the rank of $\mathrm{W} \in \mathbb{C}^{m\times m}$ is n provided the rank of $A \in \mathbb{C}^{n\times m}$ is n. A volume element on the space of positive semidefinite $m \times m$ Hermitian matrices of rank $n < m$ is introduced (see Theorem 1). The Jacobian of the change of variables in the singular value decomposition of general $m \times n$ complex matrices is derived (see Theorem 2). The density is derived for rank-n complex central Wishart distributions for all integers n, $0 < n < m$ (see Theorem 3). The joint eigenvalue density of low rank complex Wishart matrices are derived (see Theorem 4). It should be noted that singular Wishart and beta distributions are studied in [7] for real random matrices. The singular value distribution of Gaussian random matrices is given in [8].

This paper is organized as follows. Section 2 provides the necessary tools for deriving the complex singular Wishart distribution theory. Complex singular Wishart matrices are studied in Section 3. The capacity of a MIMO channel and the computational method are given in Section 4.

2 Necessary tools

In this section, we derive necessary tools for studying the singular Wishart distribution theory and MIMO channel capacity. If $0 < n < m$, then the density does not exist for $\mathrm{W} \sim \mathcal{C}W_m(n, \Sigma)$ on the space of Hermitian $m \times m$ matrices, because W is singular and of rank n almost surely. It can be shown that the density does exist on the $(2mn - n^2)$-dimensional manifold, $\mathcal{C}S_{m,n}$, of rank n of positive semidefinite $m \times m$ Hermitian matrices W with n distinct positive eigenvalues. Moreover, the set of all $m \times n$ matrices E_1 with orthonormal columns is called the *Stiefel manifold*, denoted by $\mathcal{C}V_{n,m}$. Thus,

$$\mathcal{C}V_{n,m} = \{E_1(m \times n); E_1^HE_1 = I_n\}. \tag{1}$$

The elements of E_1 can be regarded as the coordinates of a point on a $(2mn - n^2)$-dimensional surface in the $2mn$-dimensional Euclidean space.

Theorems 1 and 2 below are derived by means of the exterior product approach. See [9], p. 57, for the definition of the exterior product for real matrices, such as symmetric, skew-symmetric, upper-triangular and arbitrary matrices. On the other hand, if $X = X_r + iX_c$ is a complex matrix, then the exterior product is $(dX) = (dX_r) \wedge (dX_c)$. Jacobian formulas for some important complex matrix factorizations are given in [10].

The volume element (dW) in the reduced spectral decomposition $W = E_1\Lambda E_1^H$ is given by the following theorem.

Theorem 1 *Let m and n be two positive integers such that $0 < n < m$ and consider an $m \times m$ positive semidefinite Hermitian matrix $W \in \mathcal{C}S_{m,n}$ of rank n with decomposition $W = E_1\Lambda E_1^H$, where the diagonal elements of $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ are positive real eigenvalues in decreasing order, $\lambda_1 > \cdots > \lambda_n > 0$, and $E_1 \in \mathcal{C}V_{n,m}$. Then the* *volume element is*

$$(dW) = (2\pi)^{-n}\left(\prod_{k=1}^{n}\lambda_k^{2m-2n}\right)\prod_{k<l}^{n}(\lambda_k-\lambda_l)^2\,(d\Lambda)\wedge(E_1^H\,dE_1), \tag{2}$$

where

$$(d\Lambda) = \bigwedge_{k=l}^{n} d\lambda_k, \qquad (E_1^H\,dE_1) = \bigwedge_{k=1}^{n}\bigwedge_{l=k}^{m} e_l^H\,de_k,$$

and the matrix E_1 is appended with an $m\times(m-n)$ matrix E_2 such that the compound $m\times m$ matrix, $E=[E_1 : E_2]=[e_1,\ldots,e_n : e_{n+1},\ldots,e_m]$ is unitary.

Proof. First, we note that

$$dW = dE_1\Lambda E_1^H + E_1\,d\Lambda E_1^H + E_1\Lambda\,dE_1^H,$$

and

$$E_1^H E_1 = I_n, \qquad E_2^H E_1 = 0.$$

Therefore, we have

$$E^H\,dW E = \begin{bmatrix} E_1^H\,dE_1\Lambda + d\Lambda + \Lambda\,dE_1^H E_1 & \Lambda\,dE_1^H E_2 \\ E_2^H\,dE_1\Lambda & 0\end{bmatrix}. \tag{3}$$

The exterior product on the left side of equation (3) is equal to

$$(E^H dWE) = (\det E)^{2m}(dW) = (dW).$$

The lth row of $E_2^H\,dE_1\Lambda$ is

$$\begin{bmatrix} e_l^H\,de_1 & \cdots & e_l^H\,de_n\end{bmatrix}\Lambda, \qquad n+1\le l\le m,$$

and the exterior product of its elements is equal to

$$(\det\Lambda)^2\bigwedge_{k=1}^{n} e_l^H\,de_k = \left(\prod_{k=1}^{n}\lambda_k^2\right)\bigwedge_{k=1}^{n} e_l^H\,de_k.$$

Therefore, the exterior product of all the elements of $E_2^H\,dE_1\Lambda$ is

$$\bigwedge_{l=n+1}^{m}\left[\left(\prod_{k=1}^{n}\lambda_k^2\right)\bigwedge_{k=1}^{n} e_l^H\,de_k\right] = \left(\prod_{k=1}^{n}\lambda_k^{2m-2n}\right)\bigwedge_{k=1}^{n}\bigwedge_{l=n+1}^{m} e_l^H\,de_k. \tag{4}$$

