Université d'Ottawa • University of Ottawa

PERMISSION DE REPRODUIRE ET DE DISTRIBUER LA THÈSE

PERMISSION TO REPRODUCE AND DISTRIBUTE THE THESIS

NOM DE L'AUTEUR / NAME OF AUTHOR: Tharmalingam Ratnarajah

ADRESSE POSTALE / MAILING ADDRESS: 45 Farmfield Cr.
Kanata, Ontario K2M 2S8

GRADE / DEGREE: Ph.D (Mathematics and Statistics)

ANNÉE D'OBTENTION / YEAR GRANTED 2003

TITRE DE LA THÈSE / TITLE OF THESIS:
Topics in Complex Random Matrices and Information Theory

09 May 2003

DATE

(AUTEUR) SIGNATURE (AUTHOR)

Université d'Ottawa · University of Ottawa

Université d'Ottawa · University of Ottawa

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES

FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

RATNARAJAH, Tharmalingam

AUTEUR DE LA THÈSE - AUTHOR OF THESIS

Ph.D. (Mathematics)

GRADE - DEGREE

Mathematics and Statistics

FACULTÉ, ÉCOLE, DÉPARTEMENT - FACULTY, SCHOOL, DEPARTMENT

TITRE DE LA THÈSE - TITLE OF THE THESIS

Topics in Complex Random Matrices and Information Theory

Rémi Vaillancourt

DIRECTEUR DE LA THÈSE - THESIS SUPERVISOR

EXAMINATEURS DE LA THÈSE - THESIS EXAMINERS

A. Dabrowski

M. Zarepour

N. Balakrishnan

J.N. Pandey

J.-M. De Koninck, Ph.D.

LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES

SIGNATURE

DEAN OF THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

TOPICS IN COMPLEX RANDOM MATRICES AND INFORMATION THEORY

By

Tharmalingam Ratnarajah

May 2003

A Thesis
submitted to the School of Graduate Studies and Research
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy in Mathematics[1]

[1]The Ph.D. Program is a joint program with Carleton University, administered by the Ottawa-Carleton Institute of Mathematics and Statistics

0-612-79314-1

Abstract

The eigenvalue distribution of both central and noncentral complex Wishart matrices are investigated with the objective of studying several open problems in information theory and numerical analysis, etc. Specifically, the largest, kth largest, and the smallest eigenvalue distributions of complex Wishart matrices and the condition number distribution of complex Gaussian random matrices are derived. These distributions are represented by complex hypergeometric functions of matrix arguments, which can be expressed in terms of complex zonal polynomials. We derive several results on complex hypergeometric functions and zonal polynomials that are used to evaluate these distributions. We also give a method to compute these complex hypergeometric functions.

Then the connection between the complex Wishart matrix theory and information theory is formulated. This facilitates the evaluation of the most important information-theoretic measure, the so-called channel capacity. The capacity of the communication channel expresses the maximum rate at which information can be reliably conveyed by the channel. In particular, the capacities of the multiple input, multiple output Rayleigh and Rician distributed channels are fully investigated. We consider both correlated and uncorrelated channels and derive the corresponding channel capacity formulas. These studies show how the channel correlations degrade the capacity of the communication system.

Acknowledgements

I would like to express my gratitude to Remi Vaillancourt and Mayer Alvo for their support and encouragement during the last four years. Their provocative questions, thoughtful discussions, and careful comments and criticism, have greatly influenced me, and are reflected throughout this work. I consider it a great privilege to have had the opportunity to work with them.

I am greatly indebted to my teachers, particularly J. N. Pandey, Andre Dabrowski, Yves Bourgault and David McDonald, for their guidance, encouragement and enthusiasm. I would also like to thank Rasanayakam for introducing me to the wonderful world of mathematics.

Much credit goes to the friends and colleagues with whom I have had the pleasure of interacting throughout my stay in Ottawa. I must single out Wanye Dam, the stimulation provided by his insightful comments has been a constant source of encouragement, for which I am grateful. I would also like to thank Susanne Dalrymple and other support staffs for their help.

Finally, there has been no more constant a source of encouragement than my wife, parents, sisters, and brothers, to all of whom this effort is dedicated. They have participated in this endeavour, both by encouraging and supporting my efforts, and by making significant sacrifices. They deserve much of the credit for seeing this project through to its conclusion.

Dedication

I dedicate this work to my father and mother.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this work, we investigate the distributions of the eigenvalues and condition number of complex random matrices and their applications to information theory and numerical analysis. In contrast to the literature in [12], we consider that the elements of random matrices are complex Gaussian distributed with arbitrary mean and covariance matrices. This will enable us to consider the beautiful but difficult theory of zonal polynomials, which are symmetric polynomials in the eigenvalues of a complex matrix, see [36] and [44]. Zonal polynomials enable us to represent the distributions of the eigenvalues of these complex random matrices as infinite series.

Let an $n \times m$ complex Gaussian random matrix $\mathbf{A}$ be distributed as $\mathbf{A} \sim \mathcal{CN}(M, I_n \otimes \Sigma)$. Then $\mathbf{W} = \mathbf{A}^H \mathbf{A}$ is called a noncentral Wishart matrix. The mean and the covariance matrices of $\mathbf{A}$ are equal to M and $I_n \otimes \Sigma$, respectively, i.e., $\mathcal{E}\{\mathbf{A}\} = M$ and $\text{cov}\{\mathbf{A}\} = I_n \otimes \Sigma$. If $M = \mathbf{0}$, then $\mathbf{W}$ is called a central Wishart matrix,

The square of the condition number of a matrix A is defined by the ratio of the largest to the smallest eigenvalues of the matrix $A^H A$. The condition number of a random matrix gives valuable information on the numerical stability of the convergence analysis in optimization algorithms or, more importantly, on the reliability of the solution of linear systems of equations. We derive the eigenvalue distribution of a complex central Wishart matrix and the condition number distribution of complex random matrices in Chapter 4. Moreover, noncentral complex Wishart matrices are studied in Chapter 5.

The theory of these random matrices is used to evaluate the capacity of multiple input, multiple output (MIMO) wireless communication systems. Note that the capacity of the communication channel expresses the maximum rate at which information can be reliably conveyed by the channel [2]. A MIMO channel can be represented by an $n_r \times n_t$ complex random matrix $\mathbf{H}$, where n_t and n_r are the number of inputs (or transmitters) and outputs (or receivers) of the communication system, as shown in Figure 1.1. The complex signal received at the jth output can be written as

$$y_j = \sum_{i=1}^{n_t} h_{ij} x_i + v_j, \tag{1.1}$$

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

$$\begin{aligned} \begin{bmatrix} y_1 \\ \vdots \\ y_{n_r} \end{bmatrix} &= \begin{bmatrix} h_{11} & \dots & h_{n_t 1} \\ \vdots & \vdots & \vdots \\ h_{1 n_r} & \dots & h_{n_t n_r} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_{n_t} \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_{n_r} \end{bmatrix}, \\ y &= Hx + v, \end{aligned} \tag{1.2}$$

where $y, v \in \mathbb{C}^{n_r}$, $H \in \mathbb{C}^{n_r \times n_t}$, and $x \in \mathbb{C}^{n_t}$. The total power of the input is constrained to ρ,

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

It should be noted, in a wireless communication system, data is delivered from a transmitter to a receiver using radio waves or other electromagnetic waves. The waves, however, may be reflected off objects in the environment and scattered randomly while propagating from the transmitter to the receiver. Therefore, transmitted signals are attenuated and phase shifted during the transmission. This channel response can be modeled by complex channel coefficient. Recently researchers have exploited the use of multiple input, multiple output system in response to the demand for higher bit rates in wireless communications. These studies show that MIMO system increases capacity significantly over single input, single output (SISO) system. For example a

MIMO system achieves almost n more bits per hertz for every 3-dB increase in signal-to-noise ratio (SNR) compared to SISO case, which only achieves one additional bit per hertz for every 3-dB increase in SNR, where $n = \min\{n_t, n_r\}$. However, the channel coefficients from different transmitter antennas to a single receiver antenna can be correlated. This channel correlation degraded the capacity. One of the objectives in this work is to evaluate the capacity degradation. This have been done by deriving the closed form capacity formulas for correlated channel and evaluating numerically.

We shall deal exclusively with the linear model (1.2) and derive the capacity of different MIMO channel models in Chapters 6 and 7. In this work, we are particularly interested in two channel models, namely the *Rayleigh* and *Rician* distributed channels. The following proposition defines those channel models [34].

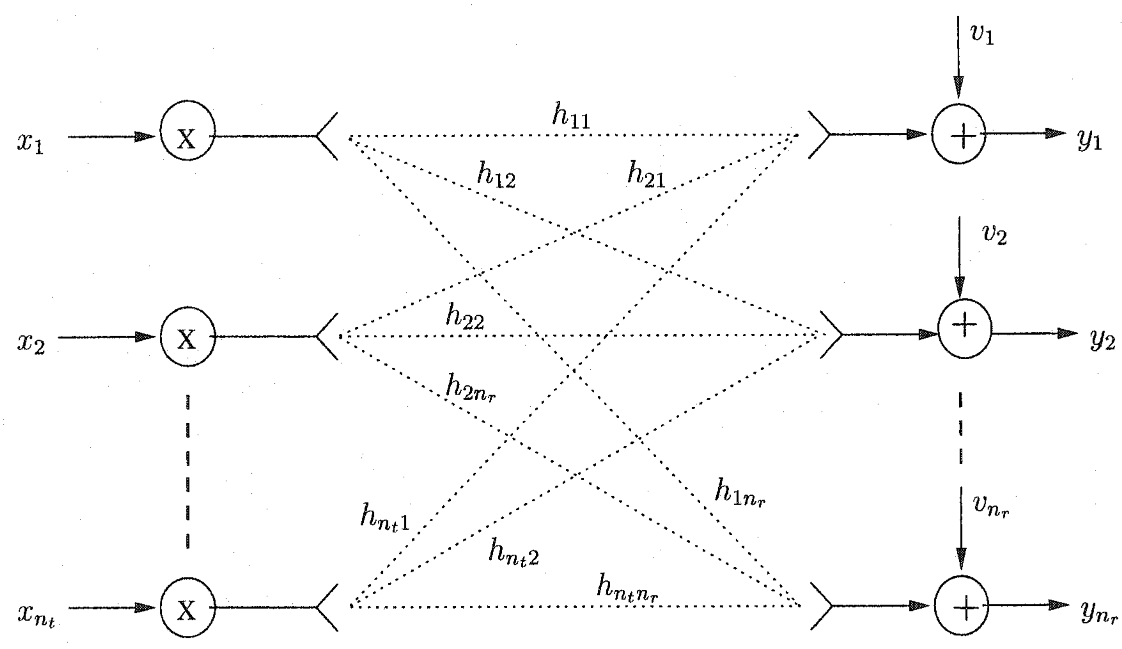

Figure 1.1: MIMO communication system.

Proposition 1.1 *Let* $z = re^{i\theta}$ $(= h_{ij}) \sim \mathcal{CN}(\mu_z, \sigma^2)$, *where* $r = |z|$ *and* $\theta = \arg z$,

$$\mathcal{E}\{z\} = \mu_z = |\mu_z| \exp(i\phi), \quad \text{var}\{z\} = \mathcal{E}|z - \mu_z|^2 = \sigma^2$$

and

$$f(z) = \frac{1}{\pi\sigma^2} \exp\left(\frac{-|z-\mu_z|^2}{\sigma^2}\right).$$

Then the joint density $g(r,\theta|\mu_z,\sigma^2)$ *of* r *and* θ *is given by*

$$g(r,\theta|\mu_z,\sigma^2) = \frac{r}{\pi\sigma^2} \exp\left(\frac{-(r^2+|\mu_z|^2)}{\sigma^2}\right) \exp\left(\frac{2|\mu_z| r\cos(\theta-\phi)}{\sigma^2}\right). \tag{1.3}$$

Therefore, the density of the magnitude or envelope r *is given by*

$$\begin{aligned} h(r|\mu_z,\sigma^2) &= \frac{r}{\pi\sigma^2} \exp\left(\frac{-(r^2+|\mu_z|^2)}{\sigma^2}\right) \int_0^{2\pi} \exp\left(\frac{2|\mu_z| r\cos(\theta-\phi)}{\sigma^2}\right) d\theta \\ &= \begin{cases} \frac{2r}{\sigma^2} \exp\left(\frac{-(r^2+|\mu_z|^2)}{\sigma^2}\right) I_0\left(\frac{2|\mu_z| r}{\sigma^2}\right) & r \geq 0 \\ 0 & r < 0, \end{cases} \end{aligned} \tag{1.4}$$

where I_0 *is the modified Bessel function of the first kind and order zero. The density of the phase* θ *is given by*

$$\begin{aligned} k(\theta|\mu_z,\sigma^2) &= \frac{1}{\pi\sigma^2} \exp\left(\frac{-|\mu_z|^2}{\sigma^2}\right) \int_0^{\infty} r\exp\left(\frac{-r^2}{\sigma^2}\right) \exp\left(\frac{2|\mu_z| r\cos(\theta-\phi)}{\sigma^2}\right) dr \\ &= \begin{cases} \frac{1}{2\pi} \exp\left(\frac{-|\mu_z|^2[1+\sin^2(\theta-\phi)]}{2\sigma^2}\right) D_{-2}\left(-|\mu_z|\left(\frac{2}{\sigma^2}\right)^{1/2}\cos(\theta-\phi)\right) \\ \qquad 0 \leq \theta < 2\pi \\ 0 \qquad \text{otherwise}, \end{cases} \end{aligned} \tag{1.5}$$

where D_{-v} *is the parabolic cylinder function. The density of* $h(r|\mu_z,\sigma^2)$ *is called the Rician density. If* $\mu_z = 0$ *then the density of* $h(r|\sigma^2)$ *is called the Rayleigh density and is given by*

$$h(r|\sigma^2) = \begin{cases} \frac{2r}{\sigma^2} \exp\left(\frac{-r^2}{\sigma^2}\right) & r \geq 0, \\ 0 & r < 0. \end{cases} \tag{1.6}$$

In this case, the distribution of the phase θ *is uniform and its density is given by*

$$k(\theta|\sigma^2) = \begin{cases} \frac{1}{2\pi} & 0 \leq \theta < 2\pi, \\ 0 & \text{otherwise}. \end{cases} \tag{1.7}$$

□

Note that the integral in (1.4) is independent of ϕ as can be seen by making the transformation $\beta = \theta - \phi$.

1.1 Contributions

The major contributions of this work are listed in this section. We believe that the theorems stated in this thesis are new and significant contributions to the literature. In addition, all the Corollaries and most of the Lemmas are the author's contributions. The statements labeled as propositions are complex analogues of the real cases, some are derived by the author and others are in the literature.

- Chapter 3 gives the theory of complex hypergeometric functions and zonal polynomials. We derived several results on these topics that are used to evaluate the distributions of random matrices and eigenvalues. We also include important results from the literature and this is the first time that these results are presented in one place. Note that some of the propositions have their real counterparts in [36], [31] and [15] as we indicate in Table 1.1.

- Chapter 4 studies the complex central Wishart matrix. Contributions of this chapter are Theorems 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, Corollaries 4.1, 4.2, and Lemmas 4.4, 4.5. Corollary 4.1 gives distribution of the largest eigenvalue of a complex central Wishart matrix, which is derived from Theorem 4.1. Likewise, Corollary 4.2 gives the distribution of the smallest eigenvalue of a complex central Wishart matrix, which is derived from Theorem 4.2. Theorems 4.3–4.6 and Lemmas 4.4 and 4.5 are related to the density of the condition number of a complex central Wishart matrix.

- Chapter 5 studies the complex noncentral Wishart matrix. The contributions are Theorems 5.1, 5.2, 5.3, 5.4 and Corollaries 5.1 and 5.2. Theorem 5.1 gives the complex noncentral Wishart density. Corollary 5.1 expresses the complex noncentral Wishart density in terms of the complex central Wishart density. Theorem 5.2 gives the joint eigenvalue density. Theorems 5.3 and 5.4 express the distribution of the largest and the smallest eigenvalue of a complex noncentral Wishart matrix, respectively.

- In Chapter 6, we provide a general framework to study the MIMO channel capacity and related information theory.

- Chapter 7 gives computational methods of channel capacity for different channel models. Theorems 7.1 – 7.3 and Lemma 7.1 are related to the correlated Rayleigh channel capacity evaluation. Theorem 7.4 gives the $n_r \times 2$ correlated Rayleigh channel capacity. Theorems 7.5 – 7.7 are related to uncorrelated Rayleigh channel capacity evaluation. Theorem 7.8 gives the $n_r \times 2$ uncorrelated Rayleigh channel capacity. Finally, Theorems 7.9 – 7.13 and Lemma 7.2 are related to Rician channel capacity.

Complex results	Real counterpart results in [36]
Proposition 2.2	Theorem 2.1.9
Proposition 2.4	Theorem 2.1.13
Proposition 2.5	Theorem 2.1.11
Proposition 2.6	Theorem 2.1.12
Proposition 2.7	Theorem 2.1.15
Corollary 2.1	Corollary 2.1.16
Proposition 3.1	Theorems 7.2.7 / 7.2.13
Proposition 3.2	Theorem 7.2.10
Proposition 3.3	Theorem 7.3.3
Proposition 3.4	Theorem 7.3.4
Proposition 3.6	Theorem 7.4.1

Table 1.1: The table lists the real counterpart results to the complex results derived in Chapters 2 and 3.

Chapter 2

Preliminaries

In this chapter, we provide the necessary tools for deriving the distribution theory in the subsequent chapters. Section 2.1 defines exterior products and Jacobians. Section 2.2 provides the complex multivariate Gamma function. Finally, in section 2.3, the Stiefel manifold and Haar measure are defined.

2.1 Exterior products and Jacobians

In order to transform density functions we need to compute the determinant of a matrix of partial derivatives (so-called Jacobian). It is tedious to explicitly write down this determinant when dealing with many variables. To overcome this difficulty an equivalent approach based on exterior product of differential forms was introduced in [20] and [36]. This approach is outlined next for a bivariate transformation. Consider the integral

$$I = \int_D f(x_1, x_2) dx_1 dx_2, \tag{2.1}$$

where $D \subset R^2$. By making the change of variables

$$x_1 = x_1(y_1, y_2) \quad \text{and} \quad x_2 = x_2(y_1, y_2)$$

Equation (2.1) becomes

$$I = \int_{D'} f(x(y)) \det \left[\left(\frac{\partial x_k}{\partial y_l} \right) \right] dy_1 dy_2, \quad k, l = 1, 2, \tag{2.2}$$

where D' denotes the image of D. The equivalent representation of Equation (2.2) based on the exterior product of differential forms is given by

$$I = \int_{D'} f(x(y)) \left(\frac{\partial x_1}{\partial y_1} dy_1 + \frac{\partial x_1}{\partial y_2} dy_2 \right) \wedge \left(\frac{\partial x_2}{\partial y_1} dy_1 + \frac{\partial x_2}{\partial y_2} dy_2 \right), \tag{2.3}$$

where the exterior product of two differentials dy_k and dy_l satisfies

(i) $dy_k \wedge dy_l = -dy_l \wedge dy_k = dy_k dy_l$ (skew-symmetric),

(ii) $dy_k \wedge dy_k = -dy_k \wedge dy_k = 0$.

For an arbitrary $n \times m$ matrix X, the symbol (dX) denotes the exterior product of the mn elements of dX

$$(dX) \equiv \bigwedge_{l=1}^{m} \bigwedge_{k=1}^{n} dx_{kl}.$$

For a symmetric $m \times m$ matrix X, the symbol (dX) denotes the exterior product of the $m(m+1)/2$ distinct elements of dX

$$(dX) \equiv \bigwedge_{1 \le k \le l \le m} dx_{kl}.$$

Similarly, if X is a skew-symmetric matrix ($X = -X^T$), then (dX) denotes the exterior product of the $m(m-1)/2$ distinct elements of dX, and if X is an upper-triangular matrix, then

$$(dX) \equiv \bigwedge_{k \le l} dx_{kl}.$$

Moreover, if $X = X_r + iX_c$ is a complex matrix, then $(dX) = (dX_r)(dX_c)$. The following proposition gives the Jacobian for a Hermitian transformation.

Proposition 2.1 *If $X = BYB^H$, where X and Y are $m \times m$ Hermitian matrices* *and B is a (fixed) nonsingular $m \times m$ matrix, then*

$$(dX) = (\det B)^{2m}(dY).$$

Proof. Note that the real and complex parts of Hermitian matrices are symmetric and skew-symmetric matrices, respectively. If $X = X_r + iX_c$ and $Y = Y_r + iY_c$, then X_r and Y_r are symmetric matrices and X_c and Y_c are skew-symmetric matrices. From [36], we have

$$(dX_r) = (\det B)^{m+1}(dY_r) \quad \text{and} \quad (dX_c) = (\det B)^{m-1}(dY_c).$$

Therefore,

$$(dX) = (dX_r)(dX_c) = (\det B)^{2m}(dY_r)(dY_c) = (\det B)^{2m}(dY).$$

The proof is complete. □

The Jacobian for the change of variables in the Cholesky factorization is given by the following proposition [16]. Here we derive this Jacobian based on the exterior product.

Proposition 2.2 *Let A be an $m \times m$ Hermitian positive definite matrix. The Cholesky factorization is given by $A = T^H T$, where T is upper triangular with real and positive diagonal elements. Then we have*

$$(dA) = 2^m \prod_{k=1}^{m} t_{kk}^{2m-2k+1}(dT).$$

Proof. Let $A = (a_{kl} + ib_{kl})$ and $T = (t_{kl} + iu_{kl})$. By considering the real part we have

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1m} & a_{2m} & \dots & a_{mm} \end{bmatrix} = \begin{bmatrix} t_{11}^2 & t_{11}t_{12} & \dots & t_{11}t_{1m} \\ \vdots & \vdots & \vdots & \vdots \\ t_{1m}t_{11} & \dots & \dots & t_{mm}^2 + \sum_{k=1}^{m} t_{km}^2 + \sum_{k=1}^{m} u_{km}^2 \end{bmatrix}.$$

For the imaginary part

$$\begin{bmatrix} 0 & b_{12} & \dots & b_{1m} \\ \vdots & \vdots & \vdots & \vdots \\ b_{1m} & b_{2m} & \dots & 0 \end{bmatrix} = \begin{bmatrix} 0 & t_{11}u_{12} & \dots & t_{11}u_{1m} \\ \vdots & \vdots & \vdots & \vdots \\ -u_{1m}t_{11} & \dots & \dots & 0 \end{bmatrix}.$$

Next, expressing the diagonal and the upper diagonal elements of A in term of elements of T and taking differentials we get the following for their real parts

$$\begin{aligned} da_{11} &= 2t_{11}dt_{11} \\ da_{12} &= t_{11}dt_{12} + \dots \\ &\vdots \\ da_{mm} &= 2t_{mm}dt_{mm} + \dots \end{aligned}$$

and for their complex parts

$$\begin{aligned} db_{12} &= t_{11}du_{12} + \dots \\ db_{13} &= t_{11}du_{13} + \dots \\ &\vdots \\ db_{m-1,m} &= t_{m-1,m-1}du_{m-1,m} + \dots. \end{aligned}$$

Note that we only retained the terms that contribute to the exterior product because the products of repeated differentials are zero. Now, by taking the exterior product we get

$$\begin{aligned} (dA) = (dA_r)(dA_c) &= \bigwedge_{k\leq l}^{m} da_{kl} \bigwedge_{k<l}^{m} db_{kl} \\ &= \left(2^m t_{11}^m t_{22}^{m-1} \cdots t_{mm} \bigwedge_{k\leq l}^{m} dt_{kl}\right)\left(t_{11}^{m-1} t_{22}^{m-2} \cdots t_{m-1,m-1} \bigwedge_{k<l}^{m} du_{kl}\right) \\ &= 2^m \prod_{k=1}^{m} t_{kk}^{2m-2k+1}(dT). \end{aligned} \tag{2.4}$$

The proof is complete. □

The Jacobian for the change of variables in the eigendecomposition is given by the following proposition, see [33] and [49]. Here we derive this Jacobian based on the exterior product.

Proposition 2.3 *Let A be an $m \times m$ Hermitian positive definite matrix. The eigen-decomposition of A is $A = E\Lambda E^H$, where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ and E is a unitary matrix. Then we have*

$$(dA) = \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 (d\Lambda)(E^H dE),$$

where

$$(E^H dE) = \bigwedge_{k<l}^{m} e_l^H de_k$$

and e_k are the columns of E.

Proof. Since E is unitary we have

$$E^H E = I \Longrightarrow E^H dE = -dE^H E = -(E^H dE)^H.$$

This implies that the real and the imaginary parts of $E^H dE$ are skew-symmetric and symmetric, respectively. Moreover,

$$dA = dE\Lambda E^H + Ed\Lambda E^H + E\Lambda dE^H.$$

Multiplying by E^H on the left and E on the right, we get

$$\begin{aligned} E^H dAE &= E^H dE\Lambda + d\Lambda + \Lambda dE^H E \\ &= E^H dE\Lambda - \Lambda E^H dE + d\Lambda. \end{aligned} \tag{2.5}$$

Using Proposition 2.1 we can write the exterior product of the Hermitian matrix on the left side of Equation (2.5) as

$$(E^H dAE) = (\det E)^{2m} (dA) = (dA).$$

On the right side of Equation (2.5), the exterior product of the diagonal elements is given by

$$(d\Lambda) = \bigwedge_{k=1}^{m} d\lambda_k. \tag{2.6}$$

Note that the diagonal elements of $(E^H dE\Lambda - \Lambda E^H dE)$ are zeros.

Now we consider the upper diagonal elements of $(E^H dE\Lambda - \Lambda E^H dE)$. First note that the matrix $\mathrm{Re}\left(E^H dE\right)$ is skew-symmetric and can be written as

$$\mathrm{Re}\left(E^H dE\right) = \begin{bmatrix} 0 & -\mathrm{Re}\left(e_2^H de_1\right) & \dots & \dots & -\mathrm{Re}\left(e_m^H de_1\right) \\ \mathrm{Re}\left(e_2^H de_1\right) & 0 & \dots & \dots & -\mathrm{Re}\left(e_m^H de_2\right) \\ \mathrm{Re}\left(e_3^H de_1\right) & \mathrm{Re}\left(e_3^H de_2\right) & 0 & \dots & -\mathrm{Re}\left(e_m^H de_3\right) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathrm{Re}\left(e_m^H de_1\right) & \mathrm{Re}\left(e_m^H de_2\right) & \dots & \dots & 0 \end{bmatrix}.$$

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

$$\mathrm{Re}\left(E^H dE\right)\Lambda - \Lambda\,\mathrm{Re}\left(E^H dE\right)$$

is given by

$$\bigwedge_{k<l}^{m} \mathrm{Re}(e_l^H de_k) \prod_{k<l}^{m} (\lambda_k - \lambda_l). \tag{2.7}$$

Similarly, $\mathrm{Im}\left(E^H dE\right)$ is a symmetric matrix and can be written as

$$\mathrm{Im}\left(E^H dE\right) = \begin{bmatrix} \mathrm{Im}\left(e_1^H de_1\right) & -\mathrm{Im}\left(e_2^H de_1\right) & \dots & \dots & -\mathrm{Im}\left(e_m^H de_1\right) \\ -\mathrm{Im}\left(e_2^H de_1\right) & \mathrm{Im}\left(e_2^H de_2\right) & \dots & \dots & -\mathrm{Im}\left(e_m^H de_2\right) \\ -\mathrm{Im}\left(e_3^H de_1\right) & -\mathrm{Im}\left(e_3^H de_2\right) & \dots & \dots & -\mathrm{Im}\left(e_m^H de_3\right) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\mathrm{Im}\left(e_m^H de_1\right) & -\mathrm{Im}\left(e_m^H de_2\right) & \dots & \dots & \mathrm{Im}\left(e_m^H de_m\right) \end{bmatrix}$$

and the exterior product of the upper diagonal elements of

$$\mathrm{Im}\left(E^H dE\right)\Lambda - \Lambda\,\mathrm{Im}\left(E^H dE\right)$$

is given by

$$\bigwedge_{k<l}^{m} \mathrm{Im}(e_l^H de_k) \prod_{k<l}^{m} (\lambda_k - \lambda_l). \tag{2.8}$$

Therefore, the exterior product of the elements of the right side of Equation (2.5) is obtained by multiplying Equations (2.6), (2.7), and (2.8), i.e.,

$$(dA) = \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 (d\Lambda)(E^H dE). \tag{2.9}$$

The proof is complete. □

The Jacobian for the change of variables in the orthonormal-triangular factorization is given by the following proposition. This result first appeared in [12] without proof. Here we derive this result based on the exterior product.