Second, we consider $E_1^H\,dE_1\Lambda + d\Lambda + \Lambda dE_1^H E_1$. Since the columns of E_1 are orthonormal, we have

$$E_1^H E_1 = I_n \Longrightarrow E_1^H\,dE_1 = -dE_1^H E_1 = -(E_1^H\,dE_1)^H.$$

This implies that the real and the imaginary parts of $E_1^H\,dE_1$ are skew-symmetric and symmetric, respectively. Moreover, we have

$$E_1^H\,dE_1\Lambda + d\Lambda + \Lambda dE_1^H E_1 = E_1^H\,dE_1\Lambda - \Lambda E_1^H\,dE_1 + d\Lambda. \tag{5}$$

The exterior product of the diagonal elements of the right side of equation (5) is given by

$$(d\Lambda) = \bigwedge_{k=1}^{n} d\lambda_k. \tag{6}$$

Note that the diagonal elements of $E_1^H\,dE_1\Lambda - \Lambda E_1^H\,dE_1$ are zeros.

Now, we consider the upper diagonal elements of $E_1^H\,dE_1\Lambda - \Lambda E_1^H\,dE_1$. Since the matrix $\operatorname{Re}\left(E_1^H\,dE_1\right)$ is skew-symmetric, it can be written as

$$\operatorname{Re}\left(E_1^H\,dE_1\right) = \begin{bmatrix} 0 & -\operatorname{Re}\left(e_2^H\,de_1\right) & \cdots & \cdots & -\operatorname{Re}\left(e_n^H\,de_1\right) \\ \operatorname{Re}\left(e_2^H\,de_1\right) & 0 & \cdots & \cdots & -\operatorname{Re}\left(e_n^H\,de_2\right) \\ \operatorname{Re}\left(e_3^H\,de_1\right) & \operatorname{Re}\left(e_3^H\,de_2\right) & 0 & \cdots & -\operatorname{Re}\left(e_n^H\,de_3\right) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \operatorname{Re}\left(e_n^H\,de_1\right) & \operatorname{Re}\left(e_n^H\,de_2\right) & \ldots & \ldots & 0\end{bmatrix}.$$

For $k<l$, the (k,l)th element of $\operatorname{Re}\left(E_1^H\,dE_1\right)\Lambda - \Lambda\operatorname{Re}\left(E_1^H\,dE_1\right)$ is given by $\operatorname{Re}\left(e_l^H\,de_k\right)(\lambda_k-\lambda_l)$. Therefore, the exterior product of the upper diagonal elements of

$$\operatorname{Re}\left(E_1^H\,dE_1\right)\Lambda - \Lambda\operatorname{Re}\left(E_1^H\,dE_1\right)$$

is given by

$$\bigwedge_{k<l}^{n} \mathrm{Re}(e_l^H\, de_k) \prod_{k<l}^{n} (\lambda_k - \lambda_l). \tag{7}$$

Similarly, since the matrix $\mathrm{Im}\left(E_1^H\, dE_1\right)$ is symmetric, it can be written as

$$\mathrm{Im}\left(E_1^H\, dE_1\right) = \begin{bmatrix} \mathrm{Im}\left(e_1^H\, de_1\right) & -\mathrm{Im}\left(e_2^H\, de_1\right) & \dots & \dots & -\mathrm{Im}\left(e_n^H\, de_1\right) \\ -\mathrm{Im}\left(e_2^H\, de_1\right) & \mathrm{Im}\left(e_2^H\, de_2\right) & \dots & \dots & -\mathrm{Im}\left(e_n^H\, de_2\right) \\ -\mathrm{Im}\left(e_3^H\, de_1\right) & -\mathrm{Im}\left(e_3^H\, de_2\right) & \dots & \dots & -\mathrm{Im}\left(e_n^H\, de_3\right) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\mathrm{Im}\left(e_n^H\, de_1\right) & -\mathrm{Im}\left(e_n^H\, de_2\right) & \dots & \dots & \mathrm{Im}\left(e_n^H\, de_n\right) \end{bmatrix}$$

and the exterior product of the upper diagonal elements of

$$\mathrm{Im}\left(E_1^H\, dE_1\right)\Lambda - \Lambda\, \mathrm{Im}\left(E_1^H\, dE_1\right)$$

is given by

$$\bigwedge_{k<l}^{n} \mathrm{Im}(e_l^H\, de_k) \prod_{k<l}^{n} (\lambda_k - \lambda_l). \tag{8}$$

Therefore, the exterior product of the elements of the right side of equation (3) is obtained by multiplying equations (4), (6), (7), and (8), i.e.,

$$(dW) = (2\pi)^{-n} \left(\prod_{k=1}^{n} \lambda_k^{2m-2n} \right) \prod_{k<l}^{n} (\lambda_k - \lambda_l)^2 (d\Lambda) \wedge (E_1^H\, dE_1). \tag{9}$$

Note that we must divide the volume element by $(2\pi)^n$ to normalize the arbitrary phases of the n elements in the first row of E_1. □

The volume element (dW) (or Jacobian) in the singular value decomposition is given by the following theorem.

Theorem 2 *Let Z be an $m \times n$ complex matrix and $Z = E_1 \Upsilon H$ the nonsingular part of the singular value decomposition, where $E_1 \in \mathcal{C}V_{n,m}$, $H \in U(n)$ and the diagonal elements of $\Upsilon = \mathrm{diag}(\upsilon_1, \dots, \upsilon_n)$ are positive real singular values with $\upsilon_1 > \cdots > \upsilon_n > 0$. Then we have*

$$\begin{aligned} (dZ) &= (2\pi)^{-n} \left(\prod_{k=1}^{n} \upsilon_k^{2m-2n+1} \right) \\ &\quad \times \prod_{k<l}^{n} (\upsilon_k^2 - \upsilon_l^2)^2 (d\Upsilon) \wedge (E_1^H\, dE_1) \wedge (H^H dH) \end{aligned} \tag{10}$$

where

$$(d\Upsilon) = \bigwedge_{k=l}^{n} d\upsilon_k, \quad (H^H dH) = \bigwedge_{k=1}^{n} \bigwedge_{l=k}^{n} h_l\, dh_k, \quad (E_1^H\, dE_1) = \bigwedge_{k=1}^{n} \bigwedge_{l=k}^{m} e_l^H\, de_k,$$

and the matrix E_1 is appended with an $m \times (m-n)$ matrix E_2 such that the compound $m \times m$ matrix, $E = [E_1 : E_2] = [e_1, \dots, e_n : e_{n+1}, \dots, e_m]$ is unitary.

Proof. First, we note that

$$dZ = dE_1 \Upsilon H^H + E_1\, d\Upsilon H^H + E_1 \Upsilon\, dH^H$$

and

$$E_1^H E_1 = I_n, \qquad E_2^H E_1 = 0.$$

Therefore, we have

$$E^H\, dZ H = \begin{bmatrix} E_1^H\, dE_1 \Upsilon + d\Upsilon + \Upsilon\, dH^H H \\ E_2^H\, dE_1 \Upsilon \end{bmatrix}. \tag{11}$$

The exterior product of the left side of equation (11) is equal to

$$(E^H\, dZ H) = (dZ).$$

As in the proof of Theorem1 (see equation (4)), the exterior product of all the elements of $E_2^H\, dE_1\Upsilon$ is

$$(E_2^H\, dE_1\Upsilon) = \left(\prod_{k=1}^{n} \upsilon_k^{2m-2n}\right) \bigwedge_{k=1}^{n} \bigwedge_{l=n+1}^{m} e_l^H\, de_k. \tag{12}$$

Second, we consider $E_1^H\, dE_1\Upsilon + d\Upsilon + \Upsilon dH^H H$. Since H is unitary we have

$$H^H H = I_n \Longrightarrow H^H dH = -dH^H H = -(H^H dH)^H.$$

This implies that the real and imaginary parts of $H^H dH$ are skew-symmetric and symmetric, respectively. Similarly, for $E_1^H\, dE_1$. Thus the $n \times n$ matrix of differential forms is

$$T = E_1^H\, dE_1\Upsilon + d\Upsilon + \Upsilon dH^H H = E_1^H\, dE_1\Upsilon - \Upsilon H^H dH + d\Upsilon, \tag{13}$$

with entries

$$T_{kl} = \begin{cases} d\upsilon_k + \upsilon_k(\mathrm{Im}(e_k^H\, de_k) - \mathrm{Im}(h_k^H\, dh_k)) & k = l, \\ e_k^H\, de_l\upsilon_l - \upsilon_k h_k^H\, dh_l & k > l, \\ e_l^H\, de_k\upsilon_l - \upsilon_k h_l^H\, dh_k & k < l. \end{cases}$$

The exterior product of the elements of T is

$$(T) = \bigwedge_{k=1}^{n} T_{kk} \bigwedge_{k<l}^{n} T_{kl} \wedge T_{lk},$$

where

$$\begin{aligned} T_{kl} \wedge T_{lk} &= \left(\upsilon_k^2 - \upsilon_l^2\right) \mathrm{Re}\left(e_l^H\, de_k \wedge h_l^H\, dh_k\right) \\ &\quad \wedge \left(\upsilon_k^2 - \upsilon_l^2\right) \mathrm{Im}\left(e_l^H\, de_k \wedge h_l^H\, dh_k\right) \\ &= \left(\upsilon_k^2 - \upsilon_l^2\right)^2 e_l^H\, de_k \wedge h_l^H\, dh_k. \end{aligned} \tag{14}$$

Therefore, the exterior product of equation (11) is

$$\begin{aligned} (dZ) &= (2\pi)^{-n}(E_2^H\, dE_1\Upsilon) \wedge (T) \\ &= (2\pi)^{-n}\left(\prod_{k=1}^{n} \upsilon_k^{2m-2n}\right) \bigwedge_{k=1}^{n} \bigwedge_{l=n+1}^{m} e_l^H\, de_k \left(\prod_{k=1}^{n} \upsilon_k\right) \prod_{k<l}^{n} \left(\upsilon_k^2 - \upsilon_l^2\right)^2 \bigwedge_{k=1}^{n} d\upsilon_k \\ &\quad \wedge \bigwedge_{k=1}^{n} \bigwedge_{l=k}^{n} e_l^H de_k \wedge \bigwedge_{k=1}^{n} \bigwedge_{l=k}^{n} h_l^H dh_k \\ &= (2\pi)^{-n}\left(\prod_{k=1}^{n} \upsilon_k^{2m-2n+1}\right) \\ &\quad \times \prod_{k<l}^{n} \left(\upsilon_k^2 - \upsilon_l^2\right)^2 (d\Upsilon) \wedge (E_1^H\, dE_1) \wedge (H^H dH). \end{aligned} \tag{15}$$

Note that we must divide the volume element by $(2\pi)^n$ to normalize the arbitrary phases of the n elements in the first row of E_1. □