Proposition 2.4 *Let A be an $(n \times m)$ $(n \geq m)$ full rank complex matrix. If $A = E_1 T$, where E_1 is an $(n \times m)$ matrix with orthonormal columns and T is an $(m \times m)$ upper triangular matrix with real and positive diagonal elements, then*

$$(dA) = \prod_{k=1}^{m} t_{kk}^{2n-2k+1} (dT)(E_1^H dE_1), \tag{2.10}$$

where

$$(E_1^H dE_1) \equiv \bigwedge_{k=1}^{m} \bigwedge_{l=k}^{n} e_l^H de_k$$

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

Proof. First, note that $dA = dE_1 T + E_1 dT$ and $E_1^H E_1 = I$ and $E_2^H E_1 = 0$. Therefore, we have

$$E^H dA = \begin{bmatrix} E_1^H \\ E_2^H \end{bmatrix} dA = \begin{bmatrix} E_1^H dE_1 T + dT \\ E_2^H dE_1 T \end{bmatrix}. \tag{2.11}$$

The exterior product on the left side of Equation (2.11) is equal to

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

The lth row of $E_2^H dE_1 T$ is

$$\left(e_l^H de_1, \ldots, e_l^H de_m\right) T, \quad m+1 \leq l \leq n,$$

and the exterior product of its elements is equal to

$$(\det T)^2 \bigwedge_{k=1}^{m} e_l^H de_k = \left(\prod_{k=1}^{m} t_{kk}^2\right) \bigwedge_{k=1}^{m} e_l^H de_k.$$

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

$$\bigwedge_{l=m+1}^{n} \left[\left(\prod_{k=1}^{m} t_{kk}^2\right) \bigwedge_{k=1}^{m} e_l^H de_k\right] = \left(\prod_{k=1}^{m} t_{kk}^{2n-2m}\right) \bigwedge_{l=m+1}^{n} \bigwedge_{k=1}^{m} e_l^H de_k. \tag{2.12}$$

Now we consider $E_1^H dE_1 T$. Since E_1 is unitary we have

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

This implies that the real and the imaginary parts of $E_1^H dE_1$ are skew-symmetric and symmetric, respectively. Therefore, the real part of the matrix $E_1^H dE_1$ can be written as

$$\operatorname{Re}\left(E_1^H dE_1\right) = \begin{bmatrix} 0 & -\operatorname{Re}\left(e_2^H de_1\right) & \ldots & \ldots & -\operatorname{Re}\left(e_m^H de_1\right) \\ \operatorname{Re}\left(e_2^H de_1\right) & 0 & \ldots & \ldots & -\operatorname{Re}\left(e_m^H de_2\right) \\ \operatorname{Re}\left(e_3^H de_1\right) & \operatorname{Re}\left(e_3^H de_2\right) & 0 & \ldots & -\operatorname{Re}\left(e_m^H de_3\right) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \operatorname{Re}\left(e_m^H de_1\right) & \operatorname{Re}\left(e_m^H de_2\right) & \ldots & \ldots & 0 \end{bmatrix}. \tag{2.13}$$

Multiplying Equation (2.13) by T and retaining the terms that contribute to the exterior product, we get

$$\begin{bmatrix} 0 & * & \dots & * & * \\ \mathrm{Re}\left(e_2^H de_1\right) t_{11} & * & \dots & * & * \\ \mathrm{Re}\left(e_3^H de_1\right) t_{11} & \mathrm{Re}\left(e_3^H de_2\right) t_{22} + * & \dots & * & * \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathrm{Re}\left(e_m^H de_1\right) t_{11} & \mathrm{Re}\left(e_m^H de_2\right) t_{22} + * & \dots & \mathrm{Re}\left(e_m^H de_{m-1}\right) t_{m-1m-1} + * & * \end{bmatrix}.$$

The exterior product of the subdiagonal elements of $\mathrm{Re}\left(E_1^H dE_1\right) T$ is given by

$$t_{11}^{m-1} \bigwedge_{l=2}^{m} \mathrm{Re}\left(e_l^H de_1\right) \wedge t_{22}^{m-2} \bigwedge_{l=3}^{m} \mathrm{Re}\left(e_l^H de_2\right) \cdots \wedge t_{m-1,m-1} \mathrm{Re}\left(e_m^H de_{m-1}\right)$$
$$= \left(\prod_{k=1}^{m} t_{kk}^{m-k}\right) \bigwedge_{k=1}^{m} \bigwedge_{l=k+1}^{m} \mathrm{Re}\left(e_l^H de_k\right). \tag{2.14}$$

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

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

and the exterior product of the diagonal and the subdiagonal elements of $\mathrm{Im}\left(E^H dE\right) T$ is given by

$$\left(\prod_{k=1}^{m} t_{kk}^{m-k+1}\right) \bigwedge_{k=1}^{m} \bigwedge_{l=k}^{m} \mathrm{Im}\left(e_l^H de_k\right). \tag{2.15}$$

It is clear that the upper diagonal elements of $E_1^H dE_1 T$ are contributing nothing to the exterior product because they have already appeared in (2.14) and (2.15). Finally, the exterior product of dT is given by

$$(dT) = \bigwedge_{k\leq l}^{m} dt_{kl}. \tag{2.16}$$

Therefore, the exterior product of the elements of the right side of Equation (2.11) is obtained by multiplying Equations (2.12), (2.14), (2.15) and (2.16), i.e.,

$$(dA) = \prod_{k=1}^{m} t_{kk}^{2n-2k+1}(dT)(E^H dE).$$

The proof is complete. □

Note that the corresponding real random matrix results for Propositions 2.2 – 2.4 are given in [36].

2.2 The complex multivariate Gamma function

Definition 2.1 *The complex multivariate gamma function is defined by*[1] *[24]*

$$\mathcal{C}\Gamma_m(a) = \int_{A^H=A>0} \operatorname{etr}(-A)(\det A)^{a-m}(dA), \quad \operatorname{Re}(a) > (m-1), \tag{2.17}$$

where $\operatorname{etr}(\cdot) \equiv e^{\operatorname{tr}(\cdot)} \equiv \exp\operatorname{tr}(\cdot)$*. The integral is over the space of Hermitian positive definite* $m \times m$ *matrices.*

Proposition 2.5 *If* Σ *is a Hermitian* $m \times m$ *matrix with* $Re(\Sigma) > 0$ *then*

$$\int_{A^H=A>0} \operatorname{etr}(-\Sigma^{-1}A)(\det A)^{a-m}(dA) = (\det \Sigma)^a \mathcal{C}\Gamma_m(a) \tag{2.18}$$

where $\operatorname{Re}(a) > (m-1)$.

[1] In the literature the notation $\widetilde{\Gamma}_m(a)$ is also used to denote the complex Gamma function.

Proof. Make the change of variables $A = \Sigma^{1/2} V \Sigma^{1/2}$. Then $(dA) = (\det \Sigma)^m (dV)$. Therefore, substituting A and (dA) in Equation (2.18) we obtain

$$(\det \Sigma)^a \int_{V^H = V > 0} \operatorname{etr}(-V)(\det V)^{a-m}(dV).$$

The result follows from Definition 2.1. □

The following proposition shows that the complex multivariate gamma function can be expressed as a product of scalar real gamma functions [24].

Proposition 2.6 *The complex multivariate gamma function $C\Gamma_m(a)$ can be written as*

$$C\Gamma_m(a) = \pi^{m(m-1)/2} \prod_{k=1}^{m} \Gamma(a-k+1), \quad \operatorname{Re}(a) > (m-1).$$

Proof. Let $A = T^H T$ where T is an upper triangular matrix with real and positive diagonal elements. We have

$$\operatorname{tr} A = \operatorname{tr} T^H T = \sum_{k \le l} |t_{kl}|^2 \quad \text{and} \quad \det A = \det T^H T = (\det T)^2 = \prod_{k=1}^{m} t_{kk}^2.$$

From Proposition 2.2, (dA) can be written as

$$(dA) = 2^m \prod_{k=1}^{m} t_{kk}^{2m-2k+1} \bigwedge_{k \le l}^{m} t_{kl}.$$

Therefore

$$\begin{aligned} C\Gamma_m(a) &= 2^m \int \cdots \int \exp\left(-\sum_{k \le l}^{m} |t_{kl}|^2\right) \prod_{k=1}^{m} t_{kk}^{2a-2k+1} \bigwedge_{k \le l}^{m} dt_{kl} \\ &= \prod_{k<l} \left[\int_{-\infty}^{\infty} \exp(-|t_{kl}|^2) dt_{kl}\right] \prod_{k=1}^{m} \left[\int_0^{\infty} \exp(-t_{kk}^2)(t_{kk}^2)^{a-k} dt_{kk}^2\right] \\ &= \pi^{m(m-1)/2} \prod_{k=1}^{m} \Gamma(a-k+1). \end{aligned} \tag{2.19}$$

The proof is complete. □

2.3 The Stiefel manifold and Haar measure

The set of all $n \times m$ $(n \geq m)$ matrices E with orthonormal columns is called the *Stiefel manifold*, denoted by $\mathcal{CV}_{m,n}$. Thus,

$$\mathcal{CV}_{m,n} = \{E(n \times m); E^H E = I_m\}.$$

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

Proposition 2.7 *The volume of the Stiefel manifold $\mathcal{CV}_{m,n}$ is given by*

$$\begin{aligned} \mathrm{Vol}(\mathcal{CV}_{m,n}) &= \int_{\mathcal{CV}_{m,n}} (E^H dE) \\ &= \frac{2^m \pi^{mn}}{\mathcal{C}\Gamma_m(n)}, \end{aligned} \tag{2.20}$$

where, by Proposition 2.6, the complex multivariate gamma function is

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

Proof. Let Z be an $n \times m$ $(n \geq m)$ complex random matrix with $\mathbf{Z} \sim \mathcal{CN}(\mathbf{0}, I_n \otimes I_m)$. The density function is given by

$$\pi^{-nm} \operatorname{etr}\left(-Z^H Z\right),$$

and

$$\pi^{nm} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \operatorname{etr}\left(-Z^H Z\right)(dZ). \tag{2.21}$$

Suppose $Z = ET$, where $E \in \mathcal{CV}_{m,n}$, and T is an upper triangular matrix with real and positive diagonal elements. Then

$$\operatorname{tr} Z^H Z = \operatorname{tr} T^H T = \sum_{k \leq l}^{m} |t_{kl}|^2.$$

From Proposition 2.4, (dZ) can be written as

$$(dZ) = \prod_{k=1}^{m} t_{kk}^{2n-2k+1} (dT)(E^H dE).$$

Equation (2.21) can be written as

$$\begin{aligned}
\pi^{nm} &= \int \cdots \int \exp\left(-\sum_{k \leq l}^{m} |t_{kl}|^2\right) \prod_{k=1}^{m} t_{kk}^{2n-2k+1} \bigwedge_{k \leq l}^{m} dt_{kl} \int_{\mathcal{CV}_{m,n}} (E^H dE) \\
&= \prod_{k<l}^{m} \left[\int_{-\infty}^{\infty} \exp\left(-|t_{kl}|^2\right) dt_{kl}\right] \prod_{k=1}^{m} \left[\int_{0}^{\infty} \exp\left(-t_{kk}^2\right) t_{kk}^{2n-2k+1} dt_{kk}\right] \int_{\mathcal{CV}_{m,n}} (E^H dE) \\
&= \prod_{k<l}^{m} [\pi] \prod_{k=1}^{m} \left[\frac{1}{2} \int_{0}^{\infty} \exp\left(-t_{kk}^2\right) (t_{kk}^2)^{n-k} dt_{kk}^2\right] \int_{\mathcal{CV}_{m,n}} (E^H dE) \\
&= 2^{-m} \pi^{m(m-1)/2} \prod_{k=1}^{m} \Gamma(n-k+1) \int_{\mathcal{CV}_{m,n}} (E^H dE) \\
&= 2^{-m} \mathcal{C}\Gamma_m(n) \int_{\mathcal{CV}_{m,n}} (E^H dE). \qquad (2.22)
\end{aligned}$$

Therefore,

$$\int_{\mathcal{CV}_{m,n}} (E^H dE) = \frac{2^m \pi^{mn}}{\mathcal{C}\Gamma_m(n)}.$$

The proof is complete. □

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

$$\mathcal{CV}_{m,m} \equiv U(m) = \{E(m \times m); E^H E = I_m\},$$

the set of unitary $m \times m$ matrices.

Corollary 2.1 *The volume of $U(m)$ is given by*

$$\mathrm{Vol}\,[U(m)] = \int_{U(m)} (E^H dE) = \frac{2^m \pi^{m^2}}{\mathcal{C}\Gamma_m(m)}. \tag{2.23}$$

□

The measure μ on the unitary manifold $U(m)$ is defined as

$$\mu(\mathcal{D}) = \int_{\mathcal{D}} (E^H dE), \quad \mathcal{D} \subset U(m). \tag{2.24}$$

This represents the volume of the region $\mathcal{D}$ on the unitary manifold. It can be shown that μ is an invariant measure on $U(m)$, i.e.,

$$\mu(E\mathcal{D}) = \mu(\mathcal{D}E) = \mu(\mathcal{D}) \quad \forall\ E \in U(m).$$

It is also called the Haar measure on $U(m)$ in honor of Haar, who proved the existence of an invariant measure in any locally compact topological group. Note that, this measure is unnormalized or so called Lebesgue measure. If we normalize it by dividing by $\mathrm{Vol}\,[U(m)]$, then we get a probability measure, or so-called Haar invariant distribution on $U(m)$. In other words, the differential form

$$(dE) \triangleq \frac{1}{\mathrm{Vol}[U(m)]}(E^H dE) = \frac{\mathcal{C}\Gamma_m(m)}{2^m \pi^{m^2}}(E^H dE) \tag{2.25}$$

has the property that

$$\int_{U(m)} (dE) = 1,$$

and it represents the normalized Haar invariant probability measure on $U(m)$. Similarly, we can obtain the Haar invariant distribution on the Stiefel manifold $\mathcal{CV}_{m,n}$.

Chapter 3

Zonal Polynomials and Hypergeometric Functions

The complex hypergeometric functions[1] ${}_pF_q$ of matrix argument which occur in complex multivariate distributions are studied in Section 3.2 by their expansions in complex zonal polynomials, as defined in Section 3.1. The algorithm for computing these polynomials is developed in Section 3.4. Section 3.3 describes invariant polynomials. The theory of hypergeometric functions of matrix argument is due to Herz [19], and the theory of zonal polynomials is due to James [20], [21], [22], [23], [24] and Constantine [8], [9].

3.1 Complex zonal polynomials

First, we define the complex multivariate hypergeometric coefficients $[a]_\kappa$ which frequently occur in integrals involving zonal polynomials. Let $\kappa = (k_1, \ldots, k_m)$ be a partition of the integer k with $k_1 \geq \cdots \geq k_m \geq 0$ and $k = k_1 + \cdots + k_m$. Then [27]

$$[a]_\kappa = \prod_{i=1}^{m} (a - i + 1)_{k_i} = \frac{C\Gamma_m(a, \kappa)}{C\Gamma_m(a)},$$

[1]Note that in the literature the real and the complex hypergeometric functions are denoted by ${}_pF_q$ and ${}_p\widetilde{F}_q$, respectively. However, we use ${}_pF_q$ for the complex case because we are not considering the real case in this work.

where $(a)_k = a(a+1)\cdots(a+k-1)$, and

$$\mathcal{C}\Gamma_m(a,\kappa) = \pi^{m(m-1)/2}\prod_{i=1}^{m}\Gamma(a+k_i-i+1), \quad \mathrm{Re}(a) > (m-1).$$

Moreover, we also have

$$\mathcal{C}\Gamma_m(a,-\kappa) = \pi^{m(m-1)/2}\prod_{i=1}^{m}\Gamma(a-m-k_i+i), \quad \mathrm{Re}(a) > (m-1)+k_1.$$

The complex zonal polynomial[2] of a complex matrix $X \in \mathbb{C}^{m\times m}$ is defined by [24]

$$C_\kappa(X) = \chi_{[\kappa]}(1)\chi_{[\kappa]}(X), \tag{3.1}$$

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

$$\chi_{[\kappa]}(1) = k!\,\frac{\prod_{i<j}^{m}(k_i-k_j-i+j)}{\prod_{i=1}^{m}(k_i+m-i)!}, \tag{3.2}$$

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

$$\chi_{[\kappa]}(X) = \frac{\det\left[\left(\lambda_i^{k_j+m-j}\right)\right]}{\det\left[\left(\lambda_i^{m-j}\right)\right]}. \tag{3.3}$$

The following basic properties are given in [24]:

$$(\mathrm{tr}\,X)^k = \sum_\kappa C_\kappa(X)$$

and

$$\int_{U(m)} C_\kappa(AXBX^H)(dX) = \frac{C_\kappa(A)C_\kappa(B)}{C_\kappa(I_m)}, \tag{3.4}$$

[2] Note that in the literature the real and the complex zonal polynomials are denoted by $C_\kappa(X)$ and $\widetilde{C}_\kappa(X)$, respectively. However, we use $C_\kappa(X)$ for the complex case because we are not considering the real case in this work.

where (dX) is the invariant measure on the unitary group $U(m)$ normalized to make the total measure unity and

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

Note that the partition κ of k has r nonzero parts. The reproductive property of the zonal polynomial is

$$\frac{1}{\mathcal{C}\Gamma_m(a)} \int_{X^H = X > 0} \operatorname{etr}(-X)(\det X)^{a-m} C_\kappa(YX)(dX) = [a]_\kappa C_\kappa(Y), \quad \operatorname{Re}(a) > (m-1).$$

Alternatively, we can write

$$\int_{X^H = X > 0} \operatorname{etr}(-X)(\det X)^{a-m} C_\kappa(YX)(dX) = \mathcal{C}\Gamma_m(a,\kappa) C_\kappa(Y), \quad \operatorname{Re}(a) > (m-1). \tag{3.5}$$

Moreover, we have

$$\int_{X^H = X > 0} \operatorname{etr}(-X)(\det X)^{a-m} C_\kappa(YX^{-1})(dX) = \mathcal{C}\Gamma_m(a,-\kappa) C_\kappa(Y), \tag{3.6}$$

where $\operatorname{Re}(a) > (m-1) + k_1$. Equations (3.5) and (3.6) are special cases of Equations (3.9) and (3.10). The following lemma is due to Muirhead [36].

Lemma 3.1 *Let* $Y = \operatorname{diag}(y_1, \ldots, y_m)$ *and* $X = (x_{ij})$ *be an* $m \times m$ *positive definite matrix. Then*

$$C_\kappa(XY) = d_\kappa y_1^{k_1} \cdots y_m^{k_m} x_{11}^{k_1-k_2} \det \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}^{k_2-k_3} \cdots \det X^{k_m} + \textit{terms of lower weight in the } y'\textit{s}, \tag{3.7}$$

and

$$C_\kappa(X^{-1}Y) = d_\kappa y_1^{k_1} \cdots y_m^{k_m} x_{11}^{-(k_1-k_2)} \det \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}^{-(k_2-k_3)} \cdots \det X^{-k_m} + \textit{terms of lower weight in the } y'\textit{s}, \tag{3.8}$$

where $\kappa = (k_1, \ldots, k_m)$ *and* d_κ *is the coefficient of the term of highest weight in* $C_\kappa(\cdot)$.

Proof. See [36]. □

Proposition 3.1 *If* Y *and* Z *are* $m \times m$ *Hermitian matrices with* $\mathrm{Re}(Z) > 0$, *then*

$$\int_{X^H = X > 0} \mathrm{etr}(-XZ)(\det X)^{a-m} C_\kappa(XY)(dX) = \mathcal{C}\Gamma_m(a, \kappa)(\det Z)^{-a} C_\kappa(YZ^{-1}), \quad (3.9)$$

where $\mathrm{Re}(a) > (m-1)$ *and*

$$\int_{X^H = X > 0} \mathrm{etr}(-XZ)(\det X)^{a-m} C_\kappa(X^{-1}Y)(dX) = \mathcal{C}\Gamma_m(a, -\kappa)(\det Z)^{-a} C_\kappa(YZ), \quad (3.10)$$

where $\mathrm{Re}(a) > (m-1) + k_1$.

Proof. If $Z = I$, then Equations (3.9) and (3.10) reduce to Equations (3.5) and (3.6), respectively. Let $f(Y)$ denote the left side of Equation (3.5). Then

$$f(EYE^H) = \int_{X^H = X > 0} \mathrm{etr}(-X)(\det X)^{a-m} C_\kappa(XEYE^H)(dX), \quad \forall\, E \in U(m). \quad (3.11)$$

If $X = EWE^H$, then $(dX) = (dW)$ and $f(EYE^H) = f(Y)$. This implies that f is a symmetric function of Y. Moreover, (dE) is the invariant measure on the unitary group $U(m)$ normalized to make the total measure unity, therefore we have

$$\begin{aligned} f(Y) &= \int_{U(m)} f(Y)(dE) \\ &= \int_{X^H = X > 0} \mathrm{etr}(-X)(\det X)^{a-m} \int_{U(m)} C_\kappa(XEYE^H)(dE)(dX) \\ &= \int_{X^H = X > 0} \mathrm{etr}(-X)(\det X)^{a-m} \frac{C_\kappa(X) C_\kappa(Y)}{C_\kappa(I_m)}(dX) \\ &= \frac{f(I_m) C_\kappa(Y)}{C_\kappa(I_m)}. \end{aligned} \quad (3.12)$$

On the one hand, we have [36]

$$f(Y) = \frac{f(I_m)}{C_\kappa(I_m)} d_\kappa y_1^{k_1} \cdots y_m^{k_m} + \text{terms of lower weight.} \tag{3.13}$$

On the other hand, using Lemma 3.1, we have

$$\begin{aligned} f(Y) &= \int_{X^H=X>0} \operatorname{etr}(-X)(\det X)^{a-m} C_\kappa(XY)(dX) \\ &= d_\kappa y_1^{k_1} \cdots y_m^{k_m} \int_{X^H=X>0} \operatorname{etr}(-X)(\det X)^{a-m} x_{11}^{k_1-k_2} \det \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}^{k_2-k_3} \\ &\quad \cdots \det X^{k_m}(dX) + \text{terms of lower weight.} \end{aligned} \tag{3.14}$$

Since the evaluation of this integral is similar to Proposition 2.6 or 2.7 in Chapter 2 (i.e., substitute $X = T^H T$), therefore, we state the result

$$f(Y) = d_\kappa y_1^{k_1} \cdots y_m^{k_m} \mathcal{C}\Gamma_m(a, \kappa) + \text{terms of lower weight.} \tag{3.15}$$

Equating the coefficients of $y_1^{k_1} \cdots y_m^{k_m}$ in Equations (3.13) and (3.15) and using Equation (3.12), we obtain

$$f(Y) = \mathcal{C}\Gamma_m(a, \kappa) C_\kappa(Y).$$

The rest of the proof for general Z is similar to Proposition 2.5 in Chapter 2 (i.e., substitute $X = Z^{-1/2} V Z^{-1/2}$). Similarly, we can prove the second part. □

The following corollary follows from the second part of Proposition 3.1 by letting $Y = I$.

Corollary 3.1 *Let Z be an $m \times m$ Hermitian matrix with* $\operatorname{Re}(Z) > 0$. *Then*

$$\int_{X^H=X>0} \operatorname{etr}(-XZ)(\det X)^{a-m} C_\kappa(X^{-1})(dX) = \frac{(-1)^k \mathcal{C}\Gamma_m(a)}{[-a+m]_\kappa} (\det Z)^{-a} C_\kappa(Z)$$

for $\operatorname{Re}(a) > k_1 + (m-1)$, *where* $\kappa = (k_1, \ldots, k_m)$.

Proof. The result follows by noting that

$$C\Gamma_m(a,-\kappa) = \frac{(-1)^k C\Gamma_m(a)}{[-a+m]_\kappa}.$$

□

Proposition 3.2 *Let Y is an $m \times m$ symmetric matrix. Then the following are true:*

$$\int_{0<X<I_m} (\det X)^{a-m} \det(I_m - X)^{b-m} C_\kappa(XY)(dX) = \frac{C\Gamma_m(a,\kappa)C\Gamma_m(b)}{C\Gamma_m(a+b,\kappa)} C_\kappa(Y) \quad (3.16)$$

for $\mathrm{Re}(a) > (m-1)$, $\mathrm{Re}(b) > (m-1)$ *and*

$$\int_{0<X<I_m} (\det X)^{a-m} \det(I_m - X)^{b-m} C_\kappa(X^{-1}Y)(dX) = \frac{C\Gamma_m(a,-\kappa)C\Gamma_m(b)}{C\Gamma_m(a+b,-\kappa)} C_\kappa(Y) \quad (3.17)$$

for $\mathrm{Re}(a) > (m-1) + k_1$ *and* $\mathrm{Re}(b) > (m-1)$.

Proof. As in the proof of Proposition 3.1, if $f(Y)$ denotes the left side of Equation (3.16) then we have

$$f(Y) = f(EYE^H)\ \forall\ E \in U(m) \quad \text{and} \quad f(Y)C_\kappa(I_m) = f(I_m)C_\kappa(Y).$$

By letting $Z = I$ and $Y = I$ in Equation (3.9) and then multiplying with $f(I_m)$ we obtain the following

$$\begin{aligned} C\Gamma_m(a+b,\kappa)f(I_m) &= \int_{W^H=W>0} \mathrm{etr}(-W)(\det W)^{a+b-m} f(W)(dW) \\ &= \int_{W^H=W>0} \mathrm{etr}(-W)(\det W)^{a+b-m} \int_{0<X<I_m} (\det X)^{a-m} \\ &\quad \cdot \det(I_m - X)^{b-m} C_\kappa(WX)(dX)(dW). \end{aligned} \quad (3.18)$$

Let $X = W^{-1/2}UW^{-1/2}$. Then $(dX) = (\det W)^{-m}(dU)$ and

$$C\Gamma_m(a+b,\kappa)f(I_m) = \int_{W^H=W>0} \mathrm{etr}(-W) \int_{0<U<W} (\det U)^{a-m}$$

$$
\begin{aligned}
&\cdot \det(W-U)^{b-m} C_\kappa(U)(dU)(dW) \\
= & \int_{U^H=U>0} \operatorname{etr}(-U)(\det U)^{a-m} C_\kappa(U)(dU) \\
& \cdot \int_{V^H=V>0} \operatorname{etr}(-V)(\det V)^{b-m}(dV) \quad (\text{letting } V = W-U) \\
= & \; \mathcal{C}\Gamma_m(a,\kappa) C_\kappa(I_m) \mathcal{C}\Gamma_m(b). \qquad (3.19)
\end{aligned}
$$

This completes the proof, i.e.,

$$
f(I_m) = \frac{\mathcal{C}\Gamma_m(a,\kappa)\mathcal{C}\Gamma_m(b)}{\mathcal{C}\Gamma_m(a+b,\kappa)} C_\kappa(I_m).
$$

Similarly, we can prove the second part. □

If $b = m$ then we have the following corollary.

Corollary 3.2 *If Y is an $m \times m$ Hermitian matrix, then*

$$
\int_{0<X<I_m} (\det X)^{a-m} C_\kappa(XY)(dX) = \frac{\mathcal{C}\Gamma_m(a)\mathcal{C}\Gamma_m(m)}{\mathcal{C}\Gamma_m(a+m)} \frac{[a]_\kappa}{[a+m]_\kappa} C_\kappa(Y)
$$

for $\mathrm{Re}(a) > (m-1)$.

Proof. The result follows by noting that

$$
\mathcal{C}\Gamma_m(a,\kappa) = [a]_\kappa \mathcal{C}\Gamma_m(a).
$$

□

3.2 Complex hypergeometric functions

The probability distributions of random matrices are often derived in terms of hypergeometric functions of matrix arguments. The following definition of hypergeometric functions with a single matrix argument is due to Constantine [8].

Definition 3.1 *The hypergeometric function of one complex matrix argument is defined as*

$$ {}_pF_q(a_1, \ldots, a_p; b_1, \ldots, b_q; X) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{[a_1]_\kappa \cdots [a_p]_\kappa}{[b_1]_\kappa \cdots [b_q]_\kappa} \frac{C_\kappa(X)}{k!}, $$

where $X \in \mathbb{C}^{m \times m}$ *and* $\{a_i\}_{i=1}^p$ *and* $\{b_i\}_{i=1}^q$ *are arbitrary complex numbers. Note that* $\sum_\kappa$ *denotes summation over all partitions* κ *of* k.

The conditions for the convergence of these hypergeometric functions are given in Section 3.4. Special cases are

$$ {}_0F_0(X) = \mathrm{etr}(X), $$

$$ {}_1F_0(a; X) = \det(I - X)^{-a}, $$

and

$$ {}_0F_1(n; ZZ^H) = \int_{U(n)} \mathrm{etr}(ZE + \overline{ZE})(dE), $$

where Z is an $m \times n$ complex matrix with $m \leq n$ and $\overline{ZE}$ denotes the complex conjugate of ZE.

Definition 3.2 *The hypergeometric function whose arguments are two complex matrices is defined by*

$$ {}_pF_q^{(m)}(a_1, \ldots, a_p; b_1, \ldots, b_q; X, Y) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{[a_1]_\kappa \cdots [a_p]_\kappa}{[b_1]_\kappa \cdots [b_q]_\kappa} \frac{C_\kappa(X) C_\kappa(Y)}{k! C_\kappa(I_m)}, $$

where $X,\ Y \in \mathbb{C}^{m \times m}$

The splitting formula is

$$ \int_{U(m)} {}_pF_q(AEBE^H)(dE) = {}_pF_q(A, B). $$

Proposition 3.3 *Let* X *be an* $m \times m$ *positive definite matrix and* Y *be an* $m \times m$

symmetric matrix. Then

$$\int_{U(m)} {}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;XEYE^H)(dE) = {}_pF_q^{(m)}(a_1,\ldots,a_p;b_1,\ldots,b_q;X,Y).$$

Proof. The result follows by expanding the integrand and integrating term by term using Equation (3.4). □

Proposition 3.4 *Let Z be an $m\times m$ complex symmetric matrix with* $\mathrm{Re}(Z)>0$, *Y be an $m\times m$ complex symmetric matrix and T is an $m\times m$ symmetric matrix. Then*

$$\begin{aligned}\int_{X^H=X>0} &\mathrm{etr}(-XZ)(\det X)^{a-m}{}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;XY)(dX)\\ &= \mathcal{C}\Gamma_m(a)(\det Z)^{-a}{}_{p+1}F_q(a_1,\ldots,a_p,a;b_1,\ldots,b_q;Z^{-1}Y),\end{aligned} \tag{3.20}$$

and

$$\begin{aligned}\int_{X^H=X>0} &\mathrm{etr}(-XZ)(\det X)^{a-m}{}_pF_q^{(m)}(a_1,\ldots,a_p;b_1,\ldots,b_q;XY,T)(dX)\\ &= \mathcal{C}\Gamma_m(a)(\det Z)^{-a}{}_{p+1}F_q^{(m)}(a_1,\ldots,a_p,a;b_1,\ldots,b_q;Z^{-1}Y,T),\end{aligned} \tag{3.21}$$

for $p<q$, $\mathrm{Re}(a)>(m-1)$*; or $p=q$,* $\mathrm{Re}(a)>(m-1)$, $\|Z^{-1}\|<1$ ($\|Y\|\le 1$).

Proof. Expanding the functions ${}_pF_q$ in the integrand and integrating term by term using Proposition 3.1 will lead to the desired results. □

Proposition 3.5 *If Y is an $m\times m$ symmetric matrix, then*

$$\begin{aligned}\int_{0<X<I_m} &(\det X)^{a-m}\det(I_m-X)^{b-m}{}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;YX)(dX)\\ &= \frac{\mathcal{C}\Gamma_m(a)\mathcal{C}\Gamma_m(b)}{\mathcal{C}\Gamma_m(a+b)}{}_{p+1}F_{q+1}(a_1,\ldots,a_p,a;b_1,\ldots,b_q,a+b;Y).\end{aligned} \tag{3.22}$$

Proof. The result follows by expanding ${}_pF_q$ in the integrand and integrating term by term using Proposition 3.2. □

Corollary 3.3 *The integral representation of the function* ${}_1F_1$ *is given by*

$$ {}_1F_1(a;c;Y) = \frac{\mathcal{C}\Gamma_m(c)}{\mathcal{C}\Gamma_m(a)\mathcal{C}\Gamma_m(c-a)} \int_0^{I_m} (\det X)^{a-m} \det(I_m - X)^{c-a-m} \operatorname{etr}(YX)(dX) \tag{3.23} $$

for all symmetric Y*. Note that* $\mathrm{Re}(a) > (m-1)$*,* $\mathrm{Re}(b) > (m-1)$*, and* $\mathrm{Re}(b-a) > (m-1)$*. Moreover, the function* ${}_2F_1$ *has the integral representation*

$$ \begin{aligned} {}_2F_1(a_1,a_2;b_1;Y) &= \frac{\mathcal{C}\Gamma_m(b_1)}{\mathcal{C}\Gamma_m(a_1)\mathcal{C}\Gamma_m(b_1-a_1)} \int_0^{I_m} \det(I-YX)^{-a_2} (\det X)^{a_1-m} \\ &\quad \cdot \det(I-X)^{b_1-a_1-m}(dX) \end{aligned} \tag{3.24} $$

for $\mathrm{Re}(X) < I$*,* $\mathrm{Re}(a) > (m-1)$ *and* $\mathrm{Re}(b-a_1) > (m-1)$*.*

Proof. Letting $p = q = 0$ and $b = c - a$ in Equation (3.22), we obtain

$$ {}_1F_1(a;c;Y) = \frac{\mathcal{C}\Gamma_m(c)}{\mathcal{C}\Gamma_m(a)\mathcal{C}\Gamma_m(c-a)} \int_0^{I_m} (\det X)^{a-m} \det(I_m - X)^{c-a-m} {}_0F_0(YX)(dX), $$

with $\mathrm{Re}(a) > (m-1)$ and $\mathrm{Re}(c-a) > (m-1)$. Equation (3.23) follows by noting that ${}_0F_0(YX) = \operatorname{etr}(YX)$. To prove Equation (3.24), letting $p=1$, $q=0$, $a_1 = a_2$, $a = a_1$ and $b = b_1 - a_1$ in Equation (3.22), we obtain

$$ \begin{aligned} {}_2F_1(a_1,a_2;b_1;Y) &= \frac{\mathcal{C}\Gamma_m(b_1)}{\mathcal{C}\Gamma_m(a_1)\mathcal{C}\Gamma_m(b_1-a_1)} \int_0^{I_m} \det(I-X)^{b_1-a_1-m} (\det X)^{a_1-m} \\ &\quad \cdot {}_1F_0(a_2;YX)(dX). \end{aligned} \tag{3.25} $$

The result follows by noting that ${}_1F_0(a_2;YX) = \det(I-YX)^{-a_2}$. □

A common method for evaluating integrals uses the multidimensional Laplace transform, which is defined next.

Definition 3.3 *If* $f(X)$ *is a function of a positive definite Hermitian matrix* X*, then the Laplace transform of* $f(X)$ *is defined as*

$$ g(Z) = \int_{X^H = X > 0} \operatorname{etr}(-XZ) f(X)(dX), \tag{3.26} $$

where $Z = U + iV$ is a complex symmetric matrix, and U and V are real matrices. It is assumed that the integral is absolutely convergent in the right half-plane $\mathrm{Re}(Z) = U > U_0$ *for some positive definite U_0.*

The hypergeometric functions ${}_1F_1$ and ${}_2F_1$ satisfy the following relations, which are easily proved by the uniqueness of the Laplace transforms of both sides.

$$\mathrm{etr}(-X){}_1F_1(a; b; X) = {}_1F_1(b-a; b; -X) \tag{3.27}$$

and

$$\begin{aligned} {}_2F_1(a_1, a_2; b; X) &= \det(I-A)^{a_2}{}_2F_1(b-a_1; b; -X(I-X)^{-1}) \\ &= \det(I-A)^{b-a_1-a_2}{}_2F_1(b-a_1, b-a_2; b; X). \end{aligned} \tag{3.28}$$

Proposition 3.6 *Let X be an $m \times n$ complex matrix with $m \leq n$ and $E_2 = [E : E_1] \in U(n)$ where E is an $n \times m$ complex matrix with orthonormal columns. Then*

$$\int_{U(n)} \mathrm{etr}(XE)(dE_2) = {}_0F_1\left(n; \frac{1}{4}XX^H\right),$$

where (dE_2) denotes the normalized invariant measure on $U(n)$.

Proof. The result follows by showing that the following equation has identical Laplace transforms on both sides.

$$(\det XX^H)^{n-m}\int_{U(n)} \mathrm{etr}(XE)(dE_2) = (\det XX^H)^{n-m}{}_0F_1\left(n; \frac{1}{4}XX^H\right). \tag{3.29}$$

The Laplace transform of the right side of Equation (3.29) is given by

$$\begin{aligned} g_R(Z) &= \int_{XX^H>0} \mathrm{etr}(-XX^H Z)(\det XX^H)^{n-m}{}_0F_1\left(n; \frac{1}{4}XX^H\right)(dXX^H) \\ &= C\Gamma_m(n)(\det Z)^{-n}{}_1F_1\left(n; n; \frac{1}{4}Z^{-1}\right) \quad \text{(by Proposition 3.4)} \\ &= C\Gamma_m(n)(\det Z)^{-n}{}_0F_0\left(\frac{1}{4}Z^{-1}\right) \end{aligned}$$

$$= \mathcal{C}\Gamma_m(n)(\det Z)^{-n} \operatorname{etr}\left(\frac{1}{4}Z^{-1}\right). \tag{3.30}$$

The Laplace transform of the left side of Equation (3.29) is given by

$$g_L(Z) = \int_{XX^H>0} \operatorname{etr}(-XX^H Z)(\det XX^H)^{n-m} \int_{U(n)} \operatorname{etr}(XE)(dE_2)(dXX^H). \tag{3.31}$$

Let $A = X^H$ and $A = HT$, where H is an $n \times m$ matrix with orthonormal columns and T is an upper triangular matrix with real and positive diagonal elements. Moreover, let $W = A^H A = XX^H = T^H T$. Now from Propositions 2.4 and 2.2 we have

$$(dA) = \prod_{k=1}^{m} t_{kk}^{2n-2k+1}(dT)(H^H dH),$$

$$(dW) = 2^m \prod_{k=1}^{m} t_{kk}^{2m-2k+1}(dT) \quad \Rightarrow \quad (dT) = 2^{-m} \prod_{k=1}^{m} t_{ii}^{-2m+2k-1}(dW),$$

$$(dA) = 2^{-m} \prod_{k=1}^{m} t_{kk}^{2n-2m}(dW)(H^H dH) = 2^{-m}(\det W)^{n-m}(dW)(H^H dH).$$

Since $\prod_{k=1}^{m} t_{kk} = \det T = (\det T^H T)^{1/2} = (\det W)^{1/2}$. In other words, we have

$$(dX) = (dX^H) = 2^{-m}(\det XX^H)^{n-m}(dXX^H)(H^H dH), \tag{3.32}$$

and

$$\int_{\mathcal{C}V_{m,n}} (H^H dH) = \frac{2^m \pi^{mn}}{\mathcal{C}\Gamma_m(n)}. \tag{3.33}$$

Substituting Equation (3.32) into Equation (3.31) and using (3.33) we obtain

$$g_L(Z) = \frac{\mathcal{C}\Gamma_m(n)}{\pi^{nm}} \int_X \operatorname{etr}(-XX^H Z) \int_{U(n)} \operatorname{etr}(XE)(dE_2)(dX). \tag{3.34}$$

Let $X = Z^{-1/2}Y$. Then $(dX) = (\det Z)^{-n}(dY)$. Hence, $g_L(Z)$ becomes

$$g_L(Z) \quad = \quad \frac{\mathcal{C}\Gamma_m(n)}{\pi^{nm}} \int_{U(n)} \int_Y \operatorname{etr}(-YY^H + Z^{-1/2}YE)(dY)(dE_2)(\det Z)^{-n}$$

$$
\begin{aligned}
&= (\det Z)^{-n} \operatorname{etr}\left(\frac{1}{4}Z^{-1}\right) \frac{\mathcal{C}\Gamma_m(n)}{\pi^{nm}} \\
&\quad \cdot \int_{U(n)} \int_Y \operatorname{etr}\left[-\left(Y - \frac{1}{2}Z^{-1/2}E^H\right)\left(Y - \frac{1}{2}Z^{-1/2}E^H\right)^H\right](dY)(dE_2) \\
&= \mathcal{C}\Gamma_m(n)(\det Z)^{-n} \operatorname{etr}\left(\frac{1}{4}Z^{-1}\right). \qquad (3.35)
\end{aligned}
$$

The last equality follows by assuming $Y \sim \mathcal{CN}\left(\frac{1}{2}Z^{-1/2}E^H, I_m \otimes I_n\right)$. This completes the proof. □

The next lemma, due to Muirhead [36], is required for the proof of Proposition 3.7.

Lemma 3.2 *Let $E_2 = [E : E_1] \in U(n)$, E be an $n \times m$ complex matrix with orthonormal columns and $G = G(E)$ be any $n \times (n-m)$ matrix with orthonormal columns orthogonal to those of E. Then we have*

$$
\int_{U(n)} f(E, E_1)(E_2^H dE_2) = \int_{E \in \mathcal{CV}_{m,n}} \int_{K \in U(n-m)} f(E, GK)(K^H dK)(E^H dE), \qquad (3.36)
$$

where $f(\cdot,\cdot)$ is arbitrary function.

Proof. See [36]. □

Proposition 3.7 *If $E_2 = [E : E_1] \in U(n)$, and E is an $n \times m$ complex matrix with orthonormal columns, then*

$$
\int_{E \in \mathcal{CV}_{m,n}} \operatorname{etr}(XE)(E^H dE) = \frac{2^m \pi^{mn}}{\mathcal{C}\Gamma_m(n)} {}_0F_1\left(n; \frac{1}{4}XX^H\right), \qquad (3.37)
$$

where X is an $m \times n$ complex matrix.

Proof. Using Lemma 3.2, we have

$$
\begin{aligned}
\int_{E_2 \in U(n)} \operatorname{etr}(XE)(E_2^H dE_2) &= \int_{E \in \mathcal{CV}_{m,n}} \int_{K \in U(n-m)} \operatorname{etr}(XE)(K^H dK)(E^H dE) \\
&= \operatorname{Vol}[U(n-m)] \int_{E \in \mathcal{CV}_{m,n}} \operatorname{etr}(XE)(E^H dE) \qquad (3.38)
\end{aligned}
$$

Since $(dE_2) \triangleq \frac{1}{\mathrm{Vol}[U(n)]}(E_2^H dE_2)$, and hence, we obtain

$$\begin{aligned}\int_{E\in\mathcal{CV}_{m,n}} \mathrm{etr}(XE)(E^H dE) &= \frac{\mathrm{Vol}[U(n)]}{\mathrm{Vol}[U(n-m)]}\int_{E_2\in U(n)} \mathrm{etr}(XE)(dE_2) \\ &= \frac{2^m \pi^{mn}}{\mathcal{C}\Gamma_m(n)}\,{}_0F_1\left(n;\frac{1}{4}XX^H\right). \qquad (3.39)\end{aligned}$$

The last equality follows from Proposition 3.6. The proof is complete. □

3.3 Invariant polynomials

In this section, we describe a class of homogeneous polynomials $C_\phi^{\kappa,\tau}(X,Y)$ of degrees k and l in the elements of the $m\times m$ symmetric complex matrices X and Y (see, [10], [11] and [7]). These polynomials are invariant under the simultaneous transformations

$$X \to E^H X E, \qquad Y \to E^H Y E, \qquad E \in U(m).$$

Moreover, these polynomials satisfy the following relationship

$$\int_{U(m)} C_\kappa(AE^H XE)C_\tau(BE^H YE)(dE) = \sum_{\phi\in\kappa.\tau} \frac{C_\phi^{\kappa,\tau}(A,B)C_\phi^{\kappa,\tau}(X,Y)}{C_\phi(I)}, \qquad (3.40)$$

where C_κ, C_τ and C_ϕ are zonal polynomials, indexed by the ordered partitions κ, τ and ϕ of the nonnegative integers k, l, and $f=k+l$, respectively, into not more than m parts. If we let $Gl(m,\mathbb{C})$ denote the group of $m\times m$ complex nonsingular matrices, then $\phi\in\kappa.\tau$ denotes that the irreducible representation of $Gl(m,\mathbb{C})$ indexed by 2ϕ that occurs in the decomposition of the Kronecker product $2\kappa\otimes 2\tau$ of the irreducible representations indexed by 2κ and 2τ [10]. Equation (3.40) can also be written as

$$\int_{U(m)} \mathrm{etr}(AE^H XE + BE^H YE)(dE) = \sum_{k,l=0}^{\infty}\sum_{\kappa,\tau;\phi\in\kappa.\tau} \frac{C_\phi^{\kappa,\tau}(A,B)C_\phi^{\kappa,\tau}(X,Y)}{k!\ l!\ C_\phi(I)}. \qquad (3.41)$$

The following statements can be shown to be true:

- $C_\phi^{\kappa,\tau}(X,X) = \theta_\phi^{\kappa,\tau} C_\phi(X)$, where $\theta_\phi^{\kappa,\tau} = C_\phi^{\kappa,\tau}(I,I)/C_\phi(I)$
- $C_\phi^{\kappa,\tau}(X,I) = \left[\theta_\phi^{\kappa,\tau} C_\phi(I)/C_\kappa(I)\right] C_\kappa(X)$
- $C_\kappa^{\kappa,0}(X,Y) = C_\kappa(X)$ and $C_\tau^{0,\tau}(X,Y) = C_\tau(Y)$
- $C_\kappa(X)C_\tau(Y) = \sum_{\phi\in\kappa.\tau} \theta_\phi^{\kappa,\tau} C_\phi^{\kappa,\tau}(X,Y)$

The incomplete gamma function can be written as [10]

$$\int_0^X \operatorname{etr}(-AY)(\det Y)^{n-m} C_\tau(BY)(dY) = \frac{\mathcal{C}\Gamma_m(n)\mathcal{C}\Gamma_m(m)}{\mathcal{C}\Gamma_m(m+n)}(\det X)^n \cdot \sum_{k=0}^{\infty} \sum_{\kappa;\phi\in\kappa.\tau} \frac{[n]_\phi \theta_\phi^{\kappa,\tau} C_\phi^{\kappa,\tau}(-AX,BX)}{k!\ [n+m]_\phi}. \tag{3.42}$$

3.4 Computation of ${}_pF_q$

Here we show that the complex hypergeometric series can be computed by truncating the series to finite terms. The accuracy of this approximation is illustrated by numerical examples, where we compare the exact and approximated values for specific functions such as ${}_0F_0$ and ${}_1F_0$. Recall that the hypergeometric function of a complex matrix argument is defined as

$${}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;X) = \sum_{k=0}^{\infty}\sum_{\kappa} \frac{[a_1]_\kappa \cdots [a_p]_\kappa}{[b_1]_\kappa \cdots [b_q]_\kappa} \frac{C_\kappa(X)}{k!}, \tag{3.43}$$

where $X \in \mathbb{C}^{m\times m}$, and $\{a_i\}_{i=1}^p$ and $\{b_i\}_{i=1}^q$ are arbitrary complex numbers. Note that none of the denominator parameters b_i is allowed to be zero or an integer or half-integer $\le (m-1)/2$. Otherwise some of the denominator terms are zero [36]. The convergence of this series can be categorized as follows [36]:

(i) If $p \le q$, then the series converges for all X.

(ii) If $p = q+1$, then the series converges for $\|X\| < 1$ (i.e., $\lambda_{\max} < 1$).

(iii) If $p > q+1$, then the series diverges for all $X \neq 0$, unless it terminates. Note that the series terminates when some of the numerators $[a_j]_\kappa$ in the series vanish.

Furthermore, the complex zonal polynomial is defined by

$$C_\kappa(X) = \chi_{[\kappa]}(1)\chi_{[\kappa]}(X), \tag{3.44}$$

where $\chi_{[\kappa]}(1)$ and $\chi_{[\kappa]}(X)$ are given in section 3.1. It is clear from Equation (3.43) that the calculation of zonal polynomials is the first step in the computation of hypergeometric functions. We have programmed an algorithm to compute these zonal polynomials based on Equation (3.44). This algorithm enables us to compute the hypergeometric functions.

Let us consider the function ${}_0F_1(b;\Lambda)$, which is going to appear in the noncentral Wishart distribution in Chapter 5. If $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2)$, then the hypergeometric function ${}_0F_1(b;\Lambda)$ can be written as

$$\begin{aligned} {}_0F_1(b;\Lambda) = 1 &+ \frac{C_{(1,0)}(\Lambda)}{[b]_{(1,0)}\ 1!} + \left[\frac{C_{(2,0)}(\Lambda)}{[b]_{(2,0)}\ 2!} + \frac{C_{(1,1)}(\Lambda)}{[b]_{(1,1)}\ 2!}\right] + \left[\frac{C_{(3,0)}(\Lambda)}{[b]_{(3,0)}\ 3!} + \frac{C_{(2,1)}(\Lambda)}{[b]_{(2,1)}\ 3!}\right] \\ &+ \left[\frac{C_{(4,0)}(\Lambda)}{[b]_{(4,0)}\ 4!} + \frac{C_{(3,1)}(\Lambda)}{[b]_{(3,1)}\ 4!} + \frac{C_{(2,2)}(\Lambda)}{[b]_{(2,2)}\ 4!}\right] + \cdots . \end{aligned} \tag{3.45}$$

Using the zonal polynomial formula given in Equation (3.1), we can show that the second term ($k = 1$) on the right hand side of Equation (3.45) is given by

$$\frac{C_{(1,0)}(\Lambda)}{[b]_{(1,0)}\ 1!} = \frac{(\lambda_1 + \lambda_2)}{b\ 1!}.$$

Similarly, the third and the fourth terms on the right hand side of Equation (3.45) are given by

$$\frac{C_{(2,0)}(\Lambda)}{[b]_{(2,0)}\ 2!} + \frac{C_{(1,1)}(\Lambda)}{[b]_{(1,1)}\ 2!} = \frac{\lambda_1^2 + \lambda_1\lambda_2 + \lambda_2^2}{b(b+1)\ 2!} + \frac{\lambda_1\lambda_2}{b^2\ 2!}$$

and

$$\frac{C_{(3,0)}(\Lambda)}{[b]_{(3,0)}\ 3!} + \frac{C_{(2,1)}(\Lambda)}{[b]_{(2,1)}\ 3!} = \frac{\lambda_1^3 + \lambda_1^2\lambda_2 + \lambda_1\lambda_2^2 + \lambda_2^3}{b(b+1)(b+2)\ 3!} + \frac{2\lambda_1^2\lambda_2 + 2\lambda_1\lambda_2^2}{b^2(b+1)\ 3!},$$

respectively. The general $(k+1)$-st term is given by

$$\frac{C_{(k,0)}(\Lambda)}{[b]_{(k,0)}\ k!} + \frac{C_{(k-1,1)}(\Lambda)}{[b]_{(k-1,1)}\ k!} + \cdots + \frac{C_{(k-l,l)}(\Lambda)}{[b]_{(k-l,l)}\ k!} + \cdots + \frac{C_{(k/2,k/2)}(\Lambda)}{[b]_{(k/2,k/2)}\ k!}$$

where $k \geq 2l$ and the zonal polynomials are given by

$$\begin{aligned}
C_{(k,0)}(\Lambda) &= \lambda_1^k + \lambda_1^{k-1}\lambda_2 + \cdots + \lambda_1\lambda_2^{k-1} + \lambda_2^k \\
C_{(k-1,1)}(\Lambda) &= (k-1)\left(\lambda_1^{k-1}\lambda_2 + \cdots + \lambda_1\lambda_2^{k-1}\right) \\
\vdots &= \vdots \\
C_{(k-l,l)}(\Lambda) &= k!\frac{k-2l+1}{(k-l+1)!\,l!}\left(\lambda_1^{k-l}\lambda_2^l + \cdots + \lambda_1^l\lambda_2^{k-l}\right) \\
\vdots &= \vdots \\
C_{(k/2,k/2)}(\Lambda) &= \frac{k!}{(k-k/2+1)!\,(k/2)!}\left(\lambda_1^{k/2}\lambda_2^{k/2}\right).
\end{aligned}$$

It is clear from Equation (3.45) that we need to truncate this series to compute this function ${}_0F_1(b;\Lambda)$. In other words, ${}_0F_1(b;\Lambda)$ can be approximated as

$$ {}_0F_1(b;\Lambda) \approx \sum_{k=1}^{M}\sum_{\kappa}\frac{C_\kappa(\Lambda)}{[b]_\kappa\, k!}. \tag{3.46}$$

Table 3.1 shows numerical examples which illustrate the accuracy of this approximation, where $\Lambda = \text{diag}(0.75, 0.25)$, $a = 2$ and $b = 2$. The series ${}_0F_1(2;\Lambda)$ converges to 1.6306 with $M = 5$ (i.e., ${}_0F_1(2;\Lambda) = 1.6306$ for $M \geq 5$). For illustrative purpose, we chose to compute the series ${}_0F_0$ and ${}_1F_0$, because there are formulas to compute exact values for these series. The first seven terms, $M = 7$, of the series ${}_0F_0(\Lambda)$ give a four decimal agreement with the exact value. The series ${}_1F_0(2;\Lambda)$ needs more terms, $M = 30$, to approach the four decimal agreement with exact value. In this case $p = q + 1$ and $\lambda_{\max} = 0.75 < 1$; therefore the series must converge (see item (*ii*) on the previous pages).

${}_0F_1(b;\Lambda)$	${}_0F_0(\Lambda)$		${}_1F_0(a;\Lambda)$	
approx., $M = 5$	$\text{etr}(\Lambda)$	approx., $M = 7$	$\det(I-\Lambda)^{-a}$	approx., $M = 30$
1.6306	2.7183	2.7183	28.4444	28.4035

Table 3.1: Numerical examples of hypergeometric series

Chapter 4

The Complex Central Wishart Matrix

The complex central Wishart matrix is studied in this chapter. Section 4.1 defines the matrix variate complex Gaussian density. Section 4.2 defines the complex central Wishart density, and in section 4.3 the joint eigenvalue density is given. The maximum eigenvalue distribution is derived in section 4.4, while, the minimum eigenvalue distribution is derived in section 4.5. Section 4.6 gives the kth largest eigenvalue distribution. Finally, in section 4.7, the density of the condition number is derived.

4.1 Matrix variate complex Gaussian distribution

Let an $r \times s$ complex random matrix $\mathbf{Y} \sim \mathcal{CN}(M, C \otimes \Sigma)$ be normally distributed, where M is $r \times s$ and C and Σ are $r \times r$ and $s \times s$ positive definite matrices, respectively. This means that $\mathcal{E}\{\mathbf{Y}\} = M$ and that $C \otimes \Sigma$ is the covariance matrix of the vector $\mathbf{y} = \operatorname{vec} \mathbf{Y}^H \in \mathbb{C}^{rs \times 1}$, i.e.,

$$\mathbf{Y} \sim \mathcal{CN}(M, C \otimes \Sigma) \iff \mathbf{y} \sim \mathcal{CN}(m, C \otimes \Sigma), \text{ where } m = \operatorname{vec} M^H.$$

The following equalities hold [36]:

(i) $\det(C \otimes \Sigma) = (\det C)^s (\det \Sigma)^r$,

(ii) $(C \otimes \Sigma)^{-1} = C^{-1} \otimes \Sigma^{-1}$,

(iii)
$$\begin{aligned} \operatorname{tr}(BX^H CX\Sigma) &= [\operatorname{vec} X]^H (B^H \Sigma^H \otimes C)[\operatorname{vec} X] \\ &= [\operatorname{vec} X]^H (\Sigma B \otimes C^H)[\operatorname{vec} X]. \end{aligned}$$

The joint density function of the elements of $\mathbf{Y}$ is given by

Proposition 4.1 *Consider the $r \times s$ matrix $\mathbf{Y} \sim \mathcal{CN}(M, C \otimes \Sigma)$, where C and Σ are $r \times r$ and $s \times s$ Hermitian positive definite matrices. Then the density of $\mathbf{Y}$ is given by*

$$f(Y) = \pi^{-rs} (\det C)^{-s} (\det \Sigma)^{-r} \operatorname{etr}\left[-C^{-1}(Y - M)\Sigma^{-1}(Y - M)^H\right]. \tag{4.1}$$

Proof. Since $\mathbf{y} \sim \mathcal{CN}(m, C \otimes \Sigma)$ with $m = \operatorname{vec} M^H$, the joint density of the elements of $\mathbf{y}$ is

$$\begin{aligned} f(Y) &= \pi^{-rs}[\det C \otimes \Sigma]^{-1} \exp\left[-(y - m)^H (C \otimes \Sigma)^{-1}(y - m)\right] \\ &= \pi^{-rs}[\det C \otimes \Sigma]^{-1} \exp\left[(\operatorname{vec} Y^H - \operatorname{vec} M^H)^H C^{-1} \otimes \Sigma^{-1} (\operatorname{vec} Y^H - \operatorname{vec} M^H)\right] \\ &= \pi^{-rs}(\det C)^{-s}(\det \Sigma)^{-r} \operatorname{etr}\left[-C^{-1}(Y - M)\Sigma^{-1}(Y - M)^H\right]. \end{aligned} \tag{4.2}$$

□

The density (4.1) is symmetrical; it is invariant under the group of transformations [24],

$$\begin{aligned} Y &\to DYE \qquad D \in GL(r, \mathbb{C}), \\ M &\to DME \qquad E \in U(s), \\ \Sigma &\to D\Sigma D^H. \end{aligned} \tag{4.3}$$

$GL(r, \mathbb{C})$ is the group of all nonsingular $r \times r$ matrices D with complex elements, i.e., the full linear group, and $U(s)$ is the group of all $s \times s$ complex unitary matrices E; $E^H E = I_s$, i.e., the unitary group.

4.2 The complex central Wishart distribution

The definition of the complex central Wishart distribution is given by

Definition 4.1 *Let* $\mathbf{W} = \mathbf{A}^H\mathbf{A}$, *where the* $n \times m$ *matrix* $\mathbf{A}$ *is distributed as* $\mathbf{A} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I_n} \otimes \Sigma)$. *Then* $\mathbf{W}$ *is said to have the complex central Wishart distribution with* n *degrees of freedom and covariance matrix* Σ, *denoted as* $\mathbf{W} \sim \mathcal{CW}_m(n, \Sigma)$.

The density of this complex central Wishart matrix is given by [24].

Proposition 4.2 *Let* $\mathbf{W} \sim \mathcal{CW}_m(n, \Sigma)$ *with* $n \geq m$. *Then the density of* $\mathbf{W}$ *is given by*

$$f(W) = \frac{1}{\mathcal{C}\Gamma_m(n)(\det \Sigma)^n} \operatorname{etr}\left(-\Sigma^{-1}W\right)(\det W)^{n-m}, \tag{4.4}$$

where $\mathcal{C}\Gamma_m(n)$ *denotes the complex multivariate gamma function given by*

$$\mathcal{C}\Gamma_m(n) = \pi^{m(m-1)/2}\prod_{k=1}^{m}\Gamma(n-k+1). \tag{4.5}$$

Proof. Let $\mathbf{W} = \mathbf{A}^H\mathbf{A}$, where the $n \times m$ matrix $\mathbf{A} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I_n} \otimes \Sigma)$. The density of $\mathbf{A}$ is

$$\begin{aligned} f(A) &= \pi^{-nm}(\det \Sigma)^{-n} \operatorname{etr}\left[-A\Sigma^{-1}A^H\right](dA) \\ &= \pi^{-nm}(\det \Sigma)^{-n} \operatorname{etr}\left[-\Sigma^{-1}A^H A\right](dA), \end{aligned} \tag{4.6}$$

where the volume element $(dA) \equiv \bigwedge_{i=1}^{n}\bigwedge_{j=1}^{m} da_{ij}$ is included to facilitate the calculation of Jacobians when we transform A. Since $n \geq m$, A has rank m with probability 1. Let $A = ET$, where E is an $n \times m$ matrix with orthonormal columns, $E^H E = I_m$. As a subspace of $\mathbb{R}^{2mn}$, these matrices form a submanifold $\mathcal{CV}_{m,n}$ (so-called Stiefel manifold) of dimension $2mn - m^2$ with volume elements $(E^H dE)$. It is shown in Proposition 2.7 that the total volume of $\mathcal{CV}_{m,n}$ is

$$\int_{\mathcal{CV}_{m,n}} (E^H dE) = \frac{2^m \pi^{mn}}{\mathcal{C}\Gamma_m(n)}. \tag{4.7}$$

Moreover, $W = A^H A = T^H T$ and from Propositions 2.4 and 2.2 we have

$$(dA) = \prod_{k=1}^{m} t_{kk}^{2n-2k+1} (dT)(E^H dE),$$

$$(dW) = 2^m \prod_{k=1}^{m} t_{kk}^{2m-2k+1} (dT) \quad \Rightarrow \quad (dT) = 2^{-m} \prod_{k=1}^{m} t_{ii}^{-2m+2k-1} (dW),$$

$$(dA) = 2^{-m} \prod_{k=1}^{m} t_{kk}^{2n-2m} (dW)(E^H dE) = 2^{-m} (\det W)^{n-m} (dW)(E^H dE).$$

Since $\prod_{k=1}^{m} t_{kk} = \det T = (\det T^H T)^{1/2} = (\det W)^{1/2}$, the joint density of W and $(E^H dE)$ is

$$\pi^{-nm} (\det \Sigma)^{-n} \operatorname{etr}\left[-\Sigma^{-1} A^H A\right] 2^{-m} (\det W)^{n-m} (dW)(E^H dE).$$

Integrating with respect to E over the Stiefel manifold $\mathcal{CV}_{m,n}$ and using Equation (4.7), we obtain the marginal density function of $\mathbf{W}$, given by Equation (4.4). The proof is complete. □

4.3 Joint eigenvalue distribution

Next, we consider the eigenvalue density of a complex Wishart matrix.

Proposition 4.3 *Let* $\mathbf{W}$ *be an arbitrary* $m \times m$ *positive definite complex random matrix with Wishart density function* $f(W)$*. Then the joint density function of the eigenvalues* $\lambda_1, \ldots, \lambda_m$ *of* $\mathbf{W}$ *is*

$$\frac{\pi^{m(m-1)}}{\mathcal{C}\Gamma_m(m)} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 \int_{U(m)} f(E \Lambda E^H)(dE), \tag{4.8}$$

where $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_m)$ *and* $W = E \Lambda E^H$ *is an eigendecomposition.*

Proof. Combining Equations (2.9) and (2.25), the Jacobian of an eigendecomposition can be written as

$$(dW) = \frac{2^m \pi^{m^2}}{C\Gamma_m(m)} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 (d\Lambda)(dE). \tag{4.9}$$

The required density can be obtained by substituting (dW) and $W = E\Lambda E^H$ in $f(W)(dW)$ and integrating with respect to (dE) over $U(m)$. Note that we must divide the density by $(2\pi)^m$ which normalizes the arbitrary phases of the m elements in the first row of E. Therefore, we have

$$\frac{\pi^{m(m-1)}}{C\Gamma_m(m)} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 \int_{U(m)} f(E\Lambda E^H)(dE). \tag{4.10}$$

□

The following proposition gives the joint density of the eigenvalues of a complex Wishart matrix [24].