The volume of the Stiefel manifold $\mathcal{C}V_{n,m}$ is given by

$$\mathrm{Vol}(\mathcal{C}V_{n,m}) = \int_{\mathcal{C}V_{n,m}} (E_1^H\, dE_1) = \frac{2^n \pi^{mn}}{\mathcal{C}\Gamma_n(m)}, \tag{16}$$

where the complex multivariate gamma function is

$$\mathcal{C}\Gamma_n(m) = \pi^{n(n-1)/2} \prod_{k=1}^{n} \Gamma(m-k+1), \qquad m > n-1.$$

The differential form

$$(dE_1) \triangleq \frac{1}{\mathrm{Vol}[\mathcal{C}V_{n,m}]} (E_1^H dE_1) = \frac{\mathcal{C}\Gamma_n(m)}{2^n \pi^{mn}} (E_1^H\, dE_1) \tag{17}$$

has the property that

$$\int_{\mathcal{C}V_{n,m}} (dE_1) = 1,$$

and it represents the Haar invariant probability measure on $\mathcal{C}V_{n,m}$.

If $m = n$, then we get a special case of Stiefel manifold, the so-called unitary manifold, defined by

$$\mathcal{C}V_{n,n} \equiv U(n) = \{E(n \times n); E^H E = I_n\},$$

that is, the set of unitary $n \times n$ matrices. The volume of $U(n)$ is given by

$$\mathrm{Vol}\,[U(n)] = \int_{U(n)} (E^H\, dE) = \frac{2^n \pi^{n^2}}{\mathcal{C}\Gamma_n(n)}. \tag{18}$$

The probability distributions of random matrices are often derived in terms of hypergeometric functions of matrix arguments. The following complex hypergeometric function of two matrix arguments is required in the sequel, i.e.,

$${}_0F_0^{(n)}(X, Y) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{C_\kappa(X) C_\kappa(Y)}{k! C_\kappa(I_m)},$$

where $X \in \mathbb{C}^{m\times m}$, $Y \in \mathbb{C}^{n\times n}$ and $0 < n < m$. Moreover, $\kappa = (k_1, \ldots, k_n)$ denotes a partition of the integer k with $k_1 \geq \cdots \geq k_n \geq 0$ and $k = k_1 + \cdots + k_n$ and $\sum_\kappa$ denotes summation over all partitions κ of k. The complex zonal polynomial (also called Schur polynomial[11]) of a complex matrix Y defined in [5] is

$$C_\kappa(Y) = \chi_{[\kappa]}(1)\chi_{[\kappa]}(Y), \tag{19}$$

where $\chi_{[\kappa]}(1)$ is the dimension of the representation $[\kappa]$ of the symmetric group,

$$\chi_{[\kappa]}(1) = k!\, \frac{\prod_{i<j}^{n}(k_i - k_j - i + j)}{\prod_{i=1}^{n}(k_i + n - i)!}, \tag{20}$$

and $\chi_{[\kappa]}(Y)$ is the character of the representation $[\kappa]$ of the linear group given as a symmetric function of the eigenvalues, $\lambda_1, \ldots, \lambda_n$, of Y by

$$\chi_{[\kappa]}(Y) = \frac{\det\left[\left(\lambda_i^{k_j+n-j}\right)\right]}{\det\left[\left(\lambda_i^{n-j}\right)\right]}. \tag{21}$$

Note that both the real and complex zonal polynomials are particular cases of the (general α) Jack polynomials $C_\kappa^{(\alpha)}(Y)$, where $\alpha = 1$ for complex and $\alpha = 2$ for real zonal polynomials, respectively. See[12] for details. In this paper we only consider the complex case; therefore, for notational simplicity we drop the superscript of Jack polynomials, as was done in equation (19), i.e., $C_\kappa(Y) := C_\kappa^{(1)}(Y)$. Finally, we have

$$C_\kappa(I_n) = 2^{2k} k! \left[\frac{1}{2}n\right]_\kappa \frac{\prod_{i<j}^{r}(2k_i - 2k_j - i + j)}{\prod_{i=1}^{r}(2k_i + r - i)!}$$

where

$$\left[\frac{1}{2}n\right]_\kappa = \prod_{i=1}^{r}\left(\frac{1}{2}(n - i + 1)\right)_{k_i}$$

for a partition κ of k with r nonzero parts and $(a)_k = a(a+1)\cdots(a+k-1)$.

3 Complex Singular Wishart Matrices

In this section, we derive the complex singular Wishart density and the joint eigenvalue density of the complex singular Wishart matrix.

Theorem 3 *Let m and n be two positive integers such that $0 < n < m$. The density of $\mathrm{W} \sim \mathcal{C}W_m(n, \Sigma)$ on the space $\mathcal{C}S_{m,n}$ of $m \times m$ positive semidefinite Hermitian matrices of rank n is given by*

$$f(W) = \frac{\pi^{n(n-m)}}{\mathcal{C}\Gamma_n(n)(\det \Sigma)^n}\, \mathrm{etr}\left(-\Sigma^{-1}W\right)(\det \Lambda)^{n-m}, \tag{22}$$

where $W = E_1 \Lambda E_1^H$, $E_1 \in \mathcal{C}V_{n,m}$, $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, and etr *denotes the exponential of the trace,* $\mathrm{etr}(\cdot) = \exp(\mathrm{tr}(\cdot))$.

Proof. Let W = $A^H A$, where the $n\times m$ matrix A is distributed as A $\sim \mathcal{C}N(0, I_n \otimes \Sigma)$ and Σ is an $m\times m$ positive definite Hermitian matrix. The density of A is

$$f(A) = \pi^{-nm}(\det\Sigma)^{-n}\operatorname{etr}\left(-A\Sigma^{-1}A^H\right)(dA)$$
$$= \pi^{-nm}(\det\Sigma)^{-n}\operatorname{etr}\left(-\Sigma^{-1}A^H A\right)(dA), \tag{23}$$

where the volume element $(dA) \equiv \bigwedge_{k=1}^{n}\bigwedge_{l=1}^{m} da_{kl}$ is included to facilitate the calculation of Jacobians when we transform A. Transforming $A^H = E_1\Upsilon H$ as in Theorem 2 results in the desired parameterization $W = E_1\Lambda E_1^H$, where $\Lambda = \Upsilon^2$ and $\det\Lambda = \det\Upsilon^2$. Thus we have

$$(d\Lambda) = \bigwedge_{k=1}^{n} d\lambda_k = 2^n\left(\prod_{k=1}^{n}\upsilon_k\right)\bigwedge_{k=1}^{n} d\upsilon_k = 2^n\left(\prod_{k=1}^{n}\upsilon_k\right)(d\Upsilon). \tag{24}$$

From Theorems 1 and 2 and equation (24) we can write (dA) as

$$(dA) = 2^{-n}\left(\prod_{k=1}^{n}\lambda_k^{n-m}\right)(dW)(H^H dH).$$

Since $\prod_{k=1}^{n}\lambda_k = \det\Lambda$, the joint density of W and H is

$$\pi^{-nm}(\det\Sigma)^{-n}\operatorname{etr}\left(-\Sigma^{-1}W\right)2^{-n}(\det\Lambda)^{n-m}(dW)(H^H dH).$$

Integrating with respect to H over the Stiefel manifold $\mathcal{C}V_{n,m}$ and using equation (16), we obtain the marginal density function of W, given by equation (22). □

The following theorem gives the joint density of the eigenvalues of a complex singular Wishart matrix.

Theorem 4 *Let m and n be two positive integers such that $0 < n < m$ and consider the $m\times m$ positive semidefinite Hermitian matrix* W $\sim \mathcal{C}W_m(n,\Sigma)$. *The joint density of the positive eigenvalues, $\lambda_1,\ldots,\lambda_n$, of* W *is*

$$f(\Lambda) = \frac{\pi^{n(n-1)}(\det\Sigma)^{-n}}{\mathcal{C}\Gamma_n(n)\mathcal{C}\Gamma_n(m)}\left(\prod_{k=1}^{n}\lambda_k^{m-n}\right)\prod_{k<l}^{n}(\lambda_k-\lambda_l)^2$$
$$\times\int_{E_1\in\mathcal{C}V_{n,m}}\operatorname{etr}\left(-\Sigma^{-1}E_1\Lambda E_1^H\right)(dE_1), \tag{25}$$

where $W = E_1\Lambda E_1^H$, $E_1 \in \mathcal{C}V_{n,m}$, and $\Lambda = \operatorname{diag}(\lambda_1,\ldots,\lambda_n)$. Moreover

$$\int_{E_1\in\mathcal{C}V_{n,m}}\operatorname{etr}\left(-\Sigma^{-1}E_1\Lambda E_1^H\right)(dE_1) = {}_0F_0^{(n)}(-\Sigma^{-1},\Lambda). \tag{26}$$

Proof. Substituting (2) into the Wishart density (22) and integrating with respect to E_1 over the Stiefel manifold $\mathcal{C}V_{n,m}$, we have

$$f(\Lambda) = \frac{\pi^{n(n-1)}(2\pi)^{-n}}{\mathcal{C}\Gamma_n(n)(\det\Sigma)^n}\left(\prod_{k=1}^{n}\lambda_k^{m-n}\right)\prod_{k<l}^{n}(\lambda_k-\lambda_l)^2$$
$$\times\int_{E_1\in\mathcal{C}V_{n,m}}\operatorname{etr}\left(-\Sigma^{-1}E_1\Lambda E_1^H\right)(E_1^H\,dE_1). \tag{27}$$

Now using (17) we obtain the desired results (25). □

Note that, if $\Sigma = \sigma^2 I_m$, then the joint density of the eigenvalues $\lambda_1,\ldots,\lambda_n$ has a simple form which does not require a complex hypergeometric function representation.

Corollary 1 *Let* W $\sim \mathcal{C}W_m(n,\sigma^2 I_m)$ *with $0 < n < m$. Then the joint density of the eigenvalues, $\lambda_1,\ldots,\lambda_n$, of* W *is*

$$g(\Lambda) = \frac{\pi^{n(n-1)}(\sigma^2)^{-nm}}{\mathcal{C}\Gamma_n(n)\mathcal{C}\Gamma_n(m)}\left(\prod_{k=1}^{n}\lambda_k^{m-n}\right)\prod_{k<l}^{n}(\lambda_k-\lambda_l)^2\exp\left(-\frac{1}{\sigma^2}\sum_{k=1}^{n}\lambda_k\right), \tag{28}$$

where $W = E_1\Lambda E_1^H$, $E_1 \in \mathcal{C}V_{n,m}$, and $\Lambda = \operatorname{diag}(\lambda_1,\ldots,\lambda_n)$.