Proposition 4.4 *Let* $\mathbf{W} \sim \mathcal{C}\mathcal{W}_m(n, \Sigma)$ *with* $n > m - 1$. *Then* $\mathbf{W}$ *is an* $m \times m$ *positive definite Hermitian matrix with real eigenvalues. The joint density of the eigenvalues* $\lambda_1, \ldots, \lambda_m$ *of* $\mathbf{W}$ *is*

$$\frac{\pi^{m(m-1)} (\det \Sigma)^{-n}}{C\Gamma_m(m) C\Gamma_m(n)} \prod_{k=1}^{m} \lambda_k^{n-m} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 \int_{U(m)} \operatorname{etr}\left(-\Sigma^{-1} E\Lambda E^H\right)(dE). \tag{4.11}$$

Moreover

$$\begin{aligned} \int_{U(m)} \operatorname{etr}\left(-\Sigma^{-1} E\Lambda E^H\right)(dE) &= \int_{U(m)} {}_0F_0\left(-\Sigma^{-1} E\Lambda E^H\right)(dE) \\ &= {}_0F_0\left(-\Lambda, \Sigma^{-1}\right) \\ &= \sum_{k=1}^{\infty} \sum_{\kappa} \frac{C_\kappa(-\Lambda) C_\kappa(\Sigma^{-1})}{k! C_\kappa(I_m)}. \end{aligned} \tag{4.12}$$

Proof. We use arguments similar to those found in [5], [33] and [49]. By substituting the Wishart density (4.4) into (4.8) and noting that $\det W = \det E\Lambda E^H = \prod_{k=1}^{m} \lambda_k$

we obtain the desired results, i.e.,

$$\frac{\pi^{m(m-1)}(\det\Sigma)^{-n}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)}\prod_{k=1}^{m}\lambda_k^{n-m}\prod_{k<l}^{m}(\lambda_k-\lambda_l)^2\int_{U(m)}\operatorname{etr}(-\Sigma^{-1}E\Lambda E^H)(dE). \tag{4.13}$$

The proof of the second part is given in Chapter 3. □

Note that the integral in Equation (4.11) depends on the population covariance matrix Σ only through its eigenvalues $v_1,\ldots,v_m$. This can be seen by writing $\Sigma = F\Upsilon F^H$, where $F \in U(m)$ and $\Upsilon = \operatorname{diag}(v_1,\ldots,v_m)$. Now we have

$$\begin{aligned}
\int_{U(m)}\operatorname{etr}\left(-\Sigma^{-1}E\Lambda E^H\right)(dE) &= \int_{U(m)}\operatorname{etr}\left(-F\Upsilon^{-1}F^H E\Lambda E^H\right)(dE)\\
&= \int_{U(m)}\operatorname{etr}\left(-\Upsilon^{-1}F^H E\Lambda E^H F\right)(dE)\\
&= \int_{U(m)}\operatorname{etr}\left(-\Upsilon^{-1}\widehat{E}\Lambda\widehat{E}^H\right)(d\hat{E})\\
&= \int_{U(m)}{}_0F_0\left(-\Upsilon^{-1}\widehat{E}\Lambda\widehat{E}^H\right)(d\hat{E})\\
&= {}_0F_0\left(-\Lambda,\Upsilon^{-1}\right)\\
&= \sum_{k=1}^{\infty}\sum_{\kappa}\frac{C_\kappa(-\Lambda)C_\kappa(\Upsilon^{-1})}{k!C_\kappa(I_m)},
\end{aligned} \tag{4.14}$$

where $\widehat{E} = F^H E \in U(m)$ and $(dE) = (d\widehat{E})$. This was observed in [36]. In general, the integral in Equation (4.11) is not easy to evaluate. An infinite series representation for this integral in terms of zonal polynomials is shown in Equation (4.12). One of the objectives of our work is to compute this kind of zonal polynomials and evaluate the exact and the limiting distributions of eigenvalues of complex central Wishart matrices.

Note that, if $\Sigma = \sigma^2 I_m$, then the joint density of the eigenvalues $\lambda_1,\ldots,\lambda_m$ has a simple form and does not require a zonal polynomial representation.

Proposition 4.5 *Let* $\mathbf{W} \sim \mathcal{CW}_m(n,\sigma^2 I_m)$ *with* $n > m-1$. *Then the joint density* *of the eigenvalues* $\lambda_1, \ldots, \lambda_m$ *of* $\mathbf{W}$ *is*

$$\frac{\pi^{m(m-1)}(\sigma^2)^{-nm}}{C\Gamma_m(m)C\Gamma_m(n)} \prod_{k=1}^{m} \lambda_k^{n-m} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 \exp\left(-\frac{1}{\sigma^2}\sum_{k=1}^{m} \lambda_k\right). \tag{4.15}$$

Proof. Putting $\Sigma = \sigma^2 I_m$ in Proposition 4.4 and noting that

$$\begin{aligned} \int_{U(m)} \operatorname{etr}\left(-\frac{1}{\sigma^2} E\Lambda E^H\right)(dE) &= \operatorname{etr}\left(-\frac{1}{\sigma^2}\Lambda\right)\int_{U(m)} (dE) \\ &= \exp\left(-\frac{1}{\sigma^2}\sum_{i=1}^{m} \lambda_i\right) \end{aligned} \tag{4.16}$$

completes the proof. □

Note that, for real Wishart matrices, the corresponding results for Propositions 4.1–4.5 are given in [36], Chapter 3. The real Wishart matrices are well studied in [36], [1] and [15].

4.4 Distribution of $\lambda_{\max}$

Theorem 4.1 *Let* $W \sim C\mathcal{W}_m(n, \Sigma)$ $(n \geq m)$ *and* Δ *be an* $m \times m$ *positive definite matrix. Then the probability* $P(W < \Delta)$ *is given by*

$$P(W < \Delta) = \frac{C\Gamma_m(m)}{C\Gamma_m(n+m)} \frac{(\det \Delta)^n}{(\det \Sigma)^n} {}_1F_1(n, n+m, -\Sigma^{-1}\Delta), \tag{4.17}$$

where

$${}_1F_1(a, b, X) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{[a]_\kappa}{[b]_\kappa} \frac{C_\kappa(X)}{k!}.$$

Proof. Using the Wishart density (4.4) we can write the probability $P(W < \Delta)$ as

$$P(W < \Delta) = \frac{1}{C\Gamma_m(n)(\det \Sigma)^n} \int_{0<W<\Delta} \operatorname{etr}(-\Sigma^{-1}W)(\det W)^{n-m}\,(dW).$$

The change of variable $W = \Delta^{1/2} X \Delta^{1/2}$ leads to the differential form $(dW) = (\det \Delta)^m (dX)$. Hence,

$$\begin{aligned} P(W < \Delta) &= \frac{(\det \Delta)^n}{C\Gamma_m(n)(\det \Sigma)^n} \int_{0<X<I} \text{etr}(-\Delta^{1/2}\Sigma^{-1}\Delta^{1/2}X)(\det X)^{n-m}\,(dX) \\ &= \frac{(\det \Delta)^n}{C\Gamma_m(n)(\det \Sigma)^n} \sum_{k=0}^{\infty}\sum_{\kappa} \frac{1}{k!} \int_0^I (\det X)^{n-m} C_\kappa(-\Delta^{1/2}\Sigma^{-1}\Delta^{1/2}X)\,(dX) \\ &= \frac{C\Gamma_m(m)}{C\Gamma_m(n+m)} \frac{(\det \Delta)^n}{(\det \Sigma)^n} \sum_{k=0}^{\infty}\sum_{\kappa} \frac{[n]_\kappa}{[n+m]_\kappa} \frac{C_\kappa(-\Sigma^{-1}\Delta)}{k!} \\ &= \frac{C\Gamma_m(m)}{C\Gamma_m(n+m)} \frac{(\det \Delta)^n}{(\det \Sigma)^n} {}_1F_1(n, n+m, -\Sigma^{-1}\Delta). \end{aligned} \tag{4.18}$$

Note that Corollary 3.2 is used in this proof. □

The following result describes the distribution of $\lambda_{\max}$, which is derived from Theorem 4.1.

Corollary 4.1 *If $W \sim C\mathcal{W}_m(n, \Sigma)$ $(n \geq m)$ and $\lambda_{\max}$ is the largest eigenvalue of W, then its distribution is given by*

$$P(\lambda_{\max} < x) = \frac{C\Gamma_m(m)}{C\Gamma_m(n+m)} \frac{x^{mn}}{(\det \Sigma)^n} {}_1F_1(n, n+m, -x\Sigma^{-1}). \tag{4.19}$$

Proof. The inequality $\lambda_{\max} < x$ is equivalent to $W < xI$. Therefore, the result follows by letting $\Delta = xI$ in Theorem 4.1. □

If $\Sigma = \sigma^2 I$, then the distribution of $\lambda_{\max}$ can be evaluated without computing the hypergeometric function. In [25], the distribution of $\lambda_{\max}$ of a Wishart matrix is obtained from the multivariate beta distribution for $\Sigma = \sigma^2 I$. Here, the result is obtained from the Wishart distribution. The following lemma is due to Khatri [29].

Lemma 4.1 *If $f_i(x)$, $i = 1, \ldots, m$, are the density functions, then we have*

$$\widetilde{\sum} \int_{D_1} \prod_{k=1}^{m} f_{i_k}(x_k)dx_k = \prod_{k=1}^{m} \left(\int_{-\infty}^{x} f_k(y)dy \right) \tag{4.20}$$

and

$$\widetilde{\sum} \int_{D_2} \prod_{k=1}^{m} f_{i_k}(x_k)dx_k = \prod_{k=1}^{m} \left(\int_{x}^{\infty} f_k(y)dy \right), \tag{4.21}$$

where $D_1 = \{x_1 < x_2 \leq \cdots \leq x_m < x\}$ *and* $D_2 = \{x \leq x_1 \leq \cdots \leq x_m\}$. *Moreover,* $\widetilde{\sum}$ *denotes summation over the permutation* $(i_1, i_2, \ldots, i_m)$ *of* $(1, 2, \ldots, m)$.

Proof. We have

$$P(x_k \leq x \ \forall \ k) = P(x_{\max} \leq x) \quad \text{and} \quad P(x_k \geq x \ \forall \ k) = P(x_{\min} \geq x).$$

The left side of Equation (4.20) is $P(x_{\max} \leq x)$, whereas its right side is $P(x_k \leq x \ \forall \ k)$ because $x_1, x_2, \ldots, x_m$ are independent variates. Similarly for Equation (4.21). □

The next lemma is needed for the derivation of $P(\lambda_{\max} \leq x)$.

Lemma 4.2 *We have*

$$\prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 = \widetilde{\sum}_k \widetilde{\sum}_t (-1)^{\mathrm{per}(t_1,\ldots,t_m)} \prod_{i=1}^{m} \lambda_{k_i}^{m-i+t_i}$$

where $\widetilde{\sum}_k$ *denotes summation over all permutations* $(k_1, \ldots, k_m)$ *of* $(1, \ldots, m)$, $\widetilde{\sum}_t$ *denotes summation over all permutations* $(t_1, \ldots, t_m)$ *of* $(0, 1, \ldots, m-1)$, *and*

$$\mathrm{per}(t_1, \ldots, t_m) = \begin{cases} 0 & \textit{even permutation}, \\ 1 & \textit{odd permutation}. \end{cases}$$

Proof. From the Vandermonde determinant, we have

$$\prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 = \det \begin{bmatrix} \lambda_1^{m-1} & \lambda_2^{m-1} & \ldots & \lambda_m^{m-1} \\ \lambda_1^{m-2} & \lambda_2^{m-2} & \ldots & \lambda_m^{m-2} \\ \vdots & \vdots & \ldots & \vdots \\ 1 & 1 & \ldots & 1 \end{bmatrix}^2$$

$$= \det \begin{bmatrix} \sum_{k=1}^m \lambda_k^{2m-2} & \sum_{k=1}^m \lambda_k^{2m-3} & \dots & \sum_{k=1}^m \lambda_k^{m-1} \\ \sum_{k=1}^m \lambda_k^{2m-3} & \sum_{k=1}^m \lambda_k^{2m-4} & \dots & \sum_{k=1}^m \lambda_k^{m-2} \\ \vdots & \vdots & \dots & \vdots \\ \sum_{k=1}^m \lambda_k^{m-1} & \sum_{k=1}^m \lambda_k^{m-2} & \dots & m \end{bmatrix}$$

$$= \widetilde{\sum_k} \det \begin{bmatrix} \lambda_{k_1}^{2m-2} & \lambda_{k_2}^{2m-3} & \dots & \lambda_{k_m}^{m-1} \\ \lambda_{k_1}^{2m-3} & \lambda_{k_2}^{2m-4} & \dots & \lambda_{k_m}^{m-2} \\ \vdots & \vdots & \dots & \vdots \\ \lambda_{k_1}^{m-1} & \lambda_{k_2}^{m-2} & \dots & \lambda_{k_m}^{0} \end{bmatrix}. \tag{4.22}$$

The last equality follows from ([35], p. 11) and the result follows from the definition of determinant. □

The next proposition gives the distribution of $\lambda_{\max}$.

Proposition 4.6 *If $\Sigma = \sigma^2 I$, then the distribution of the maximum eigenvalue of a complex central Wishart matrix is given by*

$$P(\lambda_{\max} \leq x) = K \det \Xi,$$

where

$$\begin{aligned} K &= \frac{\pi^{m(m-1)}\sigma^{-2mn}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)}, \\ \Xi &= \begin{bmatrix} \xi_0 & \dots & \xi_{m-1} \\ \vdots & \dots & \vdots \\ \xi_{m-1} & \dots & \xi_{2m-2} \end{bmatrix} = [(\xi_{i+j-2})], \\ \xi_{i+j-2} &= \int_0^x \lambda^{n-m+i+j-2} \exp\left(-\frac{1}{\sigma^2}\lambda\right) d\lambda. \end{aligned}$$

Proof. Let the domain $D_1 = \{0 < \lambda_m < \cdots < \lambda_1 \leq x\}$. Using the eigenvalue density (4.15) we can write the probability $P(\lambda_{\max} \leq x)$ as

$$P(\lambda_{\max} \leq x) = P(0 < \lambda_m < \cdots < \lambda_1 \leq x)$$

$$\begin{aligned}
&= K \int_{D_1} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 \left[\prod_{k=1}^{m} \lambda_k^{n-m} \exp\left(-\frac{1}{\sigma^2}\lambda_k\right) \right] \bigwedge_{k=1}^{m} d\lambda_k \\
&= K \widetilde{\sum}_k \widetilde{\sum}_t \int_{D_1} (-1)^{\text{per}(t_1,\ldots,t_m)} \\
&\quad \cdot \prod_{i=1}^{m} \lambda_{k_i}^{m-i+t_i} \left[\prod_{i=1}^{m} \lambda_{k_i}^{n-m} \exp\left(-\frac{1}{\sigma^2}\lambda_{k_i}\right) \right] \bigwedge_{i=1}^{m} d\lambda_{k_i} \\
&= K \widetilde{\sum}_t (-1)^{\text{per}(t_1,\ldots,t_m)} \prod_{i=1}^{m} \int_0^x \lambda_{k_i}^{n-m+m-i+t_i} \exp\left(-\frac{1}{\sigma^2}\lambda_{k_i}\right) d\lambda_{k_i} \\
&= K \widetilde{\sum}_t (-1)^{\text{per}(t_1,\ldots,t_m)} \prod_{i=1}^{m} \xi_{m-i+t_i} \\
&= K \det \Xi. \qquad (4.23)
\end{aligned}$$

Note that Lemmas 4.1 and 4.2 are used in this proof. □

4.5 Distribution of $\lambda_{\min}$

Theorem 4.2 *Let $W \sim \mathcal{C}W_m(n, \Sigma)$ $(n \geq m)$ and Δ be an $m \times m$ positive definite matrix. Then the probability $P(W > \Delta)$ can be written as a finite series, i.e.,*

$$P(W > \Delta) = \operatorname{etr}\left(-\Sigma^{-1}\Delta\right) \sum_{k=0}^{m(n-m)} \widehat{\sum}_\kappa \frac{C_k\left(\Sigma^{-1}\Delta\right)}{k!}, \qquad (4.24)$$

where $\widehat{\sum}_\kappa$ denotes summation over the partitions $\kappa = (k_1, \ldots, k_m)$ of k with $k_1 \leq n - m$.

Proof. Using the Wishart density (4.4) we can write the probability $P(W > \Delta)$ as

$$P(W > \Delta) = \frac{1}{\mathcal{C}\Gamma_m(n)(\det \Sigma)^n} \int_{W>\Delta} \operatorname{etr}\left(-\Sigma^{-1}W\right) (\det W)^{n-m} (dW). \qquad (4.25)$$

The change of variable $W = \Delta^{1/2}(I+X)\Delta^{1/2}$ leads to the differential form $(dW) = (\det \Delta)^m (dX)$. Hence,

$$\begin{aligned} P(W > \Delta) &= \frac{\operatorname{etr}(-\Sigma^{-1}\Delta)(\det\Delta)^n}{C\Gamma_m(n)(\det\Sigma)^n} \\ &\quad \cdot \int_{X>0} \operatorname{etr}\left(-\Delta^{1/2}\Sigma^{-1}\Delta^{1/2}X\right)(\det X)^{n-m}(\det(I+X^{-1}))^{n-m}(dX) \\ &= \frac{\operatorname{etr}(-\Sigma^{-1}\Delta)(\det\Delta)^n}{C\Gamma_m(n)(\det\Sigma)^n} \sum_{k=0}^{m(n-m)} \widehat{\sum_\kappa} \frac{[-(n-m)]_\kappa(-1)^k}{k!} \\ &\quad \cdot \int_{X>0} \operatorname{etr}\left(-\Delta^{1/2}\Sigma^{-1}\Delta^{1/2}X\right)(\det X)^{n-m} C_\kappa(X^{-1})(dX) \\ &= \operatorname{etr}(-\Sigma^{-1}\Delta) \sum_{k=0}^{m(n-m)} \widehat{\sum_\kappa} \frac{C_\kappa\left(\Sigma^{-1}\Delta\right)}{k!}. \end{aligned} \tag{4.26}$$

In this proof we use the following

$$\begin{aligned} \det\left(I+X^{-1}\right)^{n-m} &= {}_1F_0(-(n-m);-X^{-1}) \\ &= \sum_{k=0}^{m(n-m)} \widehat{\sum_\kappa} \frac{[-(n-m)]_\kappa C_\kappa(X^{-1})(-1)^k}{k!} \end{aligned} \tag{4.27}$$

and Corollary 3.1 in Chapter 3. Note that if any part of κ is greater that $(n-m)$ then $[-(n-m)]_\kappa = 0$. Therefore, the series ${}_1F_0$ can be written as a finite series. □

The distribution of the smallest eigenvalue is given in the following corollary.

Corollary 4.2 *If* $\mathbf{W} \sim C\mathcal{W}_m(n,\Sigma)$ *and* $\lambda_{\min}$ *is the smallest eigenvalue of* $\mathbf{W}$, *then*

$$P(\lambda_{\min} > x) = \operatorname{etr}\left(-x\Sigma^{-1}\right) \sum_{k=0}^{m(n-m)} \widehat{\sum_\kappa} \frac{C_\kappa\left(x\Sigma^{-1}\right)}{k!}, \tag{4.28}$$

where $\widehat{\sum}_\kappa$ *denotes summation over the partitions* $\kappa = (k_1,\ldots,k_m)$ *of* k *with* $k_1 \leq n-m$.

Proof. The inequality $\lambda_{\min} > x$ is equivalent to $W > xI$. Therefore, the result follows by letting $\Delta = xI$ in Theorem 4.2. □

If $\Sigma = \sigma^2 I$, then the distribution has a simple form.

Proposition 4.7 *If $\Sigma = \sigma^2 I$, then the distribution of the minimum eigenvalue of a complex central Wishart matrix is given by*

$$\begin{aligned} P(\lambda_{\min} \leq x) &= 1 - P(\lambda_{\min} \geq x) \\ &= 1 - P(x \leq \lambda_m < \cdots < \lambda_1 < \infty) \\ &= 1 - K \det \Psi, \end{aligned} \tag{4.29}$$

where

$$\begin{aligned} K &= \frac{\pi^{m(m-1)}\sigma^{-2mn}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)}, \\ \Psi &= \begin{bmatrix} \psi_0 & \dots & \psi_{m-1} \\ \vdots & \dots & \vdots \\ \psi_{m-1} & \dots & \psi_{2m-2} \end{bmatrix} = [(\psi_{i+j-2})], \\ \psi_{i+j-2} &= \int_x^\infty \lambda^{n-m+i+j-2} \exp\left(-\frac{1}{\sigma^2}\lambda\right) d\lambda. \end{aligned}$$

Proof. The proof is similar to the proof of Proposition 4.6. □

4.6 Distribution of λ_k

In this section, the distribution of the kth largest eigenvalue of the central Wishart matrix is developed, see [29]. The following lemma is due to Khatri [29].

Lemma 4.3 *If $\lambda_1 > \lambda_2 > \cdots > \lambda_m$ and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$, then*

$$C_\kappa(\Lambda) \prod_{i<j}^m (\lambda_i - \lambda_j)^2 = \chi_{[\kappa]}(1) \widetilde{\sum_r} \det \left[(\lambda_{ri}^{k_j+2m-j-i}) \right],$$

where $\widetilde{\sum}_r$ *denotes summation over all permutations* $(r_1, \ldots, r_m)$ *of* $(1, \ldots, m)$.

Proof. From the definition of $C_\kappa(\Lambda)$ (see Equation (3.1)), we have

$$\begin{aligned} C_\kappa(\Lambda)\prod_{i<j}^{m}(\lambda_i - \lambda_j)^2 &= \chi_{[\kappa]}(1)\det\left[\left(\lambda_i^{k_j+m-j}\right)\right]\det\left[\left(\lambda_i^{m-i}\right)\right] \\ &= \chi_{[\kappa]}(1)\det\left[\left(\sum_{r=1}^{m}\lambda_r^{k_j+2m-j-i}\right)\right] \\ &= \chi_{[\kappa]}(1)\widetilde{\sum}_r\det\left[\left(\lambda_{r_i}^{k_j+2m-j-i}\right)\right]. \end{aligned} \tag{4.30}$$

The last equality follows from ([35], p. 11). □

Proposition 4.8 *The distribution of the kth largest eigenvalue of a complex central Wishart matrix, where* $k > 1$, *can be obtained by considering the following probability:*

$$\begin{aligned} P(\lambda_k \le x) &= P(\lambda_{k-1} \le x) + P(\lambda_m < \cdots < \lambda_k < x < \lambda_{k-1} < \cdots < \lambda_1) \\ &= P(\lambda_{k-1} \le x) + p, \end{aligned} \tag{4.31}$$

where

$$\begin{aligned} p &= K_1\sum_{k=1}^{\infty}\sum_{\kappa}\frac{C_\kappa(I-\Sigma^{-1})}{k!C_\kappa(I)}\chi_{[\kappa]}(1)\sum_{1}\det\left[\left(\alpha_{s_i,j}(\kappa)\right)\right], \\ K_1 &= \frac{\pi^{m(m-1)}(\det\Sigma)^{-n}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)}, \\ \alpha_{s_i,j}(\kappa) &= \begin{cases} \int_x^\infty \lambda^{k_j+n+m-s_i-j}\exp(-\lambda)d\lambda & \textit{for } i = 1, \ldots, i-1, \\ \int_0^x \lambda^{k_j+n+m-s_i-j}\exp(-\lambda)d\lambda & \textit{for } i = i, \ldots, m, \end{cases} \end{aligned}$$

and $\sum_1$ *denotes summation over the combinations* $(s_1 < s_2 < \cdots < s_{i-1})$ *and* $(s_i < s_{i+1} < \cdots < s_m)$ *and* $(s_1, \ldots, s_m)$ *is a permutation of* $(1, \ldots, m)$.

Proof. First note the following

$$ {}_0F_0\left(-\Lambda, \Sigma^{-1}\right) = \int_{U(m)} \operatorname{etr}\left(-\Sigma^{-1}E\Lambda E^H\right)(dE)$$

$$
\begin{aligned}
&= \int_{U(m)} \operatorname{etr}\left(\left[I-\Sigma^{-1}\right] E\Lambda E^{H}\right) \operatorname{etr}\left(-E\Lambda E^{H}\right)(dE) \\
&= \operatorname{etr}(-\Lambda) \int_{U(m)} \operatorname{etr}\left(\left[I-\Sigma^{-1}\right] E\Lambda E^{H}\right)(dE) \\
&= \operatorname{etr}(-\Lambda)\,{}_0F_0\left(I-\Sigma^{-1}, \Lambda\right).
\end{aligned}
$$

Let the domain $D_3 = \{\lambda_m < \cdots < \lambda_k < x < \lambda_{k-1} < \cdots < \lambda_1\}$. Using the eigenvalue density (4.11) we can write the probability p as

$$
\begin{aligned}
p &= K_1 \int_{D_3} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 \left[(\det \Lambda)^{n-m}\,{}_0F_0(I-\Sigma^{-1}, \Lambda) \operatorname{etr}(-\Lambda)\right] \bigwedge_{k=1}^{m} d\lambda_k \\
&= K_1 \sum_{k=1}^{\infty} \sum_{\kappa} \frac{C_\kappa(I-\Sigma^{-1})}{k! C_\kappa(I)} \int_{D_3} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 \left[C_\kappa(\Lambda)(\det \Lambda)^{n-m} \operatorname{etr}(-\Lambda)\right] \bigwedge_{k=1}^{m} d\lambda_k \\
&= K_1 \sum_{k=1}^{\infty} \sum_{\kappa} \frac{C_\kappa(I-\Sigma^{-1})}{k! C_\kappa(I)} \chi_{[\kappa]}(1) \widetilde{\sum}_r \int_{D_3} \det\left[\left(\lambda_{r_i}^{k_j+2m-j-i}\right)\right] (\det \Lambda)^{n-m} \operatorname{etr}(-\Lambda) \bigwedge_{k=1}^{m} d\lambda_k.
\end{aligned}
$$

Now we have

$$
\det\left[\left(\lambda_{r_i}^{k_j+2m-j-i}\right)\right] = \widetilde{\sum}_t (-1)^{\operatorname{per}(s_1,\ldots,s_m)} (-1)^{\operatorname{per}(t_1,\ldots,t_m)} \prod_{i=1}^{m} \lambda_{r_{s_i}}^{k_{t_i}+2m-s_i-t_i},
$$

where $\widetilde{\sum}_t$ denotes summation over all permutations $(t_1, \ldots, t_m)$ of $(1, \ldots, m)$ and $(s_1, \ldots, s_m)$ are permutations of $(1, \ldots, m)$. Note that

$$
\widetilde{\sum}_r = \sum_1 \widetilde{\sum}_{r_{s_1}} \widetilde{\sum}_{r_{s_2}}
$$

where $\widetilde{\sum}_{r_{s_1}}$ denotes summation over the permutations $(r_{s_1}, \ldots, r_{s_{i-1}})$ of $(1, 2, \ldots, i-1)$ and $\widetilde{\sum}_{r_{s_2}}$ denotes summation over the permutations $(r_{s_i}, \ldots, r_{s_m})$ of $(i, i+1, \ldots, m)$. Therefore, we have

$$
\begin{aligned}
p &= K_1 \sum_{k=1}^{\infty} \sum_{\kappa} \frac{C_\kappa(I-\Sigma^{-1})}{k! C_\kappa(I)} \chi_{[\kappa]}(1) \sum_1 \widetilde{\sum}_t (-1)^{\operatorname{per}(s_1,\ldots,s_m)} (-1)^{\operatorname{per}(t_1,\ldots,t_m)} \\
&\quad \cdot I_1(\kappa, s, t) I_2(\kappa, s, t),
\end{aligned} \tag{4.32}
$$

where

$$
\begin{aligned}
I_1(\kappa, s, t) &= \sum_{r_{s_1}} \int_{D4} \prod_{i=1}^{i-1} \lambda_{r_{s_i}}^{k_{t_i}+n+m-s_i-t_i} \exp(-\lambda_{r_{s_i}})\, d\lambda_{r_{s_i}} \\
&= \prod_{i=1}^{i-1} \int_x^{\infty} \lambda_{r_{s_i}}^{k_{t_i}+n+m-s_i-t_i} \exp(-\lambda_{r_{s_i}})\, d\lambda_{r_{s_i}}, \quad \text{(by Lemma 4.1)}, \\
I_2(\kappa, s, t) &= \sum_{r_{s_2}} \int_{D5} \prod_{i=i}^{m} \lambda_{r_{s_i}}^{k_{t_i}+n+m-s_i-t_i} \exp(-\lambda_{r_{s_i}})\, d\lambda_{r_{s_i}} \\
&= \prod_{i=i}^{m} \int_0^{x} \lambda_{r_{s_i}}^{k_{t_i}+n+m-s_i-t_i} \exp(-\lambda_{r_{s_i}})\, d\lambda_{r_{s_i}}, \quad \text{(by Lemma 4.1)},
\end{aligned}
$$

$D_4 = \{x < \lambda_{i-1} < \cdots < \lambda_1 < \infty\}$ and $D_5 = \{0 < \lambda_m < \cdots < \lambda_i < x\}$. The result follows from the definition of the determinant. □

Proposition 4.9 *Let $\Sigma = \sigma^2 I$. The distribution of the kth largest eigenvalue, where $k > 1$, can be obtained by considering the following probability:*

$$
\begin{aligned}
P(\lambda_k \le x) &= P(\lambda_{k-1} \le x) + P(\lambda_m < \cdots < \lambda_k < x < \lambda_{k-1} < \cdots < \lambda_1) \\
&= P(\lambda_{k-1} \le x) + p, \qquad (4.33)
\end{aligned}
$$

where

$$
\begin{aligned}
p &= K \sum_1 \det\left[(\alpha_{i+j-2})\right], \\
K &= \frac{\pi^{m(m-1)} \sigma^{-2mn}}{\mathcal{C}\Gamma_m(m) \mathcal{C}\Gamma_m(n)}, \\
\alpha_{i+j-2} &= \begin{cases} \int_x^{\infty} \lambda^{n-m+i+j-2} \exp(-\frac{1}{\sigma^2}\lambda) d\lambda & \text{for } i = 1, \ldots, i-1, \\ \int_0^{x} \lambda^{n-m+i+j-2} \exp(-\frac{1}{\sigma^2}\lambda) d\lambda & \text{for } i = i, \ldots, m, \end{cases}
\end{aligned}
$$

and $\sum_1$ denotes summation over the combinations $(s_1 < s_2 < \cdots < s_{i-1})$ and $(s_i < s_{i+1} < \cdots < s_m)$ and $(s_1, \ldots, s_m)$ is a permutation of $(1, \ldots, m)$.