Proof. Putting $\Sigma = \sigma^2 I_m$ in Theorem 4 and noting that

$$\begin{aligned}\int_{E_1 \in \mathcal{C}V_{n,m}} \operatorname{etr}\left(-\frac{1}{\sigma^2} E_1 \Lambda E_1^H\right)(dE_1) &= \operatorname{etr}\left(-\frac{1}{\sigma^2}\Lambda\right)\int_{E_1 \in \mathcal{C}V_{n,m}} (dE_1) \\ &= \exp\left(-\frac{1}{\sigma^2}\sum_{i=1}^{n}\lambda_i\right) \end{aligned} \tag{29}$$

complete the proof. □

4 The MIMO Channel Capacity

Recently, in response to the demand for higher bit rates in wireless communications, industrial researchers have exploited the use of multiple input, multiple output systems, as shown in Fig. 1 below. These studies show that MIMO systems increase capacity significantly over single input, single output (SISO) systems. For example, if $n = \min\{n_t, n_r\}$, a MIMO uncorrelated Rayleigh distributed channel achieves almost n more bits per hertz for every 3-dB increase in signal-to-noise ratio (SNR) compared to a SISO system, which achieves only one additional bit per hertz for every 3-dB increase in SNR [2]. But the channel coefficients from different transmitter antennas to a single receiver antenna can be correlated. This channel correlation, which degrades the channel capacity [13], depends on the physical parameters of a MIMO system and the scatterer characteristics. The physical parameters include the antenna arrangement and spacing, the angle spread, the angle of arrival, etc. One of the objectives of this paper is to evaluate this capacity degradation for the channel matrix $\mathrm{H} \sim \mathcal{C}N(0, I_{n_r} \otimes \Sigma)$ with $n_t > n_r$. This will be done by deriving closed form ergodic capacity formulas for correlated channels and their numerical evaluation.

The complex signal received at the jth output can be written as

$$y_j = \sum_{i=1}^{n_t} h_{ij} x_i + v_j, \tag{30}$$

where h_{ij} is the complex channel coefficient between input i and output j, x_i is the complex signal at the ith input and v_j is complex Gaussian noise. The signal vector received at the output can be written as

$$\begin{bmatrix} y_1 \\ \vdots \\ y_{n_r} \end{bmatrix} = \begin{bmatrix} h_{11} & \cdots & h_{n_t 1} \\ \vdots & \vdots & \vdots \\ h_{1 n_r} & \cdots & h_{n_t n_r} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_{n_t} \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_{n_r} \end{bmatrix},$$

i.e., in vector notation,

$$y = Hx + v, \tag{31}$$

where $y, v \in \mathbb{C}^{n_r}$, $H \in \mathbb{C}^{n_r \times n_t}$, and $x \in \mathbb{C}^{n_t}$ (see Fig. 1).

The total power of the input is constrained to ρ,

$$\mathcal{E}\{x^H x\} \le \rho \quad \text{or} \quad \operatorname{tr} \mathcal{E}\{x x^H\} \le \rho.$$

In this section, we shall deal exclusively with the linear model (31) and compute the capacity of MIMO channel models for $n_t > n_r$.

We assume that H is a complex Gaussian random matrix whose realization is known to the receiver, or equivalently, the channel output consists of the pair (y, H). Note that the transmitter does not know the channel and the input power is distributed equally over all transmitting antennas. Moreover, if we assume a block-fading model and coding over many independent fading intervals, then the Shannon or ergodic capacity of the random MIMO channel [2] is given by

$$C = \mathcal{E}_{\mathrm{H}}\left\{\log\det\left(I_{n_t} + (\rho/n_t)\mathrm{H}^H\mathrm{H}\right)\right\}, \tag{32}$$

where the expectation is evaluated using a complex Gaussian density. If $\mathrm{H} \sim \mathcal{C}N(0, I_{n_r} \otimes \Sigma)$ then the channel is Rayleigh distributed and correlated at the transmitter end. This is typical of fixed or mobile communication environments. Here the covariance matrix of the rows of H is denoted by Σ, which is an $n_t \times n_t$ positive definite Hermitian matrix. Let $\mathrm{W} = \mathrm{H}^H\mathrm{H} \sim \mathcal{C}W_{n_t}(n_r, \Sigma)$ and $n_t > n_r$. Then the channel capacity can be written as

$$C = \mathcal{E}_{\mathrm{W}}\left\{\log\det\left(I_{n_t} + (\rho/n_t)\mathrm{W}\right)\right\}, \tag{33}$$

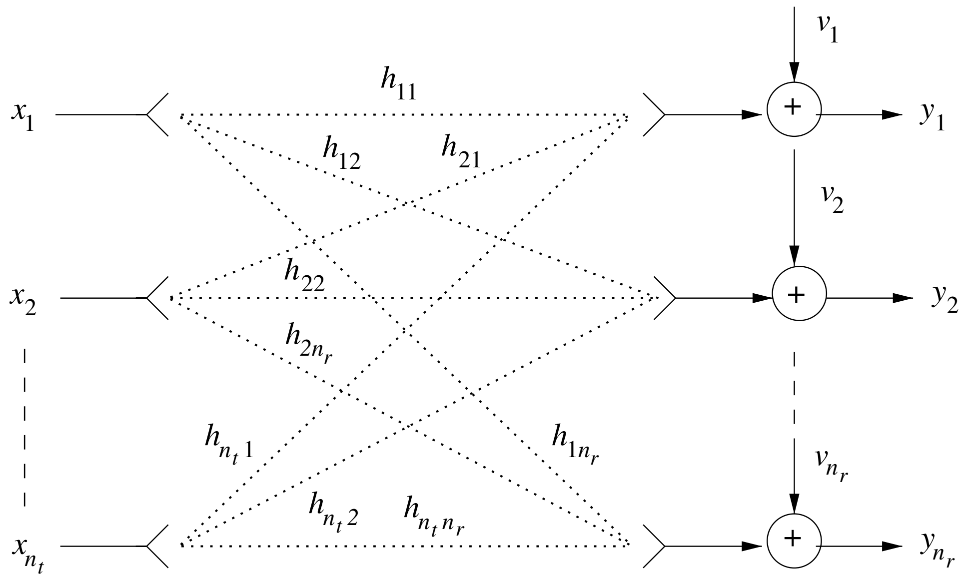

Figure 1: A MIMO communication system.