Proof. The proof is similar to the proof of Proposition 4.8. Let the domain $D_3 = \{\lambda_m < \cdots < \lambda_k < x < \lambda_{k-1} < \cdots < \lambda_1\}$. Using the eigenvalue density (4.15) we can write the probability p as

$$\begin{aligned} p &= K \int_{D_3} \prod_{i<j}^{m} (\lambda_i - \lambda_j)^2 \left[\prod_{i=1}^{m} \lambda_i^{n-m} \exp\left(-\frac{1}{\sigma^2}\lambda_i\right) \right] \bigwedge_{i=1}^{m} d\lambda_i \\ &= K \widetilde{\sum_r} \widetilde{\sum_t} \int_{D_3} (-1)^{\mathrm{per}(t_1,\ldots,t_m)} \prod_{i=1}^{m} \lambda_{r_i}^{m-i+t_i} \left[\prod_{i=1}^{m} \lambda_{r_i}^{n-m} \exp\left(-\frac{1}{\sigma^2}\lambda_{r_i}\right) \right] \bigwedge_{i=1}^{m} d\lambda_{r_i} \\ &= K \sum_{1} \widetilde{\sum_t} (-1)^{\mathrm{per}(t_1,\ldots,t_m)} I_1(s,t) I_2(s,t), \end{aligned} \tag{4.34}$$

where

$$\begin{aligned} I_1(s,t) &= \sum_{r_{s_1}} \int_{D4} \prod_{i=1}^{i-1} \lambda_{r_{s_i}}^{n-m+m-s_i-t_i} \exp\left(-\frac{1}{\sigma^2}\lambda_{r_{s_i}}\right) d\lambda_{r_{s_i}} \\ &= \prod_{i=1}^{i-1} \int_{x}^{\infty} \lambda_{r_{s_i}}^{n-m+m-s_i-t_i} \exp\left(-\frac{1}{\sigma^2}\lambda_{r_{s_i}}\right) d\lambda_{r_{s_i}}, \quad \text{(by Lemma 4.1)}, \\ I_2(s,t) &= \sum_{r_{s_2}} \int_{D5} \prod_{i=i}^{m} \lambda_{r_{s_i}}^{n-m+m-s_i-t_i} \exp\left(-\frac{1}{\sigma^2}\lambda_{r_{s_i}}\right) d\lambda_{r_{s_i}} \\ &= \prod_{i=i}^{m} \int_{0}^{x} \lambda_{r_{s_i}}^{n-m+m-s_i-t_i} \exp\left(-\frac{1}{\sigma^2}\lambda_{r_{s_i}}\right) d\lambda_{r_{s_i}}, \quad \text{(by Lemma 4.1)}, \end{aligned}$$

$D_4 = \{x < \lambda_{i-1} < \cdots < \lambda_1 < \infty\}$ and $D_5 = \{0 < \lambda_m < \cdots < \lambda_i < x\}$. The result follows from the definition of the determinant. Note that Lemma 4.2 is used in this proof. □

4.7 Distribution of α

Many scientific problems lead to solving a random system of linear equations. The condition number distribution of this random matrix provides how many digits of numerical precision are lost due to ill conditioning. In addition, if a random system is solved by an iterative technique then the condition number distribution describes the speed of convergence of this iterative method (e.g., conjugate gradient method).

It is well known that the solution to the linear system $Ax = b$ is given by $\hat{x} = (A^TA)^{-1}A^Tb$ and the residual error in this estimation is given by [14]

$$||A(A^TA)^{-1}A^T - I|| \lessapprox 200\alpha^2 n\beta^{-s},$$

where $A \in \mathbb{R}^{n\times n}$, α is the condition number, β is the base and s is the number of digits. This error bound is not accurate and the authors made further assumptions of the form

$$\lambda_{\max} \leq 100n, \;\; \lambda_{\min} \geq 1/100n, \;\text{ and } \alpha \leq 10n,$$

where $\lambda_{\max}$ and $\lambda_{\min}$ are the extreme eigenvalues of the Wishart matrix A^TA.

Other direct methods of solving linear equations are Gaussian elimination with partial pivoting and the QR factorization. In [51], the relative error is bounded by

$$\frac{||x - \hat{x}||}{||x||} \leq \alpha f(n)\epsilon,$$

where ϵ is the floating-point machine epsilon. Note that $f(n)$ is factored as $g(n)p(n)$, where $g(n)$ is the growth factor and $p(n)$ is a polynomial.

Moreover, the condition number can also be defined (see [40] and [12]) as the smallest number α such that for all x and δx, if $Ax = b$ and $A(x + \delta x) = b + \delta b$, then

$$\frac{||\delta x||}{||x||} \leq \alpha \frac{||\delta b||}{||b||}.$$

By taking the logarithm on both sides, we have

$$(\log ||\delta x|| - \log ||x||) - (\log ||\delta b|| - \log ||b||) \leq \log \alpha.$$

This shows that the number of correct digits in x can differ from the number of correct digits in b by at most $\log \alpha$. In [40], the loss of precision is denoted by $\log \alpha$. Problems where α is large are referred to as ill-conditioned and such problems are characterized by very elongated elliptical level sets. Iterative methods converge slowly for these problems [3].

This review clearly demonstrates the importance of the condition number distribution for solving random systems. If $\mathbf{A} \sim \mathcal{CN}(0, I \otimes \Sigma)$ or $\mathbf{A} \sim \mathcal{CN}(0, I \otimes \sigma^2 I)$, then the condition number densities of A and $W = A^H A$ are not available in the literature. We derive these densities in the sequel. First, we derive the joint density of the extreme eigenvalues of the complex central Wishart matrix $W = A^H A$, i.e., $f(\lambda_{\max}, \lambda_{\min})$. This will enable us to compute the density of the condition number α of the random matrix A. Note that the condition number of A is given by

$$\alpha = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}}$$

and the condition number of W is α^2. The following two lemmas are required in the sequel.

Lemma 4.4 *Let* $\Lambda = \operatorname{diag}(\lambda_1, \ldots \lambda_m)$ *and* $D_7 = \{1 > \lambda_1 > \cdots > \lambda_m > 0\}$. *Then*

$$\int_{D_7} (\det \Lambda)^{a-m} \det(I - \Lambda)^{b-m} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 C_\kappa(\Lambda) \bigwedge_{k=1}^{m} d\lambda_k = \frac{\mathcal{C}\Gamma_m(m)}{\pi^{m(m-1)}} \frac{\mathcal{C}\Gamma_m(a,\kappa)\mathcal{C}\Gamma_m(b)}{\mathcal{C}\Gamma_m(a+b,\kappa)} C_\kappa(I). \tag{4.35}$$

Proof. The result follows by letting $Y = I$ and $X = E\Lambda E^H$ in Equation (3.16) and use the differential form

$$(dX) = \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 (d\Lambda)(E^H dE) \quad \text{with} \quad \int_{U(m)} (E^H dE) = \frac{2^m \pi^{m^2}}{\mathcal{C}\Gamma_m(m)}.$$

As mentioned in the proof of Proposition 4.3, we must divide the left side of Equation (3.16) by $(2\pi)^m$. □

Lemma 4.5 *Let* $Z = \operatorname{diag}(\zeta_2, \ldots, \zeta_m)$, $Z_1 = \operatorname{diag}(1, \zeta_2, \ldots, \zeta_m)$ *and* $D_8 = \{1 > \zeta_2 >$

$\cdots > \zeta_m > 0\}$. *Then*

$$\int_{D_8} (\det Z)^{a-m} \prod_{k=2}^{m} (1-\zeta_k)^2 \prod_{k<l}^{m} (\zeta_k - \zeta_l)^2 C_\kappa(Z_1) \bigwedge_{k=2}^{m} d\zeta_k$$
$$= (ma+k)\, \frac{\mathcal{C}\Gamma_m(m)}{\pi^{m(m-1)}} \frac{\mathcal{C}\Gamma_m(a,\kappa)\mathcal{C}\Gamma_m(m)}{\mathcal{C}\Gamma_m(a+m,\kappa)} C_\kappa(I). \tag{4.36}$$

Proof. Let $b = m$ and $\zeta_k = \lambda_k/\lambda_1$, $k = 2,\ldots,m$. Then the left side of Equation (4.35) becomes

$$\int_0^1 \lambda_1^{ma+k-1} d\lambda_1 \int_{D_8} (\det Z)^{a-m} \prod_{k=2}^{m} (1-\zeta_k)^2 \prod_{k<l}^{m} (\zeta_k - \zeta_l)^2 C_\kappa(Z_1) \bigwedge_{k=2}^{m} d\zeta_k. \tag{4.37}$$

The result follows by noting that $\int_0^1 \lambda_1^{ma+k-1} d\lambda_1 = 1/(ma+k)$. □

The following theorem describes the joint density of the extreme eigenvalues of the complex central complex Wishart matrix.

Theorem 4.3 *Let* $\mathbf{W} \sim \mathcal{CW}_m(n,\Sigma)$. *The joint density function of* $\lambda_1 (= \lambda_{\max})$ *and* $\lambda_m (= \lambda_{\min})$ *of* W *is given by*

$$\begin{aligned} f(\lambda_1, \lambda_m) \;=\; & \frac{\pi^{m(m-1)} (\det \Sigma)^{-n}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)} \exp(-m\lambda_1) \sum_{k=0}^{\infty} \sum_{\kappa} \frac{\lambda_1^{mn+k-1} C_\kappa(\Sigma^{-1})}{k!\; C_\kappa(I)} \\ & \cdot \sum_{t=0}^{\infty} \sum_{\tau,\delta} \frac{(m-n)_\tau g^{\delta}_{\tau,\kappa} (1-\lambda_m/\lambda_1)^{(m-1)(m+1)+t+k-1}}{t!} \\ & \cdot [(m-1)(m+1)+k+t] \frac{\mathcal{C}\Gamma_{m-1}(m-1)}{\pi^{(m-1)(m-2)}} \\ & \cdot \frac{\mathcal{C}\Gamma_{m-1}(m+1,\delta)\mathcal{C}\Gamma_{m-1}(m-1)}{\mathcal{C}\Gamma_{m-1}(2m,\delta)} C_\delta(I), \end{aligned} \tag{4.38}$$

where $g^{\delta}_{\tau,\kappa}$ *is the coefficient of* C_δ *(defined in the proof).*

Proof. Consider Equation (4.11). By making the transformations $\lambda_1 = \lambda_1$, $\eta_k =$

$1-\lambda_k/\lambda_1$, $k=2,\ldots,m$, we obtain the joint density of λ_1, $\eta_2,\ldots,\eta_m$, as

$$\frac{\pi^{m(m-1)}(\det\Sigma)^{-n}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)}\exp(-m\lambda_1)\sum_{k=0}^{\infty}\sum_{\kappa}\lambda_1^{mn+k-1}(\det H)^2\det(I-H)^{n-m}$$
$$\cdot\frac{C_\kappa(H)C_\kappa(\Sigma^{-1})}{k!\,C_\kappa(I)}\prod_{i>j=2}^{m}(\eta_i-\eta_j)^2,\quad 0<\lambda_1<\infty,\ 0<\eta_2<\cdots<\eta_m<1, \tag{4.39}$$

where $H=\operatorname{diag}(\eta_2,\ldots,\eta_m)$. We have [28]

$$\begin{aligned}\det(I-H)^{n-m}C_\kappa(H) &= \sum_{t=0}^{\infty}\sum_{\tau}\frac{(-(n-m))_\tau C_\tau(H)C_\kappa(H)}{t!}\\ &= \sum_{t=0}^{\infty}\sum_{\tau}\sum_{\delta}\frac{(-(n-m))_\tau g_{\tau,\kappa}^{\delta}C_\delta(H)}{t!},\end{aligned} \tag{4.40}$$

where $g_{\tau,\kappa}^{\delta}$ is the coefficient of $C_\delta(H)$ in the product $C_\tau(H)C_\kappa(H)$, $\delta=(\delta_1,\ldots,\delta_m)$, $\delta_1\geq\cdots\geq\delta_m\geq 0$ and $\sum_{i=1}^{m}\delta_i=k+t$. Again, by making the transformations $\lambda_1=\lambda_1$, $\zeta_k=\eta_k/\eta_m$, $k=2,\ldots,m-1$ and $\eta_m=\eta_m$, we obtain the joint density of λ_1, $\zeta_2,\ldots,\zeta_{m-1}$, and η_m as

$$\frac{\pi^{m(m-1)}(\det\Sigma)^{-n}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)}\exp(-m\lambda_1)\sum_{k=0}^{\infty}\sum_{\kappa}\frac{\lambda_1^{mn+k-1}C_\kappa(\Sigma^{-1})}{k!\,C_\kappa(I)}$$
$$\cdot\sum_{t=0}^{\infty}\sum_{\tau,\delta}\frac{(m-n)_\tau g_{\tau,\kappa}^{\delta}\eta_m^{(m-1)(m+1)+t+k-1}}{t!}$$
$$\cdot(\det Z)^2C_\delta(Z_1)\prod_{i=2}^{m-1}(1-\zeta_i)^2\prod_{i>j=2}^{m-1}(\zeta_i-\zeta_j)^2, \tag{4.41}$$

where $Z=\operatorname{diag}(\zeta_2,\ldots,\zeta_{m-1})$ and $Z_1=\operatorname{diag}(1,\zeta_2,\ldots,\zeta_{m-1})$. Integrating with respect to $\zeta_2,\ldots,\zeta_{m-1}$ and using Lemma 4.5, we obtain the joint density of λ_1 and η_m as

$$\begin{aligned}g(\lambda_1,\eta_m) &= \frac{\pi^{m(m-1)}(\det\Sigma)^{-n}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)}\exp(-m\lambda_1)\sum_{k=0}^{\infty}\sum_{\kappa}\frac{\lambda_1^{mn+k-1}C_\kappa(\Sigma^{-1})}{k!\,C_\kappa(I)}\\ &\quad\cdot\sum_{t=0}^{\infty}\sum_{\tau,\delta}\frac{(m-n)_\tau g_{\tau,\kappa}^{\delta}\eta_m^{(m-1)(m+1)+t+k-1}}{t!}\,[(m-1)(m+1)+k+t]\end{aligned}$$

$$\cdot\frac{\mathcal{C}\Gamma_{m-1}(m-1)}{\pi^{(m-1)(m-2)}}\frac{\mathcal{C}\Gamma_{m-1}(m+1,\delta)\mathcal{C}\Gamma_{m-1}(m-1)}{\mathcal{C}\Gamma_{m-1}(2m,\delta)}C_\delta(I). \tag{4.42}$$

Finally, the result follows by substituting $\eta_m = 1 - \lambda_m/\lambda_1$. □

In the next theorem, we obtain the density of α^2 which can be used to obtain confidence bounds.

Theorem 4.4 *Let* $\mathbf{W} = \mathbf{A}^H\mathbf{A} \sim \mathcal{CW}_m(n, \Sigma)$. *Since the square of the condition number of the random matrix* $\mathbf{A}$ *is* $\alpha^2 = \lambda_1/\lambda_m$, *then the density of* $y = 1 - 1/\alpha^2$ *is given by*

$$\begin{aligned} f(y) &= \frac{\pi^{m(m-1)}(\det\Sigma)^{-n}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)}\sum_{k=0}^{\infty}\sum_{\kappa}\frac{\Gamma(mn+k)C_\kappa(\Sigma^{-1})}{m^{mn+k}\, k!\, C_\kappa(I)} \\ &\cdot\sum_{t=0}^{\infty}\sum_{\tau,\delta}\frac{(m-n)_\tau g^{\delta}_{\tau,\kappa}y^{(m-1)(m+1)+t+k-1}}{t!}[(m-1)(m+1)+k+t] \\ &\cdot\frac{\mathcal{C}\Gamma_{m-1}(m-1)}{\pi^{(m-1)(m-2)}}\frac{\mathcal{C}\Gamma_{m-1}(m+1,\delta)\mathcal{C}\Gamma_{m-1}(m-1)}{\mathcal{C}\Gamma_{m-1}(2m,\delta)}C_\delta(I). \end{aligned} \tag{4.43}$$

Proof. The result follows by integrating (4.42) with respect to λ_1 and substituting $y = \eta_m$. Note that we have

$$\int_0^\infty e^{-m\lambda_1}\lambda_1^{mn+k-1}\,d\lambda_1 = \frac{\Gamma(mn+k)}{m^{mn+k}}.$$

□

If $\Sigma = \sigma^2 I$, then the corresponding results to Theorems 4.3 and 4.4 can be derived using a similar method. However, we provide an alternative approach as follows:

Theorem 4.5 *Let* $\Sigma = \sigma^2 I$. *The joint density of* $\lambda_1(=\lambda_{\max})$ *and* $\lambda_m(=\lambda_{\min})$ *of the central Wishart matrix is given by*

$$\begin{aligned} f(\lambda_1,\lambda_m) &= \frac{\pi^{m(m-1)}(\sigma^2)^{-nm}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)}\lambda_1^{(m-1)(n-m-1)+m}\exp\left\{-\frac{1}{\sigma^2}[(m-1)\lambda_1-\lambda_m]\right\}\lambda_m^{n-m} \\ &\cdot(\lambda_1-\lambda_m)^{m^2-2}\varrho(\psi; m-2, 2, 0, 1), \quad 0<\lambda_m<\lambda_1<\infty, \end{aligned} \tag{4.44}$$

where

$$\varrho(\psi, m, r, L, U) = \int_{D_6} \prod_{k=1}^{m} (x_k^r \psi(x_k)) \prod_{k>l=1}^{m} (x_k - x_l)^2 \bigwedge_{k=1}^{m} dx_k, \tag{4.45}$$

and $\psi(x) = (1-x)^2(1-x-(\lambda_m/\lambda_1)x)^{n-m} \exp\left(\frac{1}{\sigma^2}(\lambda_1 - \lambda_m)x\right)$ *with* $D_6 = \{L \le x_1 \le \cdots \le x_q \le U\}$.

Proof. Consider Equation (4.15). By making the transformations $\lambda_1 = \lambda_1$, $\eta_k = 1 - \lambda_k/\lambda_1$ $k = 2, \ldots, m$, we obtain the joint density of $\lambda_1, \eta_2, \ldots, \eta_m$ as

$$\frac{\pi^{m(m-1)}(\sigma^2)^{-nm}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)} \lambda_1^{mn-1} \exp\left(-\frac{1}{\sigma^2} m\lambda_1\right) \cdot \prod_{k=2}^{m} \left[\eta_k^2 (1-\eta_k)^{n-m} \exp\left(\frac{1}{\sigma^2}\lambda_1\eta_k\right)\right] \prod_{k>l=2}^{m} (\eta_k - \eta_l)^2, \tag{4.46}$$

where $0 < \lambda_1 < \infty$ and $0 < \eta_2 < \cdots < \eta_m < 1$. Again, by making the transformations $\lambda_1 = \lambda_1$, $\zeta_k = \eta_k/\eta_m$, $k = 2, \ldots, m-1$ and $\eta_m = \eta_m$, we obtain the joint density of λ_1, $\zeta_2, \ldots, \zeta_{m-1}$ and η_m as

$$\frac{\pi^{m(m-1)}(\sigma^2)^{-nm}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)} \lambda_1^{mn-1} \exp\left\{-\frac{1}{\sigma^2}\lambda_1 (m - \eta_m)\right\} \eta_m^{m^2-2} (1-\eta_m)^{n-m} \cdot \prod_{k=2}^{m-1} \left[\zeta_k^2 (1-\zeta_k)^2 (1-\eta_m\zeta_k)^{n-m} \exp\left(\frac{1}{\sigma^2}\lambda_1\eta_m\zeta_k\right)\right] \prod_{k>l=2}^{m-1} (\zeta_k - \zeta_l)^2, \tag{4.47}$$

where $0 < \lambda_1 < \infty$, $0 < \zeta_2 < \cdots < \zeta_{m-1} < 1$ and $0 < \eta_m < 1$. Integrating with respect to $\zeta_2, \ldots, \zeta_{m-1}$, the joint density of λ_1 and η_m, denoted by $g(\lambda_1, \eta_m)$, is given by

$$\frac{\pi^{m(m-1)}(\sigma^2)^{-nm}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)} \lambda_1^{mn-1} \exp\left\{-\frac{1}{\sigma^2}\lambda_1 (m - \eta_m)\right\} \eta_m^{m^2-2} (1-\eta_m)^{n-m} \varrho(\psi; m-2, 2, 0, 1) \tag{4.48}$$

where $\psi(x) = (1-x)^2(1-\eta_m x)^{n-m} \exp\left(\frac{1}{\sigma^2}\lambda_1\eta_m x\right)$, $0 < \lambda_1 < \infty$, and $0 < \eta_m < 1$. Now, the result follows by substituting $\eta_m = 1 - \lambda_m/\lambda_1$. □

Theorem 4.6 *Let* $\mathbf{W} = \mathbf{A}^H\mathbf{A} \sim \mathcal{C}\mathcal{W}_m(n, \sigma^2 I)$*. Since the square of the condition number of* A *is* $\alpha^2 = \lambda_1/\lambda_m$*, then the density of* $y = 1 - 1/\alpha^2$ *is given by*

$$f(y) = \int_0^\infty g(\lambda_1, \eta_m)\, d\lambda_1. \tag{4.49}$$

Proof. The proof is obvious from Equation (4.48). □

It should be noted that the joint density of the extreme eigenvalues of the real central Wishart matrix is studied in [42], [47] and [48].

Chapter 5

Complex Noncentral Wishart Matrix

The complex noncentral Wishart matrix is studied in this chapter. Section 5.1 defines the complex noncentral Wishart distribution. Section 5.2 gives the joint eigenvalue density. The maximum eigenvalue distribution is derived in section 5.3, while, the minimum eigenvalue distribution is derived in section 5.4.

5.1 The complex noncentral Wishart distribution

The definition of the complex noncentral Wishart distribution is given by

Definition 5.1 *Let* $\mathbf{W} = \mathbf{A}^H\mathbf{A}$ *, where the* $n \times m$ *matrix* $\mathbf{A}$ *is distributed as* $\mathbf{A} \sim \mathcal{CN}(M, I_n \otimes \Sigma)$. *Then* $\mathbf{W}$ *is said to have the complex noncentral Wishart distribution with* n *degrees of freedom, covariance matrix* Σ*, and matrix of noncentrality parameters* $\Omega = \Sigma^{-1}M^H M$. *We shall write* $\mathbf{W} \sim \mathcal{CW}_m(n, \Sigma, \Omega)$.

The density of this noncentral Wishart is given by:

Theorem 5.1 *Let* $\mathbf{W} \sim \mathcal{CW}_m(n, \Sigma, \Omega)$ *with* $n \geq m$. *Then the density of* $\mathbf{W}$ *is given by*

$$f(W) = \frac{1}{\mathcal{C}\Gamma_m(n)(\det \Sigma)^n} \operatorname{etr}\left(-\Sigma^{-1}W\right)(\det W)^{n-m} \operatorname{etr}(-\Omega)\, {}_0F_1\left(n; \Omega\Sigma^{-1}W\right), \quad (5.1)$$

where $\Omega = \Sigma^{-1}M^H M$.

Proof. The density of **A** is

$$\pi^{-nm}(\det\Sigma)^{-n}\operatorname{etr}\left(-\Sigma^{-1}A^H A\right)\operatorname{etr}\left(-\Sigma^{-1}M^H M\right)\operatorname{etr}\left(2\Sigma^{-1}M^H A\right)(dA).$$

Let $A = ET$, where E is an $n\times m$ matrix with $E^H E = I_m$ and T is an upper triangular matrix with real and positive diagonal elements. We have $W = A^H A = T^H T$ and from Proposition 4.2 in Chapter 4

$$(dA) = 2^{-m}(\det W)^{n-m}(dW)(E^H dE).$$

Therefore, the density becomes

$$2^{-m}\pi^{-nm}(\det\Sigma)^{-n}\operatorname{etr}\left(-\Sigma^{-1}W\right)\operatorname{etr}(-\Omega)(\det W)^{n-m}\operatorname{etr}\left(2\Sigma^{-1}M^H ET\right)(dW)(E^H dE).$$

Integrating with respect to E over the Stiefel manifold $\mathcal{C}V_{m,n}$ and using Proposition 3.7 we obtain

$$\int_{E\in\mathcal{C}V_{m,n}}\operatorname{etr}\left(2\Sigma^{-1}M^H ET\right)(E^H dE) = \frac{2^m\pi^{nm}}{\mathcal{C}\Gamma_m(n)}\,{}_0F_1(n;\Omega\Sigma^{-1}W).$$

The result follows. □

Corollary 5.1 *The complex noncentral Wishart density can be expressed in terms of the complex central Wishart density, i.e.,*

$$\mathcal{C}W_m(n,\Sigma,\Omega) = \mathcal{C}W_m(n,\Sigma,0)\sum_{k,l=0}^{\infty}\sum_{\kappa,\tau;\phi\in\kappa.\tau}\frac{\theta_\phi^{\kappa,\tau}C_\phi^{\kappa,\tau}(-\Omega,\Omega\Sigma^{-1}W)}{k!\;l!\;[n]_\tau},\tag{5.2}$$

where $C_\phi^{\kappa,\tau}$ *is an invariant polynomial, indexed by the ordered partitions* κ, τ *and* ϕ *of the nonnegative integers* k, l, *and* $f = k + l$, *respectively, into no more than* m *parts and* $\theta_\phi^{\kappa,\tau} = C_\phi^{\kappa,\tau}(I,I)/C_\phi(I)$.

Proof. We have

$$
\begin{aligned}
\mathcal{C}\mathcal{W}_m(n,\Sigma,\Omega) &= \mathcal{C}\mathcal{W}_m(n,\Sigma,0)\,\mathrm{etr}(-\Omega)_0F_1(n;\Omega\Sigma^{-1}W) \\
&= \mathcal{C}\mathcal{W}_m(n,\Sigma,0)\sum_{k,l=0}^{\infty}\sum_{\kappa,\tau}\frac{C_\kappa(-\Omega)}{k!}\frac{C_\tau(\Omega\Sigma^{-1}W)}{l!\,[n]_\tau} \\
&= \mathcal{C}\mathcal{W}_m(n,\Sigma,0)\sum_{k,l=0}^{\infty}\sum_{\kappa,\tau;\phi\in\kappa.\tau}\frac{\theta_\phi^{\kappa,\tau}C_\phi^{\kappa,\tau}(-\Omega,\Omega\Sigma^{-1}W)}{k!\;l!\;[n]_\tau}. \qquad (5.3)
\end{aligned}
$$

The result follows from Section 3.3. □

5.2 Joint eigenvalue distribution

The eigenvalue density of a complex noncentral Wishart matrix is given by the following theorem.

Theorem 5.2 *Let* $\mathbf{W}\sim\mathcal{C}\mathcal{W}_m(n,\Sigma,\Omega)$ *with* $n>m-1$. *Then* $\mathbf{W}$ *is an* $m\times m$ *positive definite Hermitian matrix with real eigenvalues. The joint density of the eigenvalues,* $\lambda_1,\ldots,\lambda_m$, *of* $\mathbf{W}$ *is*

$$
\begin{aligned}
f(\Lambda) = &\frac{\pi^{m(m-1)}(\det\Sigma)^{-n}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)}\,\mathrm{etr}(-\Omega)\prod_{k=1}^{m}\lambda_k^{n-m}\prod_{k<l}^{m}(\lambda_k-\lambda_l)^2 \\
&\sum_{k,l=0}^{\infty}\sum_{\kappa,\tau;\phi\in\kappa.\tau}\frac{C_\phi^{\kappa,\tau}(-\Sigma^{-1},\Omega\Sigma^{-1})C_\phi^{\kappa,\tau}(\Lambda,\Lambda)}{k!\;l!\;[n]_\tau\;C_\phi(I_m)}, \qquad (5.4)
\end{aligned}
$$

where $C_\phi^{\kappa,\tau}$ *is an invariant polynomial, indexed by the ordered partitions* κ, τ *and* ϕ *of the nonnegative integers* k, l, *and* $f=k+l$, *respectively, into no more than* m *parts.*

Proof. From Proposition 4.3, we obtain

$$
\begin{aligned}
f(\Lambda) &= \frac{\pi^{m(m-1)}}{\mathcal{C}\Gamma_m(m)}\prod_{k<l}^{m}(\lambda_k-\lambda_l)^2\int_{U(m)}f(E\Lambda E^H)(dE) \\
&= \frac{\pi^{m(m-1)}(\det\Sigma)^{-n}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)}\,\mathrm{etr}(-\Omega)\prod_{k=1}^{m}\lambda_k^{n-m}\prod_{k<l}^{m}(\lambda_k-\lambda_l)^2
\end{aligned}
$$

$$
\begin{aligned}
&\cdot \int_{U(m)} \operatorname{etr}\left(-\Sigma^{-1} E \Lambda E^H\right) {}_0F_1\left(n; \Omega\Sigma^{-1} E \Lambda E^H\right) (dE) \\
= &\ \frac{\pi^{m(m-1)} (\det \Sigma)^{-n}}{C\Gamma_m(m) C\Gamma_m(n)} \operatorname{etr}(-\Omega) \prod_{k=1}^{m} \lambda_k^{n-m} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 \\
&\cdot \sum_{k,l=0}^{\infty} \sum_{\kappa,\tau} \frac{1}{k!\ l!\ [n]_\tau} \int_{U(m)} C_\kappa(-\Sigma^{-1} E \Lambda E^H) C_\tau(\Omega\Sigma^{-1} E \Lambda E^H)(dE). \quad (5.5)
\end{aligned}
$$

The result follows from Equation (3.40). □

5.3 Distribution of $\lambda_{\max}$

The distribution of the largest eigenvalue of a complex noncentral Wishart matrix is derived from the following theorem.

Theorem 5.3 *Let* $\mathbf{W} \sim \mathcal{CW}_m(n, \Sigma, \Omega)$ $(n > m-1)$ *and* Δ *be an* $m \times m$ *positive definite matrix. Then the probability* $P(W < \Delta)$ *is given by*

$$
P(W < \Delta) = \frac{C\Gamma_m(m) (\det \Delta)^n \operatorname{etr}(-\Omega)}{C\Gamma_m(n+m) (\det \Sigma)^n} \sum_{k,l=0}^{\infty} \sum_{\kappa,\tau;\phi\in\kappa.\tau} \frac{[n]_\phi \theta_\phi^{\kappa,\tau} C_\phi^{\kappa,\tau}(-\Sigma^{-1}\Delta, \Omega\Sigma^{-1}\Delta)}{k!\ l!\ [n]_\tau\ [n+m]_\phi}, \quad (5.6)
$$

where $C_\phi^{\kappa,\tau}$ *is an invariant polynomial, indexed by the ordered partitions* κ, τ *and* ϕ *of the nonnegative integers* k, l, *and* $f = k + l$, *respectively, into no more than* m *parts and* $\theta_\phi^{\kappa,\tau} = C_\phi^{\kappa,\tau}(I, I)/C_\phi(I)$.

Proof. Using the noncentral Wishart density (5.1) we can write the probability $P(W < \Delta)$ as

$$
\begin{aligned}
P(W < \Delta) &= \frac{\operatorname{etr}(-\Omega)}{C\Gamma_m(n)(\det \Sigma)^n} \int_0^\Delta \operatorname{etr}(-\Sigma^{-1} W)(\det W)^{n-m} {}_0F_1(n; \Omega\Sigma^{-1} W)(dW) \\
&= \frac{\operatorname{etr}(-\Omega)}{C\Gamma_m(n)(\det \Sigma)^n} \sum_{l=0}^{\infty} \sum_{\tau} \int_0^\Delta \frac{\operatorname{etr}(-\Sigma^{-1} W)(\det W)^{n-m} C_\tau(\Omega\Sigma^{-1} W)}{l!\ [n]_\tau} (dW).
\end{aligned}
$$

The result follows from Equation (3.42). □

The following result describes the distribution of $\lambda_{\max}$, which follows from Theorem 5.3.

Corollary 5.2 *If $W \sim \mathcal{CW}_m(n, \Sigma, \Omega)$ and $\lambda_{\max}$ is the largest eigenvalue of W, then its distribution is given by*

$$P(\lambda_{\max} < y) = \frac{y^{mn} \mathcal{C}\Gamma_m(m)\,\mathrm{etr}(-\Omega)}{\mathcal{C}\Gamma_m(n+m)(\det \Sigma)^n} \sum_{k,l=0}^{\infty} \sum_{\kappa,\tau;\phi\in\kappa.\tau} \frac{[n]_\phi \theta_\phi^{\kappa,\tau} C_\phi^{\kappa,\tau}(-y\Sigma^{-1}, y\Omega\Sigma^{-1})}{k!\; l!\; [n]_\tau\; [n+m]_\phi}, \quad (5.7)$$

where $C_\phi^{\kappa,\tau}$ is an invariant polynomial, indexed by the ordered partitions κ, τ and ϕ of the nonnegative integers k, l, and $f = k + l$, respectively, into no more than m parts and $\theta_\phi^{\kappa,\tau} = C_\phi^{\kappa,\tau}(I, I)/C_\phi(I)$.

Proof. The inequality $\lambda_{\max} < y$ is equivalent to $W < yI$. Therefore, the result follows by letting $\Delta = yI$ in Theorem 5.3. □

5.4 Distribution of $\lambda_{\min}$

The distribution of the smallest eigenvalue of a complex noncentral Wishart matrix is derived from the following theorem.

Theorem 5.4 *Let $W \sim \mathcal{CW}_m(n, \Sigma, \Omega)$ $(n > m-1)$ and Δ be an $m \times m$ positive definite matrix. Then the probability $P(W > \Delta)$ can be written as*

$$P(W > \Delta) = \frac{\mathrm{etr}(-\Omega)\,\mathrm{etr}\,(-\Sigma^{-1}\Delta)\,(\det \Delta)^n}{\mathcal{C}\Gamma_m(n)(\det \Sigma)^n} \sum_{l=0}^{\infty} \sum_{\tau} \frac{1}{l!\; [n]_\tau}$$
$$\cdot \int_{X>0} \mathrm{etr}(-\Delta^{1/2}\Sigma^{-1}\Delta^{1/2}X) \det(I+X)^{n-m} C_\tau(\Omega\Sigma^{-1}\Delta^{1/2}(I+X)\Delta^{1/2})(dX), \quad (5.8)$$

where τ is a partition of l.

Proof. Using the noncentral Wishart density (5.1) we can write the probability $P(W > \Delta)$ as

$$P(W > \Delta) = \frac{\text{etr}(-\Omega)}{\mathcal{C}\Gamma_m(n)(\det \Sigma)^n} \int_{W>\Delta} \text{etr}\left(-\Sigma^{-1}W\right)(\det W)^{n-m} {}_0F_1\left(n; \Omega\Sigma^{-1}W\right)(dW). \tag{5.9}$$

The change of variable $W = \Delta^{1/2}(I + X)\Delta^{1/2}$ leads to the differential form $(dW) = (\det \Delta)^m (dX)$. Hence,

$$P(W > \Delta) = \frac{\text{etr}(-\Omega)\,\text{etr}(-\Sigma^{-1}\Delta)(\det \Delta)^n}{\mathcal{C}\Gamma_m(n)(\det \Sigma)^n}$$
$$\cdot \int_{X>0} \text{etr}\left(-\Delta^{1/2}\Sigma^{-1}\Delta^{1/2}X\right)\det(I + X)^{n-m} {}_0F_1(n; \Omega\Sigma^{-1}\Delta^{1/2}(I + X)\Delta^{1/2})(dX).$$

The result follows by expanding ${}_0F_1(n; \Omega\Sigma^{-1}\Delta^{1/2}(I + X)\Delta^{1/2})$. □

The distribution of the smallest eigenvalue is given in the following corollary.

Corollary 5.3 *If* $\mathbf{W} \sim \mathcal{CW}_m(n, \Sigma, \Omega)$ *and* $\lambda_{\min}$ *is the smallest eigenvalue of* $\mathbf{W}$*, then*

$$P(\lambda_{\min} > y) = \frac{y^{mn}\,\text{etr}(-\Omega)\,\text{etr}\left(-y\Sigma^{-1}\right)}{\mathcal{C}\Gamma_m(n)(\det \Sigma)^n} \sum_{l=0}^{\infty}\sum_{\tau} \frac{1}{l!\,[n]_\tau}$$
$$\cdot \int_{X>0} \text{etr}(-y\Sigma^{-1}X)\det(I + X)^{n-m} C_\tau(y\Omega\Sigma^{-1}(I + X))(dX), \tag{5.10}$$

where τ *is a partition of* l.

Proof. The inequality $\lambda_{\min} > y$ is equivalent to $W > yI$. Therefore, the result follows by letting $\Delta = yI$ in Theorem 5.4. □

Chapter 6

The Channel Capacity

In this and the subsequent chapter, we investigate a multiple input, multiple output communication system over complex additive Gaussian noise with Rayleigh and Rician distributed channels. We derive formulas for the capacities and error exponents of such channels, and describe computational methods to evaluate such formulas. As mentioned in Chapter 1, we deal exclusively with a linear model in which the output vector $y \in \mathbb{C}^{n_r}$ depends on the input vector $x \in \mathbb{C}^{n_t}$ via the linear system

$$y = Hx + v, \tag{6.1}$$

where H is an $n_r \times n_t$ complex matrix and v is a complex random vector noise with mean μ_v and covariance matrix R_{vv}, i.e., $\mathbf{v} \sim (\mu_v, R_{vv})$. The total power of the input is constrained to ρ,

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

We shall consider the cases where the matrix H is either deterministic or random. If H is a random matrix, it will be denoted in bold letter by $\mathbf{H}$. This chapter is organized as follows. Section 6.1 defines the differential entropy. Section 6.2 studies the case where H is a deterministic matrix. The corresponding error exponent is given in section 6.3. Finally, in section 6.4, the case where H is a random matrix is studied.

6.1 Differential entropy

Definition 6.1 *The differential entropy $\mathcal{H}(\mathbf{x})$ of a continuous random variable $\mathbf{x}$ with density $f(x)$ is defined as [2]*

$$\begin{aligned} \mathcal{H}(\mathbf{x}) = \mathcal{H}(f) &= -\int_S f(x) \log f(x) dx \\ &= \mathcal{E}\{-\log f(x)\}, \end{aligned} \tag{6.2}$$

where S is the support of the random variable.

Proposition 6.1 *The differential entropy of the complex random vector $\mathbf{x} \sim \mathcal{CN}(\mu_x, R_{xx})$ is given by*

$$\mathcal{H}(f) = \log\det(\pi e R_{xx}). \tag{6.3}$$

Proof. The probability density of a complex Gaussian random vector $\mathbf{x} \in \mathbb{C}^{n_t}$ with mean μ_x and covariance R_{xx} is given by

$$\begin{aligned} f(x|\mu_x, R_{xx}) &= [\pi^{n_t} \det(R_{xx})]^{-1} \exp\left\{-(x-\mu_x)^H R_{xx}^{-1}(x-\mu_x)\right\} \\ &= [\det(\pi R_{xx})]^{-1} \exp\left\{-(x-\mu_x)^H R_{xx}^{-1}(x-\mu_x)\right\}. \end{aligned} \tag{6.4}$$

From Definition 6.1, the differential entropy $\mathcal{H}(f)$ can be written as

$$\begin{aligned} \mathcal{H}(f) &= \mathcal{E}\{-\log f(x|\mu_x, R_{xx})\} \\ &= \log\det(\pi R_{xx}) + (\log e)\mathcal{E}\left\{(x-\mu_x)^H R_{xx}^{-1}(x-\mu_x)\right\} \\ &= \log\det(\pi R_{xx}) + (\log e)\,\mathrm{tr}\left(\mathcal{E}\left\{(x-\mu_x)(x-\mu)^H\right\} R_{xx}^{-1}\right) \\ &= \log\det(\pi R_{xx}) + (\log e)\,\mathrm{tr}\, I \\ &= \log\det(\pi e R_{xx}). \end{aligned} \tag{6.5}$$

□

Lemma 6.1 *Let $f(x)$ and $g(x)$ be arbitrary probability density functions.*

(a) *If* $-\int_{\mathbb{C}^{n_t}} f(x)\log g(x)dx$ *is finite, then* $-\int_{\mathbb{C}^{n_t}} f(x)\log f(x)dx$ *exists, and furthermore*

$$-\int_{\mathbb{C}^{n_t}} f(x)\log f(x)dx \leq -\int_{\mathbb{C}^{n_t}} f(x)\log g(x)dx \tag{6.6}$$

with equality iff $f(x) = g(x)$ *for almost all* x *with respect to Lebesgue measure.*

(b) *If* $-\int_{\mathbb{C}^{n_t}} f(x)\log f(x)dx$ *is finite, then* $-\int_{\mathbb{C}^{n_t}} f(x)\log g(x)dx$ *exists, and Equation* (6.6) *holds.*

Proof. We have $f(x)\log\left[\frac{g(x)}{f(x)}\right] \leq g(x) - f(x)$, with equality iff $f(x) = g(x)$, because, $\log\alpha \leq \alpha - 1$ with equality iff $\alpha = 1$. Therefore

$$\int_{\mathbb{C}^{n_t}} f(x)\log\frac{g(x)}{f(x)}dx \leq \int_{\mathbb{C}^{n_t}} g(x)dx - \int_{\mathbb{C}^{n_t}} f(x)dx = 1 - 1 = 0 \tag{6.7}$$

with equality iff $f(x) = g(x)$ for almost all x. Now

$$-f(x)\log f(x) = f(x)\log\frac{g(x)}{f(x)} - f(x)\log g(x). \tag{6.8}$$

It follows that

$$-\int_{\mathbb{C}^{n_t}} f(x)\log f(x)dx \leq -\int_{\mathbb{C}^{n_t}} f(x)\log g(x)dx. \tag{6.9}$$

If equality holds in Equation (6.6), then the finiteness of $-\int_{\mathbb{C}^{n_t}} f(x)\log g(x)dx$ allows us to conclude from Equation (6.8) that

$$-\int_{\mathbb{C}^{n_t}} f(x)\log\frac{g(x)}{f(x)}dx = 0 \tag{6.10}$$

and hence that $f(x) = g(x)$ for almost all x. The argument for (b) is quite similar. □

The following proposition demonstrates the importance of complex Gaussian random vectors.

Proposition 6.2 *Let* $\mathbf{x}$ *be an* n_t*-dimensional absolutely continuous complex random vector with density* $f(x)$ *(*$\mathbf{x} \sim (\mu_x, R_{xx})$*). If* $\mathbf{x}$ *has positive definite covariance* R_{xx}

and mean μ_x*, then* $\mathcal{H}(\mathbf{x})$ *exists, and* $\mathcal{H}(\mathbf{x}) \leq \log\det(\pi e R_{xx})$*, with equality if and only if* $\mathbf{x}$ *is a complex Gaussian random vector with* $\mathbf{x} \sim \mathcal{CN}(\mu_x, R_{xx})$*.*

Proof. Let $\mathbf{x}$ be an arbitrary random variable with density $f(x)$, mean μ_x, and covariance matrix R_{xx}. Let $g(x) = [\det(\pi R_{xx})]^{-1}\exp\left\{-(x-\mu_x)^H R_{xx}^{-1}(x-\mu_x)\right\}$. But

$$
\begin{aligned}
& - \int_{\mathbb{C}^{n_t}} f(x)\log g(x)dx \\
&= \int_{\mathbb{C}^{n_t}} f(x)\left[\log\det(\pi R_{xx}) + \log e \operatorname{tr}\left(\left\{(x-\mu_x)(x-\mu_x)^H\right\} R_{xx}^{-1}\right)\right] dx \\
&= \log\det(\pi e R_{xx}).
\end{aligned} \tag{6.11}
$$

The result now follows from Lemma 6.1. □

It should be noted that, for a scalar real random variable, the corresponding results for Lemma 6.1 and Proposition 6.2 are given in [2].

6.2 Known deterministic matrix channel

In this section, the channel capacity for a known deterministic matrix channel H is derived. First, we define the mutual information between input x and output y of the channel and then the channel capacity.

Definition 6.2 *The information processed by the channel or mutual information* $\mathcal{I}(\mathbf{x};\mathbf{y})$ *is defined as [2]*

$$
\begin{aligned}
\mathcal{I}(\mathbf{x};\mathbf{y}) &= \mathcal{H}(\mathbf{x}) - \mathcal{H}(\mathbf{x}|\mathbf{y}) \\
&= \mathcal{H}(\mathbf{y}) - \mathcal{H}(\mathbf{y}|\mathbf{x}) \\
&= \mathcal{H}(\mathbf{y}) - \mathcal{H}(\mathbf{v}),
\end{aligned} \tag{6.12}
$$

where $\mathcal{H}(\mathbf{x}|\mathbf{y}) = -\int f(x,y)\log f(x|y)$ *and* $y = Hx + v$*.*

It is important to notice that the information processed by a channel depends on the input distribution $f(x)$. We may vary the input distribution until the information reaches a maximum; the maximum information is called the *channel capacity.*

Definition 6.3 *The channel capacity C is defined as [2]*

$$C = \max_{f(x)} \mathcal{I}(\mathbf{x};\mathbf{y}).$$

From Proposition 6.2, the maximum entropy is achieved when the distribution is a complex Gaussian. Therefore, in the next proposition we assume that $\mathbf{x}$ and $\mathbf{v}$ are complex Gaussian vectors [45]. Hence $\mathbf{y}$ is a complex Gaussian vector.

Proposition 6.3 *Consider the model* (6.1). *If the n_t-dimensional complex random vector $\mathbf{x}$ is distributed as $\mathbf{x} \sim \mathcal{CN}(\mu_x, R_{xx})$ and the n_r-dimensional complex random vector $\mathbf{v}$ is distributed as $\mathbf{v} \sim \mathcal{CN}(\mu_v, R_{vv})$, then the n_r-dimensional complex random vector is distributed as $\mathbf{y} \sim \mathcal{CN}(\mu_y, HR_{xx}H^H + R_{vv})$ and the mutual information is given by*

$$\mathcal{I}(\mathbf{x};\mathbf{y}) = \log\left(\frac{\det(R_{vv} + HR_{xx}H^H)}{\det R_{vv}}\right). \tag{6.13}$$

Moreover, if the transmitter (or input) has perfect channel knowledge and $R_{vv} = I$, then $\max_{R_{xx}} \mathcal{I}(x;y)$ subject to $\mathrm{tr}(R_{xx}) \le \rho$ is given by

$$C = \sum_{i}^{n_t} \log(\mu\lambda_i)^+, \tag{6.14}$$

where μ is chosen to meet the power constraint and a^+ denotes $\max\{0, a\}$.

Proof. From Definition 6.2, we have

$$\begin{aligned}\mathcal{I}(\mathbf{x};\mathbf{y}) &= \mathcal{H}(\mathbf{y}) - \mathcal{H}(\mathbf{v}) \\ &= \log\det\left(\pi e\left(R_{vv} + HR_{xx}H^H\right)\right) - \log\det(\pi e R_{vv}) \\ &= \log\left(\frac{\det(R_{vv} + HR_{xx}H^H)}{\det R_{vv}}\right).\end{aligned} \tag{6.15}$$

If $R_{vv} = I_{n_r}$, then

$$\mathcal{I}(\mathbf{x};\mathbf{y}) = \log\det\left(I_{n_r} + HR_{xx}H^H\right) = \log\det\left(I_{n_t} + R_{xx}H^H H\right). \tag{6.16}$$

Let $H^H H = E^H \Lambda E$, where $E \in U(n_t)$ and $\Lambda = (\lambda_1, \ldots, \lambda_{n_t})$. Now, we have

$$\begin{aligned}
\det\left(I_{n_r} + H R_{xx} H^H\right) &= \det\left(I_{n_t} + \Lambda^{1/2} E R_{xx} E^H \Lambda^{1/2}\right) \\
&= \det\left(I_{n_t} + \Lambda^{1/2} \tilde{R}_{xx} \Lambda^{1/2}\right) \\
&\leq \prod_{i=1}^{n_t} \left[1 + \left(\tilde{R}_{xx}\right)_{ii} \lambda_i\right], \qquad (6.17)
\end{aligned}$$

where, $\tilde{R}_{xx} = E R_{xx} E^H$, $R_{xx} \geq 0 \Rightarrow \tilde{R}_{xx} = E R_{xx} E^H \geq 0$ and $\operatorname{tr} \tilde{R}_{xx} = \operatorname{tr} R_{xx}$. Equality in (6.17) occurs when $\tilde{R}_{xx}$ is diagonal. Therefore,

$$\arg\max_{R_{xx}} \mathcal{I}(\mathbf{x}; \mathbf{y}) = \operatorname{diag}\left(\left(\tilde{R}_{xx}\right)_{11}, \ldots, \left(\tilde{R}_{xx}\right)_{n_t n_t}\right) = \operatorname{diag}\left((\mu - \lambda_1^{-1})^+, \ldots, (\mu - \lambda_{n_t}^{-1})^+\right)$$

and the capacity C is

$$C = \log \prod_{i=1}^{n_t} \left[1 + \lambda_i \left(\mu - \lambda_i^{-1}\right)^+\right] = \sum_{i=1}^{n_t} \log\left(\mu \lambda_i\right)^+,$$

where μ is chosen to satisfy $\sum_{i=1}^{n_t} \left(\tilde{R}_{xx}\right)_{ii} = \sum_{i=1}^{n_t} \left(\mu - \lambda_i^{-1}\right)^+ = \rho$. □

6.3 Error exponents for deterministic channel

Recall that the capacity of a channel expresses the maximum rate at which information can be reliably conveyed by the channel. Knowing the capacity is not always sufficient, because it is difficult to get close to the maximum rate. Error exponents provide a partial answer to this difficulty by giving an upper bound to the probability of the error achievable by block codes of a given length n and rate $\mathcal{R}$. The upper bound is known as the random coding bound [13].

Definition 6.4 *Let n be a block codes length and $\mathcal{R}$ be a rate. Then the upper bound to the probability of error is given by [13], [45]*

$$P(\text{error}) \leq \exp\{-nE(\mathcal{R})\}. \qquad (6.18)$$

The random coding exponent $E(\mathcal{R})$ is defined by

$$E(\mathcal{R}) = \max_{0\leq\alpha\leq 1}\left[-\alpha\mathcal{R} + \max_{f(x),\mathrm{tr}(R_{xx})\leq\rho} E_0(\alpha, f(x))\right], \tag{6.19}$$

where $f(x)$ is the input distribution,

$$E_0(\alpha, f(x)) = -\log\int\left[\int f(x)p(y|x)^{1/(1+\alpha)}\,dx\right]^{1+\alpha} dy, \tag{6.20}$$

and $p(y|x)$ is the transition probability.

Shannon has shown in [39] that by choosing codes appropriately, the error probability can be made to approach zero exponentially with increasing block length n for any rate less than the capacity, $\mathcal{R} < C$. The following proposition gives the random coding bound for a known deterministic channel matrix H [45].

Proposition 6.4 *Let $x \sim \mathcal{CN}(0, R_{xx})$ and*

$$p(y|x) = \det(\pi I)\exp\left\{-(y - Hx)^H(y - Hx)\right\}.$$

Then

$$E_0(\alpha, f(x)) = \alpha\log\det\left(I + (1+\alpha)^{-1}HR_{xx}H^H\right) \tag{6.21}$$

and

$$E_0(\alpha) = \max_{f(x),\mathrm{tr}(R_{xx})\leq\rho} E_0(\alpha, f(x)) = \alpha\sum_{i=1}^{n_t}\log\left(\mu(1+\alpha)^{-1}\lambda_i\right)^+. \tag{6.22}$$

The upper bound to the probability of error is given by

$$P(\text{error}) \leq \exp\left\{-n\left[\max_{0\leq\alpha\leq 1}(E_0(\alpha) - \alpha R)\right]\right\}.$$

Proof. Let $g(x, y) = f(x)p(y|x)^{1/a}$ and $a = 1 + \alpha$. Then we have

$$g(x,y) = \frac{1}{\det(\pi R_{xx})}\left(\frac{1}{\det(\pi I_{n_r})}\right)^{1/a}\exp - \left\{x^H R_{xx}^{-1}x + (y - Hx)^H(y - Hx)/a\right\}$$

$$= \frac{1}{\det(\pi R_{xx})}\left(\frac{1}{\det(\pi I_{n_r})}\right)^{1/a} \exp -\frac{1}{a}\left\{\left(x - \left(aR_{xx}^{-1}+H^H H\right)^{-1} H^H y\right)^H \right.$$
$$\cdot \left(aR_{xx}^{-1}+H^H H\right)\left(x - \left(aR_{xx}^{-1}+H^H H\right)^{-1} H^H y\right) +$$
$$\left. y^H\left(I - H\left(aR_{xx}^{-1}+H^H H\right)^{-1} H^H\right) y\right\}. \tag{6.23}$$

Now Equation (6.20) is written as

$$E_0(\alpha, f(x)) = -\log\left(\frac{1}{\det\left(I + a^{-1} H R_{xx} H^H\right)^a}\,\frac{1}{\det(I - H\left(aR_{xx}^{-1}+H^H H\right)^{-1} H^H)}\right).$$

Using the matrix identity $(A+BCD)^{-1} = A^{-1} - A^{-1}B\left(C^{-1}+DA^{-1}B\right)^{-1} DA^{-1}$, we obtain

$$\left(I + a^{-1} H R_{xx} H^H\right)^{-1} = (I - H\left(aR_{xx}^{-1}+H^H H\right)^{-1} H^H).$$

Hence, Equation (6.21) follows. The maximization problem (6.22) is similar to Proposition 6.3. Therefore, we obtain the following

$$E_0(\alpha) = \alpha \sum_{i=1}^{n_t} \log\left(\mu(1+\alpha)^{-1}\lambda_i\right)^+.$$

□

6.4 Random matrix channel

Here we assume that the matrix $\mathbf{H} \in \mathbb{C}^{n_r \times n_t}$ is a complex random matrix. Since the receiver knows the realization of $\mathbf{H}$, the channel output is the pair $(\mathbf{y}, \mathbf{H}) = (\mathbf{Hx}+\mathbf{v}, \mathbf{H})$. Note that the transmitter does not know the channel. The mutual information between input and output is then given by

$$\mathcal{I}(\mathbf{x}; (\mathbf{y}, H)) = \mathcal{E}_{\mathbf{H}}\{\mathcal{I}(\mathbf{x}; \mathbf{y}|\mathbf{H} = H)\}. \tag{6.24}$$

Maximizing this mutual information with respect to the input distribution is given by the following proposition [45].

Proposition 6.5 *Let* $\mathbf{H} \in \mathbb{C}^{n_r \times n_t}$ *be a random matrix channel. The capacity of this channel is achieved when* $\mathbf{x}$ *is complex Gaussian with covariance* $(\rho/n_t)I_{n_t}$*. The capacity is given by*

$$\mathcal{E}_{\mathbf{H}}\left\{\log\det\left(I_{n_r} + (\rho/n_t)\mathbf{H}\mathbf{H}^H\right)\right\} \quad or \quad \mathcal{E}_{\mathbf{H}}\left\{\log\det\left(I_{n_t} + (\rho/n_t)\mathbf{H}^H\mathbf{H}\right)\right\}. \quad (6.25)$$

Proof. Just a sketch of the proof is given here. See [45] for details. From Proposition 6.3, the maximum mutual information

$$\mathcal{I}(\mathrm{x}; (\mathrm{y}, \mathbf{H})) = \mathcal{E}_{\mathbf{H}}\left\{\log\det\left(I_{n_r} + \mathbf{H}R_{xx}\mathbf{H}^H\right)\right\}$$

is achieved if the input distribution is a complex Gaussian. This leads to the following problem:

$$\begin{array}{ll} \max_{R_{xx}} & \mathcal{E}_{\mathbf{H}}\left\{\log\det\left(I_{n_r} + \mathbf{H}R_{xx}\mathbf{H}^H\right)\right\} \\ \text{subject to} & \mathrm{tr}(R_{xx}) \le \rho. \end{array} \quad (6.26)$$

Let $R_{xx} = E\Lambda E^H$, where $E \in U(n_t)$. Hence the distribution of $\mathbf{H}$ and $\mathbf{H}E$ are the same. Therefore,

$$\mathcal{E}_{\mathbf{H}}\left\{\log\det\left(I_{n_r} + (\mathbf{H}E)\Lambda(\mathbf{H}E)^H\right)\right\} = \mathcal{E}_{\mathbf{H}}\left\{\log\det\left(I_{n_r} + \mathbf{H}\Lambda\mathbf{H}^H\right)\right\}.$$

It is now clear that the optimum solution to Problem (6.26) is achieved when $R_{xx} = (\rho/n_t)I$. □

In the calculation of capacity, the following two cases are considered:

(*i*) $n_r \ge n_t$ – In this case, the distribution of the channel matrix is given by $\mathbf{H} \sim \mathcal{CN}(M_1, I_{n_r} \otimes \Sigma_1)$. Therefore, the distribution of an $n_t \times n_t$ Wishart matrix is given by $\mathbf{W} = \mathbf{H}^H\mathbf{H} \sim \mathcal{CW}_{n_t}(n_r, \Sigma_1, \Omega)$ with $\Omega = \Sigma_1^{-1}M_1^H M_1$. Here the covariance matrix of the rows of $\mathbf{H}$ is denoted by Σ_1, which is an $n_t \times n_t$ Hermitian matrix.

(*ii*) $n_t \geq n_r$ – In this case, the distribution of the channel matrix is given by $\mathbf{H}^H \sim \mathcal{CN}(M_2, I_{n_t} \otimes \Sigma_2)$. Therefore, the distribution of an $n_r \times n_r$ Wishart matrix is given by $\mathbf{W} = \mathbf{H}\mathbf{H}^H \sim \mathcal{CW}_{n_r}(n_t, \Sigma_2, \Omega)$ with $\Omega = \Sigma_2^{-1} M_2^H M_2$. Note that the covariance matrix of the columns of $\mathbf{H}$ is denoted by Σ_2, which is an $n_r \times n_r$ Hermitian matrix.

We mainly consider case (*i*), $n_r \geq n_t$. Case (*ii*) is a straightforward extension of case (*i*).

Chapter 7

Computation of the Capacity

In this chapter, we evaluate the channel capacity for both correlated and uncorrelated Rayleigh fading channels and Rician fading channels. Section 7.1 studies the correlated Rayleigh channel, while, in section 7.2, the uncorrelated Rayleigh channel is investigated. The Rician channel capacity is evaluated in section 7.3. Finally, the error exponent for a random matrix channel is given in section 7.4.

7.1 Correlated Rayleigh channel

In a correlated Rayleigh channel, the distribution of an $n_r \times n_t$ channel matrix $\mathbf{H}$ is given by $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, I_{n_r} \otimes \Sigma_1)$, with $n_r \geq n_t$. Note that the off diagonal elements of an $n_t \times n_t$ Hermitian matrix Σ_1 are non zero for correlated channels. The following lemma is required in the sequel.

Lemma 7.1 *If X is an $n \times m$ $(n \geq m)$ full rank matrix and the function $f(X)$ depends on X through $X^H X$, then*

$$\int_{X^H X = A} f(X^H X)(dX) = \frac{\pi^{nm}}{\mathcal{C}\Gamma_m(n)} (\det A)^{n-m} f(A).$$

Proof. Since $X^H X = A$, we have

$$\int_{X^H X = A} f(X^H X)(dX) = f(A) \int_{X^H X = A} (dX).$$

Let $X = ET$ and $A = T^H T$, where $E^H E = I_m$ and T is an upper triangular matrix with real and positive diagonal elements. Then, from Propositions 2.4 and 2.2 we have

$$(dX) = \prod_{k=1}^{m} t_{kk}^{2n-2k+1}(dT)(E^H dE)$$

and

$$(dA) = 2^m \prod_{k=1}^{m} t_{kk}^{2m-2k+1}(dT).$$

Hence

$$(dX) = 2^{-m} \prod_{k=1}^{m} t_{kk}^{2n-2m}(dA)(E^H dE).$$

Moreover, we have

$$\prod_{k=1}^{m} t_{kk} = \det T = (\det T^H T)^{1/2} = (\det A)^{1/2}.$$

Therefore,

$$\begin{aligned} f(A) \int_{X^H X = A} (dX) &= 2^{-m} f(A) (\det A)^{n-m} \int_{\mathcal{CV}_{m,n}} (E^H dE) \\ &= \frac{\pi^{nm}}{\mathcal{C}\Gamma_m(n)} (\det A)^{n-m} f(A). \end{aligned} \quad (7.1)$$

The last equality follows from Proposition 2.7. □

The channel capacity is given by the following theorem. We shall assume that the realization of $\mathbf{H}$ is known to the receiver, or equivalently, the channel output consists of the pair $(y, \mathbf{H})$.

Theorem 7.1 *Consider the correlated Rayleigh channel, i.e.,* $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, I_{n_r} \otimes \Sigma_1)$*, with* $n_r \geq n_t$. *If the input power is constrained by* ρ*, then the capacity* C *is given by*

$$\frac{1}{\mathcal{C}\Gamma_{n_t}(n_r)(\det \Sigma_1)^{n_r}} \int_{W>0} \log\det\left[I_{n_t} + (\rho/n_t)W\right] (\det W)^{n_r - n_t} \operatorname{etr}\left(-\Sigma_1^{-1} W\right) (dW), \quad (7.2)$$

where $W = H^H H$.

Proof. From the capacity formula (6.25), the capacity C is given by

$$C = \int_H \log\det\left[I_{n_t} + (\rho/n_t)H^H H\right] f(H)(dH),$$

where

$$f(H) = \pi^{-n_r n_t}\det(\Sigma_1)^{-n_r}\operatorname{etr}\left(-\Sigma_1^{-1}H^H H\right).$$

Using Lemma 7.1, we can write C as

$$\begin{aligned} C &= \pi^{-n_r n_t}\det(\Sigma_1)^{-n_r}\int_H \log\det\left[I_{n_t} + (\rho/n_t)H^H H\right]\operatorname{etr}\left(-\Sigma_1^{-1}H^H H\right)(dH) \\ &= \pi^{-n_r n_t}\det(\Sigma_1)^{-n_r}\int_{W>0}\int_{H^H H = W}\log\det\left[I_{n_t} + (\rho/n_t)H^H H\right]\operatorname{etr}\left(-\Sigma_1^{-1}H^H H\right)(dH)(dW) \\ &= \frac{1}{\mathcal{C}\Gamma_{n_t}(n_r)(\det\Sigma_1)^{n_r}}\int_{W>0}\log\det\left[I_{n_t} + (\rho/n_t)W\right](\det W)^{n_r - n_t}\operatorname{etr}\left(-\Sigma_1^{-1}W\right)(dW). \end{aligned}$$

□

Theorem 7.1 can also be obtained by using Proposition 6.5 and the central Wishart density given in Proposition 4.2. Moreover, we assume that the Wishart matrix is nonsingular and its eigenvalues are positive. The knowledge of the distribution of the eigenvalues can be used to test hypotheses about their values. Specifically this can be used to test how many transmitted signals are actually received at the output.

Now using the eigenvalue density of a central Wishart matrix, the correlated Rayleigh channel capacity can be expressed as follows.