where the expectation is evaluated using a complex singular Wishart density given in Theorem 3. Let $\lambda_1 > \cdots > \lambda_{n_r}$ be the eigenvalues of W and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_{n_r})$. Then the capacity can also be computed using a joint eigenvalue density, $f(\Lambda)$, or the single unordered eigenvalue density, $f(\lambda)$, i.e.,

$$\begin{aligned} C &= \mathcal{E}_\Lambda \left\{ \log \left(\prod_{k=1}^{n_r} [1 + (\rho/n_t)\lambda_k] \right) \right\} \\ &= \sum_{k=1}^{n_r} \mathcal{E}_{\lambda_k} \{\log(1 + (\rho/n_t)\lambda_k)\} \\ &= n_r \mathcal{E}_\lambda \{\log(1 + (\rho/n_t)\lambda)\}. \end{aligned} \tag{34}$$

The joint eigenvalue density of a complex singular Wishart matrix is given in Theorem 4. From this joint eigenvalue density we can obtain a single unordered eigenvalue density, $f(\lambda)$, by dividing $f(\Lambda)$ by $n_r!$ and integrating with respect to $\lambda_2, \ldots, \lambda_{n_r}$.

As a numerical example, we compute the channel capacity of a correlated 2×4 channel matrix ($n_r = 2$ and $n_t = 4$), i.e., $\mathrm{H} \sim \mathcal{C}N(0, I_2 \otimes \Sigma)$, where we assume the eigenvalues of the positive definite Hermitian matrix Σ are

$$1.8090, \quad 1.3090, \quad 0.6910, \quad 0.1910.$$

In this case W is 4×4 complex singular Wishart matrix with two nonzero eigenvalues λ_1 and λ_2. The joint eigenvalue distribution is given by

$$f(\lambda_1, \lambda_2) = \frac{1}{12(\det \Sigma)^2} (\lambda_1\lambda_2)^2(\lambda_1 - \lambda_2)^2 {}_0F_0^{(2)}(-\Sigma^{-1}, \Lambda), \tag{35}$$

where $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2)$. The capacity of the correlated channel, $\mathrm{H} \sim \mathcal{C}N(0, I_2 \otimes \Sigma)$, is given by

$$C_c = \int_0^\infty \int_0^{\lambda_1} [\log(1 + (\rho/4)\lambda_1) + \log(1 + (\rho/4)\lambda_2)] \, f(\lambda_1, \lambda_2) \, d\lambda_2 \, d\lambda_1. \tag{36}$$

For comparison purpose we also compute the capacity of the uncorrelated 2×4 channel matrix, i.e., $\mathrm{H} \sim \mathcal{C}N(0, I_2 \otimes \sigma^2 I_4)$. In this case, the joint eigenvalue density of the complex singular Wishart matrix is given by

$$g(\lambda_1, \lambda_2) = \frac{1}{12\sigma^{16}} (\lambda_1\lambda_2)^2(\lambda_1 - \lambda_2)^2 e^{-(\lambda_1+\lambda_2)/\sigma^2}. \tag{37}$$

From equation (37), we can evaluate the single unordered eigenvalue density, $f(\lambda)$, which is given by

$$g(\lambda) = \frac{\lambda^2 e^{-\lambda/\sigma^2}}{12\sigma^6} \left(\frac{\lambda^2}{\sigma^4} - 6\frac{\lambda}{\sigma^2} + 12 \right). \tag{38}$$

The capacity of this uncorrelated channel $\mathrm{H} \sim \mathcal{C}N(0, I_2 \otimes \sigma^2 I_4)$ is given by

$$C_u = 2\int_0^{\infty} \log(1 + (\rho/4)\lambda)\, g(\lambda)\, d\lambda. \tag{39}$$

Figure 2 shows the capacity in nats[5] vs signal-to-noise ratio for a 2×4 correlated/uncorrelated Rayleigh fading channel matrix. From this figure we note the following: (i) the capacity is decreasing with channel correlation, (ii) the capacity is increasing with SNR.

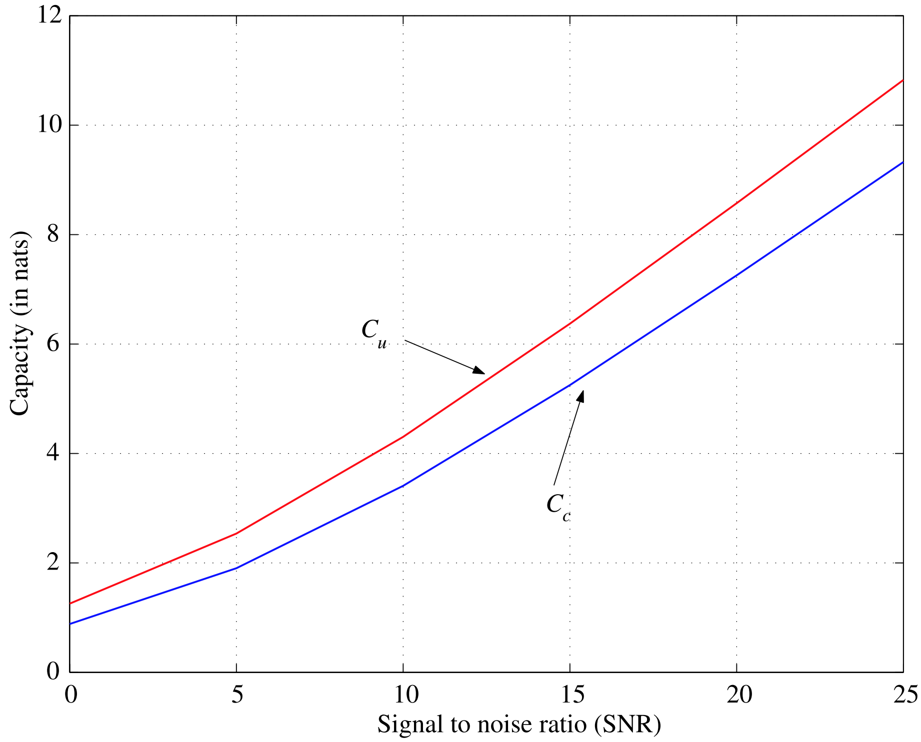

Figure 2: Capacity vs SNR for $n_t = 4$ and $n_r = 2$, i.e., 2×4 channel matrix. C_u and C_c denote the capacity of uncorrelated and correlated Rayleigh channels, respectively.

5 Conclusion

In this paper, we studied the complex singular Wishart distribution and its application. In particular, we derived the complex singular Wishart density and joint eigenvalue density of a complex singular Wishart matrix. Using these distributions, both correlated and uncorrelated MIMO Rayleigh channel capacity formulas were obtained. The capacities of 2×4 MIMO Rayleigh channel matrices were computed for b correlated and uncorrelated channels. It was also shown how channel correlation degrades the capacity of the communication systems.

Appendix

The real singular Wishart density is derived in [7, Theorem 6]. In this appendix we give the joint eigenvalue density of a real singular Wishart matrix. Note that we follow the notation of [7] and [9].

Theorem 5 *Let m and n be two positive integers such that $0 < n < m$ and consider an $m \times m$ positive semidefinite symmetric Wishart matrix $\mathrm{W} \sim \mathcal{W}_m(n, \Sigma)$. Then the joint density of the positive eigenvalues, $\lambda_1, \ldots, \lambda_n$, of W is*

$$\frac{\pi^{n^2/2} 2^{-mn/2} (\det \Sigma)^{-n/2}}{\Gamma_n(n/2)\Gamma_n(m/2)} \left(\prod_{k=1}^{n} \lambda_k^{(m-n-1)/2} \right) \prod_{k<l}^{n} (\lambda_k - \lambda_l)$$
$$\times \int_{E_1 \in \mathcal{V}_{n,m}} \operatorname{etr}\left(-\frac{1}{2}\, \Sigma^{-1} E_1 \Lambda E_1^H \right) (dE_1),$$

[5]In equation (34), if we use $\log_e$ then the capacity is measured in nats. If we use $\log_2$ then the capacity is measured in bits. Thus, one nat is equal to e bits/sec/Hz ($e = 2.718\ldots$).

where $W = E_1 \Lambda E_1^H$, $E_1 \in \mathcal{V}_{n,m}$ *and* $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. *Moreover*

$$\int_{E_1 \in \mathcal{V}_{n,m}} \mathrm{etr}\left(-\frac{1}{2}\,\Sigma^{-1} E_1 \Lambda E_1^H\right)(dE_1) = {}_0F_0^{(n)}\left(-\frac{1}{2}\,\Sigma^{-1}, \Lambda\right).$$

References

[1] R. B. Ash, *Information Theory*, Dover, New York, 1965.

[2] I. E. Telatar, *Capacity of multi-antenna Gaussian channels*, Eur. Trans. Telecom, **10** (1999) 585–595.

[3] T. Ratnarajah, R. Vaillancourt and M. Alvo, *Complex random matrices and Rayleigh channel capacity*, Communications in Information & Systems, **3** (Oct. 2003) 119-138.

[4] N. R. Goodman, *Statistical analysis based on a certain multivariate complex Gaussian distribution (An introduction)*, Ann. Math. Statist., **34** (1963) 152–177.

[5] A. T. James, *Distributions of matrix variate and latent roots derived from normal samples*, Ann. Math. Statist., **35** (1964) 475–501.

[6] T. Ratnarajah, R. Vaillancourt and M. Alvo, *Eigenvalues and condition numbers of complex random matrices*, SIAM Journal on Matrix Analysis and Applications, to appear.

[7] H. Uhlig, *On singular Wishart and singular multivariate beta distributions*, Ann. of Statist., **22** (1994) 395–405.

[8] J. Shen, *On the singular values of Gaussian random matrices*, Linear Alg. Appl., **326** (2001) 1–14.

[9] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, Wiley, New York, 1982.

[10] T. Ratnarajah, R. Vaillancourt and M. Alvo, *Jacobians and hypergeometric functions in complex multivariate analysis*, Canadian Applied Mathematics Quarterly, to appear.

[11] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, Oxford University Press, New York, 1995.

[12] T. Baker and P. Forrester, *The Calogero-Sutherland model and generalized classical polynomials*, Commun. Math. Phys., **188** (1997) 175–216.

[13] C. N. Chuah, D. Tse, J. M. Kahn and R. A. Valenzuela, *Capacity scaling in MIMO wireless systems under correlated fading*, IEEE Trans. on Information Theory, **48** (March 2002) 637–650.