Theorem 7.2 *Consider the correlated Rayleigh channel, i.e.,* $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, I_{n_r}\otimes\Sigma_1)$*, with* $n_r \geq n_t$*. If the input power is constrained by* ρ*, then using the eigenvalue density of the central Wishart matrix* $\mathbf{W} = \mathbf{H}^H\mathbf{H}$ *we can write the capacity* C *as*

$$K\int_{\Lambda>0}\log\left\{\prod_{k=1}^{n_t}\left[1 + (\rho/n_t)\lambda_k\right]\right\}\prod_{k=1}^{n_t}\lambda_k^{n_r - n_t}\prod_{k<l}^{n_t}(\lambda_k - \lambda_l)^2{}_0F_0\left(-\Sigma_1^{-1}, \Lambda\right)\bigwedge_{k=1}^{n_t} d\lambda_k, \quad (7.3)$$

where $(\lambda_1, \ldots, \lambda_{n_t})$ *are the eigenvalues of* $\mathbf{W}$, $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_{n_t})$, *and*

$$K = \frac{\pi^{n_t(n_t-1)}(\det \Sigma_1)^{-n_r}}{\mathcal{C}\Gamma_{n_t}(n_t)\mathcal{C}\Gamma_{n_t}(n_r)}.$$

Proof. From Theorem 7.1, the capacity C is given by

$$\begin{aligned} C &= \mathcal{E}_{\mathbf{W}}\left\{\log\det\left(I_{n_t} + (\rho/n_t)\mathbf{W}\right)\right\} \\ &= \mathcal{E}_{\Lambda}\left\{\log\left(\prod_{k=1}^{n_t}[1 + (\rho/n_t)\lambda_k]\right)\right\}. \end{aligned} \tag{7.4}$$

The result follows by using the following eigenvalue density (see Propositions 4.4)

$$f(\lambda_1, \ldots, \lambda_{n_t}) = \frac{\pi^{n_t(n_t-1)}(\det \Sigma_1)^{-n_r}}{\mathcal{C}\Gamma_{n_t}(n_t)\mathcal{C}\Gamma_{n_t}(n_r)} \prod_{k=1}^{n_t} \lambda_k^{n_r-n_t} \prod_{k<l}^{n_t}(\lambda_k - \lambda_l)^2 {}_0F_0\left(-\Sigma_1^{-1}, \Lambda\right). \tag{7.5}$$

□

As mentioned in Equation (4.14), the joint eigenvalue density of a central Wishart matrix depends on the population covariance matrix Σ_1 only through its eigenvalues $\upsilon_1, \ldots, \upsilon_{n_t}$, i.e.,

$${}_0F_0\left(-\Sigma_1^{-1}, \Lambda\right) = {}_0F_0\left(-\Upsilon^{-1}, \Lambda\right),$$

where $\Upsilon = \mathrm{diag}(\upsilon_1, \ldots, \upsilon_{n_t})$. Let $\Upsilon^{-1} = \mathrm{diag}(a_1, \ldots, a_{n_t})$. Then ${}_0F_0\left(-\Upsilon^{-1}, \Lambda\right)$ can be written as [29]

$${}_0F_0\left(-\Upsilon^{-1}, \Lambda\right) = \frac{\mathcal{C}\Gamma_{n_t}(n_t)\det\left[\left(\exp\left(-a_i\lambda_j\right)\right)\right]}{\pi^{n_t(n_t-1)/2}\prod_{k<l}^{n_t}(\lambda_k - \lambda_l)\prod_{k<l}^{n_t}(a_l - a_k)}. \tag{7.6}$$

Theorem 7.3 *Consider the correlated Rayleigh channel, i.e.,* $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, I_{n_r} \otimes \Sigma_1)$, *with* $n_r \geq n_t$. *If the input power is constrained by* ρ, *then using the eigenvalue density we can write the capacity* C *as*

$$C = n_t \mathcal{E}_{\lambda_1}\left[\log(1 + (\rho/n_t)\lambda_1)\right]. \tag{7.7}$$

The density $f(\lambda_1)$ *given by*

$$f(\lambda_1) = \frac{\pi^{n_t(n_t-1)/2}\prod_{k=1}^{n_t} a_k^{n_r}}{n_t! C\Gamma_{n_t}(n_r)\prod_{k<l}^{n_t}(a_l - a_k)} \int \left\{ \widetilde{\sum}_i (-1)^{\mathrm{per}(i_1,\ldots,i_{n_t})} \exp\left(\sum_{j=1}^{n_t} -a_{i_j}\lambda_j\right)\right\} \cdot \left\{ \widetilde{\sum}_k (-1)^{\mathrm{per}(k_1,\ldots,k_{n_t})} \prod_{l=1}^{n_t} \lambda_l^{n_r-n_t+k_l} \right\} \bigwedge_{k=2}^{n_t} d\lambda_k, \tag{7.8}$$

where $\widetilde{\sum}_i$ *denotes summation over all permutations* $(i_1,\ldots,i_{n_t})$ *of* $(1,\ldots,n_t)$, $\widetilde{\sum}_k$ *denotes summation over all permutations* $(k_1,\ldots,k_{n_t})$ *of* $(0,\ldots,n_t-1)$ *and* $\mathrm{per}(k_1,\ldots,k_{n_t})$ *is* 0 *or* 1 *depending on the permutation being even or odd. Similarly for* $\mathrm{per}(i_1,\ldots,i_{n_t})$.

Proof. From Equation (7.4), C can be written as

$$\begin{aligned} C &= \sum_{k=1}^{n_t} \mathcal{E}_{\lambda_k}\left[\log(1 + (\rho/n_t)\lambda_k)\right] \\ &= n_t \mathcal{E}_{\lambda_1}\left[\log(1 + (\rho/n_t)\lambda_1)\right] \end{aligned} \tag{7.9}$$

where the expectation is with respect to λ_1. By substituting (7.6) in (7.5) and integrating with respect to $\lambda_2,\ldots,\lambda_{n_t}$, i.e.,

$$f(\lambda_1) = \frac{\pi^{n_t(n_t-1)/2}\prod_{k=1}^{n_t} a_k^{n_r}}{C\Gamma_{n_t}(n_r)\prod_{k<l}^{n_t}(a_l - a_k)} \int \det\left[(\exp(-a_i\lambda_j))\right] \prod_{k<l}^{n_t}(\lambda_k - \lambda_l) \prod_{k=1}^{n_t} \lambda_k^{n_r-n_t} \bigwedge_{k=2}^{n_t} d\lambda_k. \tag{7.10}$$

As in [33], the integrand in Equation (7.10) can be written as

$$\begin{aligned} &\det\left[(\exp(-a_i\lambda_j))\right] \prod_{k<l}^{n_t}(\lambda_k - \lambda_l) \prod_{k=1}^{n_t} \lambda_k^{n_r-n_t} \\ &= \frac{1}{n_t!} \det \begin{bmatrix} e^{-a_1\lambda_1} & \cdots & e^{-a_1\lambda_{n_t}} \\ e^{-a_2\lambda_1} & \cdots & e^{-a_2\lambda_{n_t}} \\ \vdots & \vdots & \vdots \\ e^{-a_{n_t}\lambda_1} & \cdots & e^{-a_{n_t}\lambda_{n_t}} \end{bmatrix} \det \begin{bmatrix} 1 & \cdots & 1 \\ \lambda_1 & \cdots & \lambda_{n_t} \\ \vdots & \vdots & \vdots \\ \lambda_1^{n_t-1} & \cdots & \lambda_{n_t}^{n_t-1} \end{bmatrix} \prod_{k=1}^{n_t} \lambda_k^{n_r-n_t} \\ &= \frac{1}{n_t!} \det \begin{bmatrix} e^{-a_1\lambda_1} & \cdots & e^{-a_1\lambda_{n_t}} \\ \vdots & \vdots & \vdots \\ e^{-a_{n_t}\lambda_1} & \cdots & e^{-a_{n_t}\lambda_{n_t}} \end{bmatrix} \det \begin{bmatrix} \lambda_1^{n_r-n_t} & \cdots & \lambda_{n_t}^{n_r-n_t} \\ \vdots & \vdots & \vdots \\ \lambda_1^{n_r-1} & \cdots & \lambda_{n_t}^{n_r-1} \end{bmatrix} \end{aligned}$$

$$= \frac{1}{n_t!}\left\{\widetilde{\sum_i}(-1)^{\text{per}(i_1,\ldots,i_{n_t})}\exp\left(\sum_{j=1}^{n_t}-a_{i_j}\lambda_j\right)\right\}\left\{\widetilde{\sum_k}(-1)^{\text{per}(k_1,\ldots,k_{n_t})}\prod_{l=1}^{n_t}\lambda_l^{n_r-n_t+k_l}\right\}$$

The result follows. □

7.1.1 Correlated Rayleigh $n_r \times 2$ channel matrix

In this section, a numerical evaluation of correlated Rayleigh $n_r \times 2$ channel matrix is given. Thus, we assume that we have a two-input ($n_t = 2$), n_r-output communication system operating over a correlated Rayleigh fading environment (typical mobile wireless environment). As mentioned before, the joint eigenvalue density of a central Wishart matrix depends on the population covariance matrix Σ_1 only through its eigenvalues $\upsilon_1, \ldots, \upsilon_{n_t}$, i.e.,

$$_0F_0\left(-\Sigma_1^{-1}, \Lambda\right) = {_0F_0}\left(-\Upsilon^{-1}, \Lambda\right),$$

where $\Upsilon = \text{diag}(\upsilon_1, \ldots, \upsilon_{n_t})$.

Let $n_t = 2$ and $\Upsilon^{-1} = \text{diag}(a_1, a_2)$. Then we have [29]

$$_0F_0(-\Upsilon^{-1}, \Lambda) = \frac{1}{(a_2 - a_1)(\lambda_1 - \lambda_2)}\left[\exp\left\{-(a_1\lambda_1 + a_2\lambda_2)\right\} - \exp\left\{-(a_1\lambda_2 + a_2\lambda_1)\right\}\right]. \tag{7.11}$$

The following theorem gives the correlated Rayleigh channel capacity for a $n_r \times 2$ matrix.

Theorem 7.4 *Consider the two-input correlated Rayleigh channel, i.e.,* $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, I_{n_r} \otimes \Sigma_1)$, *with* $n_r \geq 2$. *If the input power is constrained by* ρ, *then the capacity* C *is given by*

$$\begin{aligned} C &= \frac{a_1^{n_r} a_2}{(a_2 - a_1)\Gamma(n_r)} \int_0^\infty \log[1 + (\rho/2)\lambda_1]\lambda_1^{n_r-1} e^{-a_1\lambda_1} d\lambda_1 \\ &\quad - \frac{a_1 a_2^{n_r}}{(a_2 - a_1)\Gamma(n_r)} \int_0^\infty \log[1 + (\rho/2)\lambda_1]\lambda_1^{n_r-1} e^{-a_2\lambda_1} d\lambda_1 \end{aligned}$$

$$- \frac{a_1^{n_r}}{(a_2 - a_1)\Gamma(n_r - 1)} \int_0^\infty \log[1 + (\rho/2)\lambda_1]\lambda_1^{n_r-2} e^{-a_1\lambda_1} d\lambda_1$$
$$+ \frac{a_2^{n_r}}{(a_2 - a_1)\Gamma(n_r - 1)} \int_0^\infty \log[1 + (\rho/2)\lambda_1]\lambda_1^{n_r-2} e^{-a_2\lambda_1} d\lambda_1, \qquad (7.12)$$

where λ_1 is an eigenvalue of $W = H^H H$ and (a_1, a_2) are eigenvalues of Σ_1^{-1}.

Proof. Using Equation (7.11), the eigenvalue density of W is given by

$$f(\lambda_1, \lambda_2) = \frac{(a_1 a_2)^{n_r}(\lambda_1\lambda_2)^{n_r-2}(\lambda_1 - \lambda_2)}{2(a_2 - a_1)\Gamma(n_r)\Gamma(n_r - 1)} \left[e^{-a_1\lambda_1 - a_2\lambda_2} - e^{-a_1\lambda_2 - a_2\lambda_1}\right]. \qquad (7.13)$$

Now, integrating with respect to λ_2 and noting that

$$\int_0^\infty x^{a-1} e^{-x/b} dx = \Gamma(a) b^a,$$

we obtain the density of $f(\lambda_1)$. Thus we have

$$f(\lambda_1) = \frac{1}{2(a_2 - a_1)} \left\{ \frac{a_1^{n_r} a_2 \lambda_1^{n_r-1} e^{-a_1\lambda_1}}{\Gamma(n_r)} - \frac{a_1 a_2^{n_r} \lambda_1^{n_r-1} e^{-a_2\lambda_1}}{\Gamma(n_r)} - \frac{a_1^{n_r}\lambda_1^{n_r-2} e^{-a_1\lambda_1}}{\Gamma(n_r - 1)} + \frac{a_2^{n_r}\lambda_1^{n_r-2} e^{-a_2\lambda_1}}{\Gamma(n_r - 1)} \right\}. \qquad (7.14)$$

It is easy to see that $\int_0^\infty f(\lambda_1) d\lambda_1 = 1$. Finally, evaluating Equation (7.7) with $f(\lambda_1)$ gives Equation (7.12). □

Tables 7.1 and 7.2 show the capacity in nats[1] for a $n_r \times 2$ correlated Rayleigh fading channel matrix with correlation coefficients 0.2 and 0.9, respectively. Note that each column represents different levels of input power or signal to noise ratio (SNR) in dB. Figures 7.1 and 7.3 show the capacity in nats vs n_r for the correlation coefficients 0.2 and 0.9, respectively. Figure 7.4 shows the capacity vs the correlation coefficient and Figure 7.2 shows the capacity vs SNR. From these tables and figures we note the following: (*i*) the capacity is decreasing with increasing channel correlation, (*ii*) the

[1]In Equation (7.12), 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 ($e = 2.718\ldots$).

capacity is increasing with increasing n_r and SNR.

	ρ in dB							
n_r	0 dB	5 dB	10 dB	15 dB	20 dB	25 dB	30 dB	35 dB
2	1.1613	2.2752	3.8147	5.6745	7.7246	9.8622	12.0327	14.2143
4	1.9691	3.5653	5.5446	7.7229	9.9827	12.2713	14.5694	16.8706
6	2.5660	4.3816	6.4904	8.7254	11.0058	13.3013	15.6017	17.9035
8	3.0336	4.9685	7.1343	9.3909	11.6786	13.9764	16.2775	18.5796
10	3.4160	5.4244	7.6217	9.8894	12.1808	14.4798	16.7813	19.0835
12	3.7387	5.7965	8.0136	10.2882	12.5818	14.8815	17.1832	19.4855
14	4.0175	6.1107	8.3413	10.6205	12.9156	15.2158	17.5177	19.8200
16	4.2626	6.3824	8.6229	10.9054	13.2016	15.5021	17.8041	20.1064
18	4.4813	6.6218	8.8697	11.1547	13.4517	15.7525	18.0545	20.3569
20	4.6787	6.8356	9.0895	11.3764	13.6740	15.9750	18.2770	20.5795

Table 7.1: The capacity in nats for a two-input, n_r-output communication system operating over a correlated Rayleigh fading channel, where ρ is signal to noise ratio in dB and the correlation coefficient equal to 0.2.

Note that the covariance matrix is $\Sigma_1 = \begin{bmatrix} 1 & 0.2 \\ 0.2 & 1 \end{bmatrix}$, its eigenvalues are $\Upsilon = \text{diag}(1.2, 0.8)$ and $a_1 = 1/1.2, a_2 = 1/0.8$. Note that the diagonal element of Σ gives the correlation between the channel coefficient from different transmitter antennas to a single receiver antenna. This diagonal element is called a correlation coefficient.

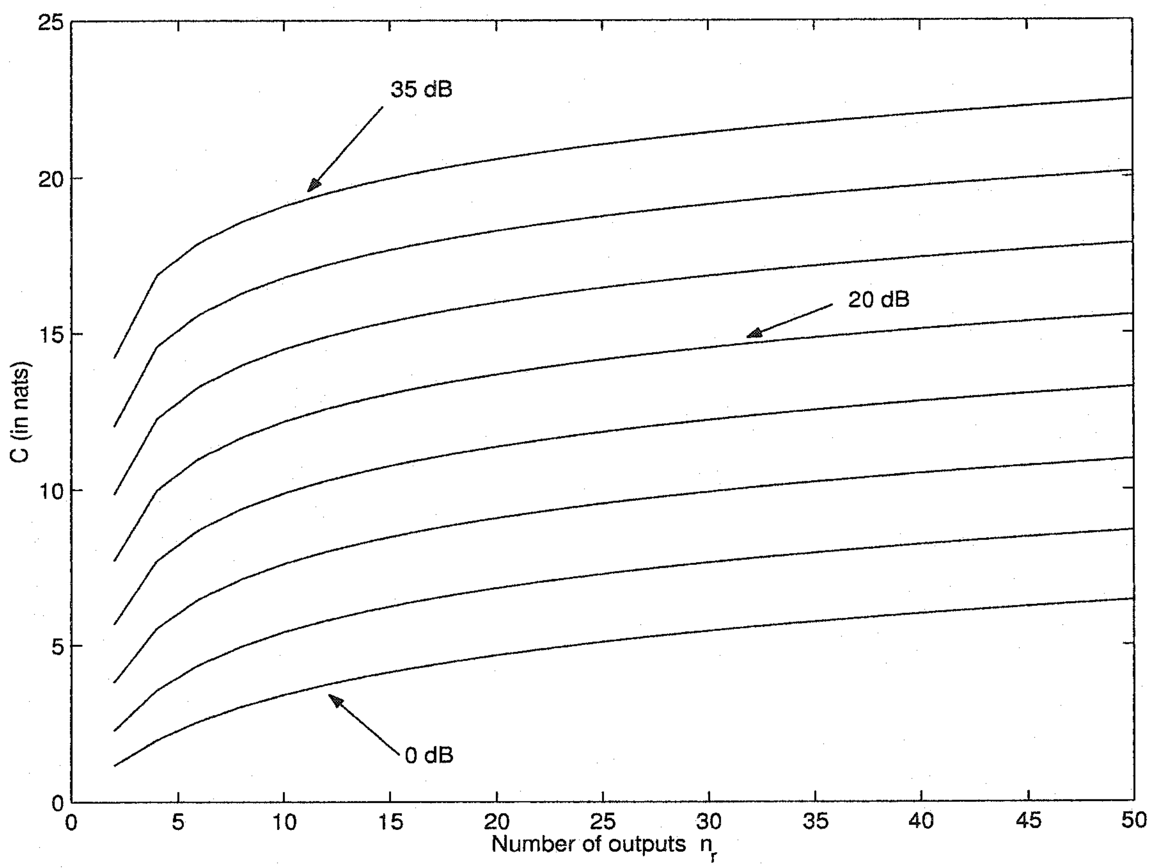

Figure 7.1: Capacity vs number of outputs for SNR= $0, 5, 10, 15, 20, 25, 30, 35$ dB. Note that $\mathbf{H}$ is a $n_r \times 2$ correlated Rayleigh fading channel matrix with correlation coefficient equal to 0.2.

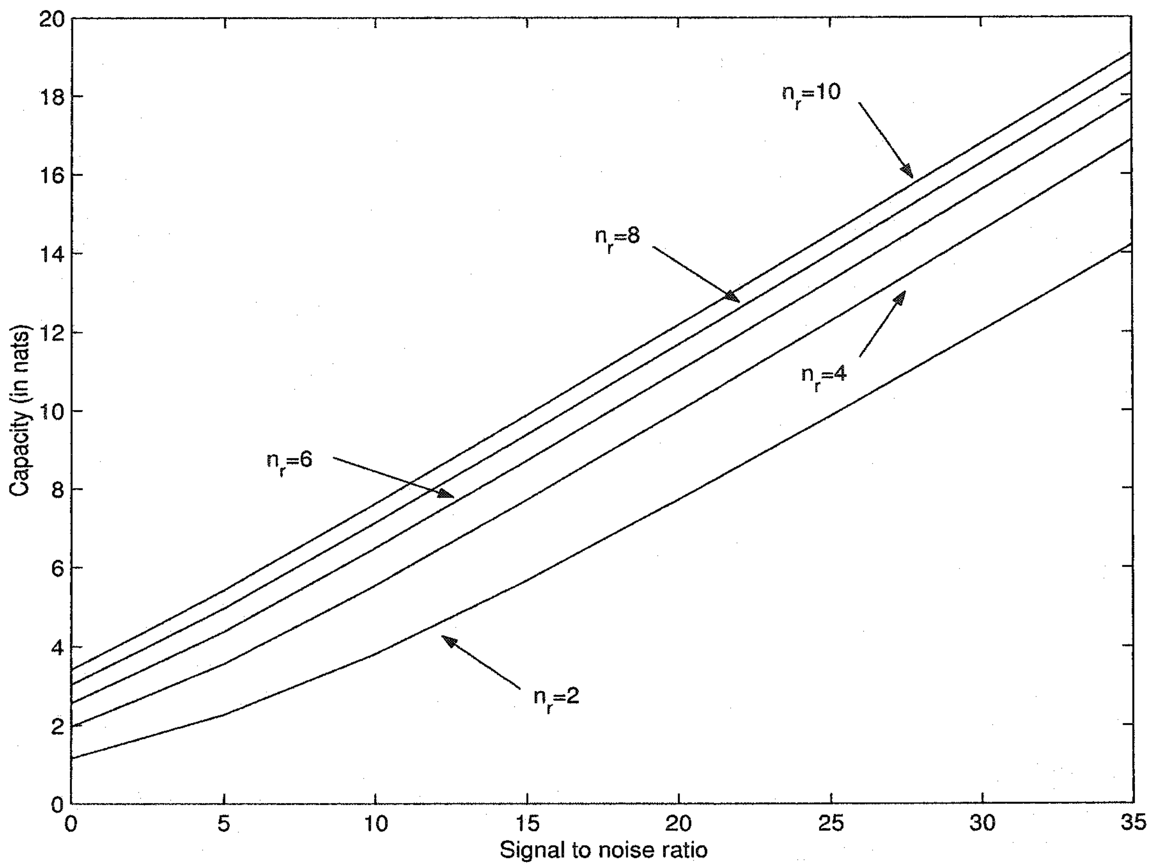

Figure 7.2: Capacity vs SNR for correlation coefficient 0.2 and $n_t = 2$ and $n_r = 2, 4, 6, 8, 10$, i.e., $\mathbf{H}$ is a $n_r \times 2$ correlated Rayleigh fading channel matrix.

	ρ in dB							
n_r	0 dB	5 dB	10 dB	15 dB	20 dB	25 dB	30 dB	35 dB
2	1.0326	1.9252	3.1157	4.5641	6.2023	7.9419	9.7221	11.5165
4	1.6408	2.8426	4.4118	6.3154	8.4439	10.6803	12.9577	15.2490
6	2.0685	3.4398	5.1852	7.2250	9.4266	11.6948	13.9863	16.2855
8	2.4033	3.8917	5.7454	7.8540	10.0862	12.3653	14.6604	16.9606
10	2.6804	4.2568	6.1838	8.3330	10.5817	12.8666	15.1635	17.4643
12	2.9179	4.5639	6.5437	8.7196	10.9786	13.2669	15.5650	17.8661
14	3.1265	4.8293	6.8489	9.0437	11.3096	13.6003	15.8992	18.2005
16	3.3129	5.0631	7.1139	9.3226	11.5936	13.8860	16.1853	18.4869
18	3.4817	5.2722	7.3479	9.5674	11.8422	14.1359	16.4357	18.7373
20	3.6361	5.4612	7.5574	9.7855	12.0634	14.3580	16.6581	18.9599

Table 7.2: The capacity in nats for a two-input, n_r-output communication system operating over a correlated Rayleigh fading channel, where ρ is signal to noise ratio in dB and the correlation coefficient equal to 0.9.

Note that the covariance matrix is $\Sigma_1 = \begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}$, its eigenvalues are $\Upsilon = \operatorname{diag}(1.9, 0.1)$ and $a_1 = 1/1.9, a_2 = 1/0.1$.

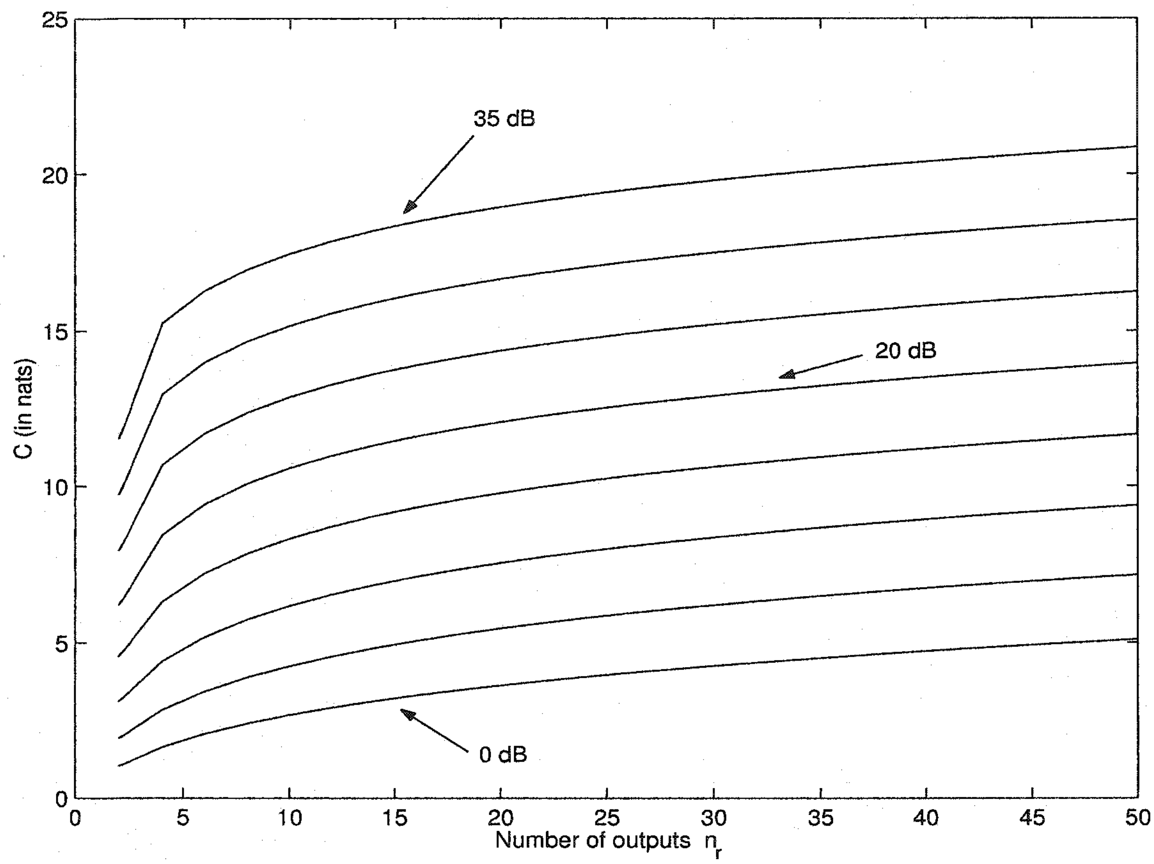

Figure 7.3: Capacity vs number of outputs for SNR= $0, 5, 10, 15, 20, 25, 30, 35$ dB. Note that $\mathbf{H}$ is a $n_r \times 2$ correlated Rayleigh fading channel matrix with the correlation coefficient is equal to 0.9.

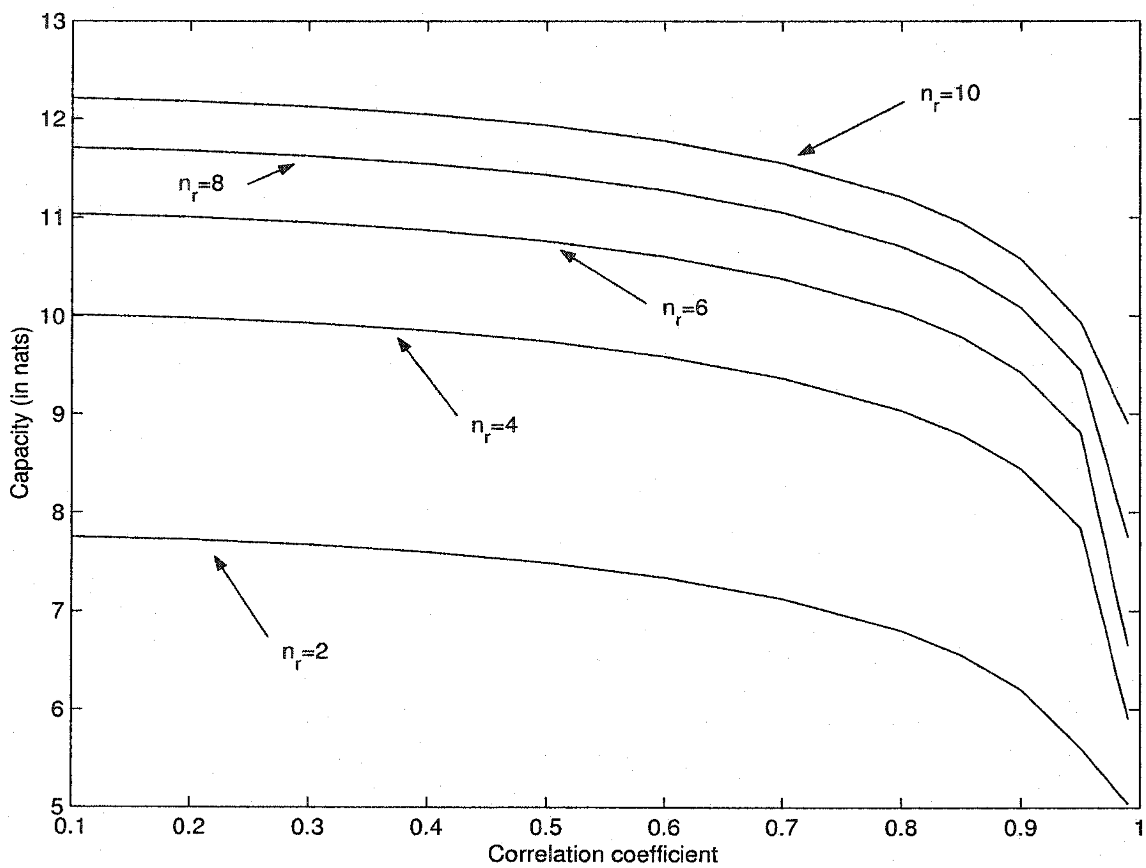

Figure 7.4: Capacity vs correlation coefficient for SNR=20dB and $n_t = 2$ and $n_r = 2, 4, 6, 8, 10$, i.e., $\mathbf{H}$ is a $n_r \times 2$ correlated Rayleigh fading channel matrix.

7.2 Uncorrelated Rayleigh channel

It should be noted that the capacity evaluation for an uncorrelated Rayleigh channel does not require the computation of the hypergeometric function or zonal polynomial and this was investigated in [45]. Here we assume that $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, I_{n_r} \otimes (\sigma^2 I_{n_t}))$. In other words, the elements of $\mathbf{H}$ are iid (independent and identically distributed) and $h_{ij} \sim \mathcal{CN}(0, \sigma^2)$, i.e, $\mathrm{Re}(h_{ij}), \mathrm{Im}(h_{ij}) \sim \mathcal{N}(0, \sigma^2/2)$. In this case, the magnitude or absolute value $|h_{ij}|$ is distributed as a Rayleigh distribution and $\arg h_{ij}$ is distributed as uniform $[0, 2\pi]$ (see Equations (1.6) and (1.7)). This choice models a *Rayleigh fading environment* with enough separation within the output and the input of the communication system such that the fades for each input-output pair are independent. We shall assume that the realization of $\mathbf{H}$ is known to the receiver, or equivalently, the channel output consists of the pair $(y, \mathbf{H})$. The following theorem gives the capacity formula for this scenario.

Theorem 7.5 *Consider the uncorrelated Rayleigh channel, i.e.,* $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, I_{n_r} \otimes \sigma^2 I_{n_t})$*, with* $n_r \geq n_t$*. If the input power is constrained by* ρ*, then the capacity* C *is given by*

$$\frac{(\sigma^2)^{-n_r n_t}}{C\Gamma_{n_t}(n_r)} \int_{W>0} \log\det\left[I_{n_t} + (\rho/n_t)W\right] (\det W)^{n_r - n_t} \operatorname{etr}\left(-\frac{1}{\sigma^2}W\right)(dW) \qquad (7.15)$$

where $W = H^H H$.

Proof. The proof is similar to that of Theorem 7.1. □

The following theorem corresponds to Theorem 7.2.

Theorem 7.6 *Consider the uncorrelated Rayleigh channel, i.e.,* $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, I_{n_r} \otimes \sigma^2 I_{n_t})$*, with* $n_r \geq n_t$*. If the input power is constrained by* ρ*, then using the eigenvalue density of the central Wishart matrix* $\mathbf{W} = \mathbf{H}^H\mathbf{H}$*, we can write the capacity* C *as*

$$K \int_{\Lambda>0} \log\left\{\prod_{k=1}^{n_t}\left[1 + (\rho/n_t)\lambda_k\right]\right\} \prod_{k=1}^{n_t} \lambda_k^{n_r - n_t} \prod_{k<l}^{n_t} (\lambda_k - \lambda_l)^2 \exp\left(-\frac{1}{\sigma^2}\sum_{k=1}^{n_t}\lambda_k\right) \bigwedge_{k=1}^{n_t} d\lambda_k, \qquad (7.16)$$

where $(\lambda_1, \ldots, \lambda_{n_t})$ *are the eigenvalues of* $\mathbf{W}$, $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_{n_t})$, *and*

$$K = \frac{\pi^{n_t(n_t-1)}(\sigma^2)^{-n_r n_t}}{\mathcal{C}\Gamma_{n_t}(n_t)\mathcal{C}\Gamma_{n_t}(n_r)}.$$

Proof. The proof is similar to that of Theorem 7.2. □

The following theorem corresponds to Theorem 7.3.

Theorem 7.7 *Consider the uncorrelated Rayleigh channel, i.e.,* $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, I_{n_r} \otimes \sigma^2 I_{n_t})$, *with* $n_r \geq n_t$. *If the input power is constrained by* ρ, *then we can write the capacity* C *as*

$$C = n_t \mathcal{E}_{\lambda_1}\left[\log(1 + (\rho/n_t)\lambda_1)\right]. \tag{7.17}$$

The density $f(\lambda_1)$ *is given by*

$$f(\lambda_1) = \frac{1}{n_t}\sum_{k=1}^{n_t}\left[\varphi_k(\lambda_1)\right]^2 \tag{7.18}$$

and φ_k *form the orthonormal set which can be obtained by applying the Gram–Schmidt procedure to the sequence of functions*

$$\lambda^{(n_r-n_t)/2}e^{-\lambda/(2\sigma^2)}, \lambda^{(n_r-n_t)/2+1}e^{-\lambda/(2\sigma^2)}, \lambda^{(n_r-n_t)/2+2}e^{-\lambda/(2\sigma^2)}, \ldots, \lambda^{(n_r+n_t)/2-1}e^{-\lambda/(2\sigma^2)}.$$

Proof. The first part of the proof is the same as in Theorem 7.3 and the second part, Equation (7.18), is given below. See [49] for similar work. The joint eigenvalue density can be written as

$$\begin{aligned} f(\Lambda) &= \frac{K}{n_t!}\left(\det\begin{bmatrix} 1 & \cdots & 1 \\ \lambda_1 & \cdots & \lambda_{n_t} \\ \vdots & & \vdots \\ \lambda_1^{n_t-1} & \cdots & \lambda_{n_t}^{n_t-1}\end{bmatrix}\right)^2 \left[\prod_{k=1}^{n_t}\lambda_k^{n_r-n_t}\exp\left(-\frac{1}{\sigma^2}\lambda_k\right)\right] \\ &= \frac{K}{n_t!}\left(\det\begin{bmatrix} \lambda_1^{(n_r-n_t)/2}e^{-\frac{\lambda_1}{2\sigma^2}} & \cdots & \lambda_{n_t}^{(n_r-n_t)/2}e^{-\frac{\lambda_{n_t}}{2\sigma^2}} \\ \vdots & & \vdots \\ \lambda_1^{(n_r+n_t)/2-1}e^{-\frac{\lambda_1}{2\sigma^2}} & \cdots & \lambda_{n_t}^{(n_r+n_t)/2-1}e^{-\frac{\lambda_{n_t}}{2\sigma^2}}\end{bmatrix}\right)^2 \end{aligned}$$

$$= \frac{K_1}{n_t!}\left(\det\begin{bmatrix}\varphi_1(\lambda_1) & \ldots & \varphi_1(\lambda_{n_t})\\ \vdots & & \vdots \\ \varphi_{n_t}(\lambda_1) & \ldots & \varphi_{n_t}(\lambda_{n_t})\end{bmatrix}\right)^2, \tag{7.19}$$

where φ_k is defined in Theorem 7.7 and satisfies

$$\int \varphi_k(\lambda)\varphi_l(\lambda)d\lambda = \delta_{kl}.$$

The determinant squared in Equation (7.19) can be expanded as

$$\widetilde{\sum}_{r,s}(-1)^{\mathrm{per}(r_1,\ldots,r_{n_t})}(-1)^{\mathrm{per}(s_1,\ldots,s_{n_t})}\prod_k \varphi_{r_k}(\lambda_k)\varphi_{s_k}(\lambda_k)$$

where $\widetilde{\sum}_{r,s}$ denotes summation over all permutations $(r_1,\ldots,r_{n_t})$ and $(s_1,\ldots,s_{n_t})$ of $(1,\ldots,n_t)$ and $\mathrm{per}(r_1,\ldots,r_{n_t})$ is 0 or 1 depending on the permutation being even or odd. Similarly for $\mathrm{per}(s_1,\ldots,s_{n_t})$. Hence, $f(\lambda_1)$ can be obtained by integrating (7.19) with respect to $\lambda_2,\ldots,\lambda_{n_t}$, i.e.,

$$\begin{aligned}
f(\lambda_1) &= \int\cdots\int f(\Lambda)\bigwedge_{k=2}^{n_t} d\lambda_k\\
&= \frac{K_1}{n_t!}\widetilde{\sum}_{r,s}(-1)^{\mathrm{per}(r_1,\ldots,r_{n_t})}(-1)^{\mathrm{per}(s_1,\ldots,s_{n_t})}\int\cdots\int\prod_k \varphi_{r_k}(\lambda_k)\varphi_{s_k}(\lambda_k)\bigwedge_{k=2}^{n_t} d\lambda_k\\
&= \frac{K_1}{n_t!}\widetilde{\sum}_{r,s}(-1)^{\mathrm{per}(r_1,\ldots,r_{n_t})}(-1)^{\mathrm{per}(s_1,\ldots,s_{n_t})}\varphi_{r_1}(\lambda_1)\varphi_{s_1}(\lambda_1)\prod_{k\geq 2}^{n_t}\delta_{r_k s_k}\\
&= \frac{K_1(n_t-1)!}{n_t!}\sum_{k=1}^{n_t}[\varphi_k(\lambda_1)]^2\\
&= \frac{1}{n_t}\sum_{k=1}^{n_t}[\varphi_k(\lambda_1)]^2.
\end{aligned} \tag{7.20}$$

Since $\int f(\lambda_1)d\lambda_1 = 1$, therefore, $K_1 = 1$. The proof is complete. □

Remark 7.1 *The evaluation of* (7.18) *for* $\sigma^2 = 1$ *is given in [5], [33] and [45], where* *the following formula*

$$f(\lambda_1) = \frac{1}{n_t}\sum_{l=1}^{n_t}[\varphi_l(\lambda_1)]^2 \tag{7.21}$$

is obtained with

$$\begin{aligned}\varphi_{k+1}(\lambda) &= \lambda^{(n_r-n_t)/2}e^{-\lambda/2}\left[\frac{k!}{(k+n_r-n_t)!}\right]^{1/2}\frac{1}{k!}e^{\lambda}\lambda^{n_t-n_r}\frac{d^k}{d\lambda^k}(e^{-\lambda}\lambda^{n_r-n_t+k})\\ &= \lambda^{(n_r-n_t)/2}e^{-\lambda/2}\left[\frac{k!}{(k+n_r-n_t)!}\right]^{1/2}L_k^{n_r-n_t}(\lambda), \quad k=0,\ldots,n_t-1,\end{aligned}$$

and $L_k^{n_r-n_t}(\lambda)$ *is the generalized Laguerre polynomial of order* k.

For smaller n_t, the integration (7.18) can be done without going through the Gram–Schmidt orthogonalization. This is done in the next section.

7.2.1 Uncorrelated Rayleigh $n_r \times 2$ channel matrix

In this section, the numerical evaluation of an uncorrelated Rayleigh $n_r \times 2$ channel matrix is given. In other words, we assumed we have a two-input ($n_t = 2$), n_r-output communication system operating over an uncorrelated Rayleigh fading environment, which is a typical fixed wireless environment. The following theorem gives the capacity expression.

Theorem 7.8 *Consider the two-input uncorrelated Rayleigh channel, i.e.,* $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, I_{n_r}\otimes \sigma^2 I_2)$*, with* $n_r \geq 2$*. If the input power is constrained by* ρ*, then the capacity* C *is given by*

$$\begin{aligned} C &= \frac{(\sigma^2)^{-n_r-1}}{\Gamma(n_r)}\int_0^\infty \log[1+(\rho/2)\lambda_1]\lambda_1^{n_r}e^{-\lambda_1/\sigma^2}d\lambda_1\\ &\quad -\frac{2(\sigma^2)^{-n_r}}{\Gamma(n_r-1)}\int_0^\infty \log[1+(\rho/2)\lambda_1]\lambda_1^{n_r-1}e^{-\lambda_1/\sigma^2}d\lambda_1\\ &\quad +\frac{(\sigma^2)^{-n_r+1}\Gamma(n_r+1)}{\Gamma(n_r)\Gamma(n_r-1)}\int_0^\infty \log[1+(\rho/2)\lambda_1]\lambda_1^{n_r-2}e^{-\lambda_1/\sigma^2}d\lambda_1 \end{aligned} \tag{7.22}$$

where λ_1 *is an eigenvalue of* $W = H^H H$.

Proof. The eigenvalue density of W is given as

$$f(\lambda_1, \lambda_2) = \frac{(\sigma^2)^{-2n_r}(\lambda_1\lambda_2)^{n_r-2}(\lambda_1-\lambda_2)^2}{2\Gamma(n_r)\Gamma(n_r-1)} e^{-(\lambda_1+\lambda_2)/\sigma^2}. \tag{7.23}$$

Integrating with respect to λ_2 and noting that $\int_0^\infty x^{a-1}e^{-x/b}dx = \Gamma(a)b^a$, we obtain the density of λ_1, i.e.,

$$f(\lambda_1) = \frac{(\sigma^2)^{-n_r-1}}{2\Gamma(n_r)}\lambda_1^{n_r}e^{-\lambda_1/\sigma^2} - \frac{(\sigma^2)^{-n_r}}{\Gamma(n_r-1)}\lambda_1^{n_r-1}e^{-\lambda_1/\sigma^2} + \frac{(\sigma^2)^{-n_r+1}\Gamma(n_r+1)}{2\Gamma(n_r)\Gamma(n_r-1)}\lambda_1^{n_r-2}e^{-\lambda_1/\sigma^2} \tag{7.24}$$

It is easy to see that $\int_0^\infty f(\lambda_1)d\lambda_1 = 1$. Finally, evaluating Equation (7.17) with $f(\lambda_1)$ gives Equation (7.22). □

Tables 7.3 shows the capacity in nats for a $n_r \times 2$ uncorrelated Rayleigh fading channel matrix with different levels of input power. Figure 7.5 shows the capacity in nats vs n_r for different signal to noise ratios. It is clearly seen from the table and figure that the capacity is increasing with increasing n_r and SNR.

	ρ in dB							
n_r	0 dB	5 dB	10 dB	15 dB	20 dB	25 dB	30 dB	35 dB
2	1.1671	2.2890	3.8382	5.7066	7.7633	9.9062	12.0815	14.2676
4	1.9831	3.5910	5.5788	7.7614	10.0227	12.3119	14.6102	16.9114
6	2.5857	4.4125	6.5274	8.7649	11.0462	13.3420	15.6425	17.9444
8	3.0573	5.0020	7.1725	9.4308	11.7191	14.0172	16.3183	18.6204
10	3.4425	5.4595	7.6605	9.9296	12.2214	14.5206	16.8221	19.1244
12	3.7672	5.8326	8.0528	10.3285	12.6225	14.9223	17.2240	19.5263
14	4.0475	6.1475	8.3808	10.6609	12.9563	15.2566	17.5585	19.8608
16	4.2939	6.4197	8.6626	10.9458	13.2423	15.5429	17.8449	20.1473
18	4.5136	6.6595	8.9096	11.1952	13.4924	15.7933	18.0953	20.3977
20	4.7117	6.8736	9.1294	11.4169	13.7147	16.0158	18.3179	20.6203

Table 7.3: The capacity in nats for a two-input, n_r-output communication system operating over an uncorrelated Rayleigh fading channel, where ρ is signal to noise ratio in dB.

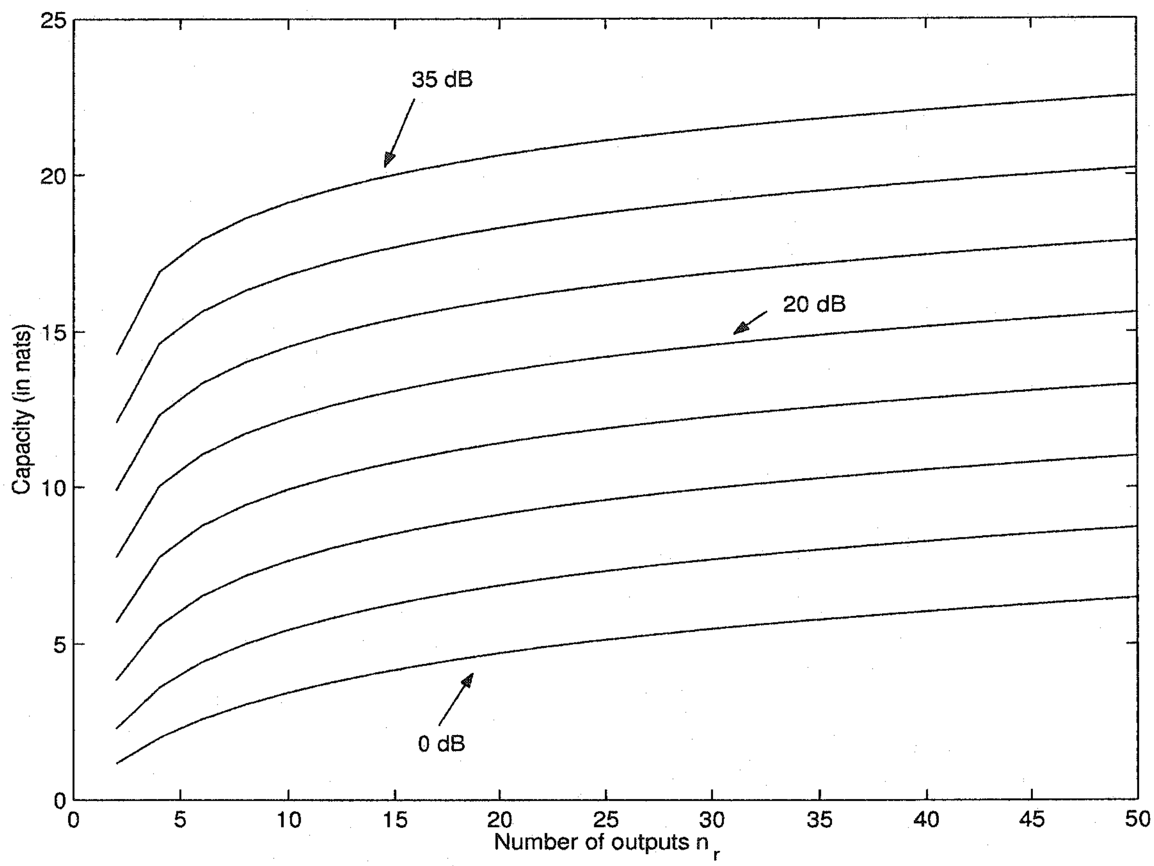

Figure 7.5: Capacity vs number of outputs for SNR=0, 5, 10, 15, 20, 25, 30, 35 dB. Note that $\mathbf{H}$ is a $n_r \times 2$ uncorrelated Rayleigh fading channel matrix.

7.3 Rician channel

In this section, we evaluate the channel capacity for a Rician channel. In a Rician channel, the distribution of an $n_r \times n_t$ channel matrix $\mathbf{H}$ is given by $\mathbf{H} \sim \mathcal{CN}(M, I_{n_r} \otimes \Sigma_1)$, with $n_r \geq n_t$. The channel capacity is given by the following theorem.

Theorem 7.9 *Consider a Rician channel, i.e.,* $\mathbf{H} \sim \mathcal{CN}(M, I_{n_r} \otimes \Sigma_1)$*, with* $n_r \geq n_t$*. If the input power is constrained by* ρ*, then using the complex noncentral Wishart density we can write the capacity as*

$$K_1 \int_{W>0} \log\det\left[I_{n_t} + (\rho/n_t)W\right] (\det W)^{n_r - n_t} \operatorname{etr}\left(-\Sigma_1^{-1}W\right) {}_0F_1\left(n_r; \Omega\Sigma_1^{-1}W\right)(dW), \tag{7.25}$$

where $W = H^H H$*,* $\Omega = \Sigma_1^{-1} M^H M$ *and*

$$K_1 = \frac{\operatorname{etr}(-\Omega)}{\mathcal{C}\Gamma_{n_t}(n_r)(\det \Sigma_1)^{n_r}}.$$

Proof. The proof is similar to that of Theorem 7.1. □

The following theorem expresses the Rician channel capacity in terms of the eigenvalue density of a complex noncentral Wishart matrix.

Theorem 7.10 *Consider the Rician channel, i.e.,* $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, I_{n_r} \otimes \Sigma_1)$*, with* $n_r \geq n_t$*. If the input power is constrained by* ρ*, then using the eigenvalue distribution of the Wishart matrix* $\mathbf{W} = \mathbf{H}^H\mathbf{H}$ *we can write the capacity as*

$$\int_{\Lambda>0} \log\left\{\prod_{k=1}^{n_t}\left[1 + (\rho/n_t)\lambda_k\right]\right\} f(\Lambda)d\Lambda, \tag{7.26}$$

where $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_{n_t})$ *and*

$$f(\Lambda) = \frac{\pi^{n_t(n_t-1)}(\det\Sigma_1)^{-n_r}}{\mathcal{C}\Gamma_{n_t}(n_t)\mathcal{C}\Gamma_{n_t}(n_r)} \operatorname{etr}(-\Omega) \prod_{k=1}^{n_t} \lambda_k^{n_r - n_t} \prod_{k<l}^{n_t} (\lambda_k - \lambda_l)^2$$
$$\sum_{k,l=0}^{\infty} \sum_{\kappa,\tau;\phi\in\kappa.\tau} \frac{C_\phi^{\kappa,\tau}(-\Sigma_1^{-1}, \Omega\Sigma_1^{-1}) C_\phi^{\kappa,\tau}(\Lambda,\Lambda)}{k!\ l!\ [n_r]_\tau\ C_\phi(I_{n_t})} \quad (\lambda_1 > \cdots > \lambda_{n_t} > 0). \tag{7.27}$$

Proof. The proof is similar to that of Theorem 7.2. □

The Rician channel capacity formulas given in Theorems 7.9 and 7.10 are difficult to compute. This difficulty motivates us to consider the approximate capacity evaluation, or specifically, finding the bound on the Rician capacity, which is studied next. The following lemma is required in the sequel.

Lemma 7.2 *The following inequality holds*

$$ {}_0F_1(b;X) < {}_0F_0\left(X/b\right), \tag{7.28} $$

where X is an $m \times m$ complex matrix and b is an arbitrary complex number.

Proof. The functions ${}_0F_1(b;X)$ and ${}_0F_0\left(X/b\right)$ are given by

$$ {}_0F_1(b;X) = \sum_{k=0}^{\infty}\sum_{\kappa}\frac{C_\kappa(X)}{[b]_\kappa\ k!} $$

and

$$ \begin{aligned} {}_0F_0\left(X/b\right) &= \sum_{k=0}^{\infty}\sum_{\kappa}\frac{C_\kappa(X/b)}{k!} \\ &= \sum_{k=0}^{\infty}\sum_{\kappa}\frac{C_\kappa(X)}{b^k\ k!}. \end{aligned} $$

Since $[b]_\kappa \geq b^k$, therefore, we have

$$ {}_0F_1(b;X) < {}_0F_0\left(X/b\right). $$

A numerical evaluation shows that this bound is tight. The proof is complete. □

The joint eigenvalue density of a complex noncentral Wishart matrix can be expressed by a bounded density function, which is given by the following theorem.

Theorem 7.11 *Let* $\mathbf{W} \sim \mathcal{C}\mathcal{W}_m(n, \Sigma_1, \Omega)$ *with* $n > m-1$. *Then* $\mathbf{W}$ *is an* $m \times m$ *positive definite Hermitian matrix with real eigenvalues. The joint density of the eigenvalues,* $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$, *of* $\mathbf{W}$ *satisfies the inequality*

$$f(\Lambda) < \frac{\pi^{m(m-1)}(\det \Sigma_1)^{-n}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)} \mathrm{etr}(-\Omega) \prod_{k=1}^{m} \lambda_k^{n-m} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 {}_0F_0(-\Psi, \Lambda), \tag{7.29}$$

where the diagonal elements of $\Psi = \mathrm{diag}(\psi_1, \ldots, \psi_m)$ *are the eigenvalues of the matrix* $\left(\Sigma_1^{-1} - \Omega\Sigma_1^{-1}/n\right)$ *and* $\Omega = \Sigma^{-1}M^H M$.

Proof. From Proposition 4.3, we obtain

$$\begin{aligned}
f(\Lambda) &= \frac{\pi^{m(m-1)}}{\mathcal{C}\Gamma_m(m)} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 \int_{U(m)} f(E\Lambda E^H)(dE) \\
&= \frac{\pi^{m(m-1)}(\det \Sigma_1)^{-n}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)} \mathrm{etr}(-\Omega) \prod_{k=1}^{m} \lambda_k^{n-m} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 \\
&\quad \cdot \int_{U(m)} \mathrm{etr}\left(-\Sigma_1^{-1}E\Lambda E^H\right) {}_0F_1\left(n; \Omega\Sigma_1^{-1}E\Lambda E^H\right)(dE) \\
&< \frac{\pi^{m(m-1)}(\det \Sigma_1)^{-n}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)} \mathrm{etr}(-\Omega) \prod_{k=1}^{m} \lambda_k^{n-m} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 \\
&\quad \cdot \int_{U(m)} \mathrm{etr}\left(-\Sigma_1^{-1}E\Lambda E^H\right) {}_0F_0\left(\Omega\Sigma_1^{-1}E\Lambda E^H/n\right)(dE) \\
&< \frac{\pi^{m(m-1)}(\det \Sigma_1)^{-n}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)} \mathrm{etr}(-\Omega) \prod_{k=1}^{m} \lambda_k^{n-m} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 \\
&\quad \cdot \int_{U(m)} \mathrm{etr}\left(-\left(\Sigma_1^{-1} - \Omega\Sigma_1^{-1}/n\right)E\Lambda E^H\right)(dE) \\
&< \frac{\pi^{m(m-1)}(\det \Sigma_1)^{-n}}{\mathcal{C}\Gamma_m(m)\mathcal{C}\Gamma_m(n)} \mathrm{etr}(-\Omega) \prod_{k=1}^{m} \lambda_k^{n-m} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 {}_0F_0(-\Psi, \Lambda). \qquad (7.30)
\end{aligned}$$

The result follows from Equation (4.14). □

Note that the function ${}_0F_0\left(-\Psi,\Lambda\right)$ can be written as [29]

$$ {}_0F_0\left(-\Psi,\Lambda\right)=\frac{C\Gamma_m(m)\det\left[\left(\exp\left(-\psi_i\lambda_j\right)\right)\right]}{\pi^{m(m-1)/2}\prod_{k<l}^{m}(\lambda_k-\lambda_l)\prod_{k<l}^{m}(\psi_l-\psi_k)}. \tag{7.31} $$

The following theorem corresponds to Theorem 7.3.

Theorem 7.12 *Consider the Rician channel, i.e.,* $\mathbf{H}\sim\mathcal{CN}(M,I_{n_r}\otimes\Sigma_1)$, *with* $n_r\geq n_t$. *If the input power is constrained by* ρ, *then using the eigenvalue density function we can bound the capacity* C *as*

$$ C<n_t\mathcal{E}_{\lambda_1}\left[\log(1+(\rho/n_t)\lambda_1)\right]. \tag{7.32} $$

The density $f(\lambda_1)$ *satisfies the inequality*

$$ f(\lambda_1)<\frac{\pi^{n_t(n_t-1)/2}(\det\Sigma_1)^{-n_r}\operatorname{etr}(-\Omega)}{n_t!C\Gamma_{n_t}(n_r)\prod_{k<l}^{n_t}(\psi_l-\psi_k)}\int\left\{\widetilde{\sum}_i(-1)^{\operatorname{per}(i_1,\ldots,i_{n_t})}\exp\left(\sum_{j=1}^{n_t}-\lambda_j\psi_{i_j}\right)\right\} \cdot\left\{\widetilde{\sum}_k(-1)^{\operatorname{per}(k_1,\ldots,k_{n_t})}\prod_{l=1}^{n_t}\lambda_l^{n_r-n_t+k_l}\right\}\bigwedge_{k=2}^{n_t}d\lambda_k, \tag{7.33} $$

where $\widetilde{\sum}_i$ *denotes summation over all permutations* $(i_1,\ldots,i_{n_t})$ *of* $(1,\ldots,n_t)$, $\widetilde{\sum}_k$ *denotes summation over all permutations* $(k_1,\ldots,k_{n_t})$ *of* $(0,\ldots,n_t-1)$ *and* $\operatorname{per}(k_1,\ldots,k_{n_t})$ *is* 0 *or* 1 *depending on the permutation being even or odd. Similarly for* $\operatorname{per}(i_1,\ldots,i_{n_t})$.

Proof. The proof is similar to that of Theorem 7.3. □

7.3.1 Rician $n_r\times 2$ channel matrix

In this section, a numerical evaluation of a $n_r\times 2$ Rician channel matrix is given. Thus, we assume that we have a two-input ($n_t=2$), n_r-output communication system operating over a Rician fading environment (typical satellite communication environment).

Let $n_t = 2$ and $\Psi = \text{diag}(\psi_1, \psi_2)$. Then we have [29]

$$ {}_0F_0(-\Psi, \Lambda) = \frac{1}{(\psi_2 - \psi_1)(\lambda_1 - \lambda_2)} \left[\exp\left\{-(\psi_1\lambda_1 + \psi_2\lambda_2)\right\} - \exp\left\{-(\psi_1\lambda_2 + \psi_2\lambda_1)\right\}\right]. \tag{7.34} $$

The following theorem gives the Rician channel capacity for a $n_r \times 2$ matrix.

Theorem 7.13 *Consider the two-input Rician channel, i.e.,* $\mathbf{H} \sim \mathcal{CN}(M, I_{n_r} \otimes \Sigma_1)$, *with* $n_r \geq 2$. *If the input power is constrained by* ρ, *then the capacity* C *satisfies the inequality*

$$ \begin{aligned} C \; < \; & \frac{(\det \Sigma_1)^{-n_r} \operatorname{etr}(-\Omega)}{(\psi_2 - \psi_1)} \left[\frac{1}{\Gamma(n_r)} \int_0^\infty \log[1 + (\rho/2)\lambda_1] \lambda_1^{n_r - 1} e^{-\psi_1 \lambda_1} d\lambda_1 \right. \\ & - \frac{1}{\Gamma(n_r)} \int_0^\infty \log[1 + (\rho/2)\lambda_1] \lambda_1^{n_r - 1} e^{-\psi_2 \lambda_1} d\lambda_1 \\ & - \frac{1}{\Gamma(n_r - 1)} \int_0^\infty \log[1 + (\rho/2)\lambda_1] \lambda_1^{n_r - 2} e^{-\psi_1 \lambda_1} d\lambda_1 \\ & \left. + \frac{1}{\Gamma(n_r - 1)} \int_0^\infty \log[1 + (\rho/2)\lambda_1] \lambda_1^{n_r - 2} e^{-\psi_2 \lambda_1} d\lambda_1 \right], \end{aligned} \tag{7.35} $$

where λ_1 *is an eigenvalue of* $W = H^H H$ *and* (ψ_1, ψ_2) *are the eigenvalues of* $\left(\Sigma_1^{-1} - \Omega\Sigma_1^{-1}/n_r\right)$.

Proof. Using Equation (7.34), the eigenvalue density of W satisfies the inequality

$$ f(\lambda_1, \lambda_2) < \frac{(\det \Sigma_1)^{-n_r} \operatorname{etr}(-\Omega)(\lambda_1\lambda_2)^{n_r - 2}(\lambda_1 - \lambda_2)}{2(\psi_2 - \psi_1)\Gamma(n_r)\Gamma(n_r - 1)} \left[e^{-\psi_1\lambda_1 - \psi_2\lambda_2} - e^{-\psi_1\lambda_2 - \psi_2\lambda_1}\right]. \tag{7.36} $$

Now, integrating with respect to λ_2 and noting that

$$ \int_0^\infty x^{a-1} e^{-x/b} dx = \Gamma(a) b^a, $$

we obtain the density of $f(\lambda_1)$. Thus we have

$$ \begin{aligned} f(\lambda_1) < & \frac{(\det \Sigma_1)^{-n_r} \operatorname{etr}(-\Omega)}{2(\psi_2 - \psi_1)} \\ & \cdot \left\{ \frac{\lambda_1^{n_r - 1} e^{-\psi_1\lambda_1}}{\Gamma(n_r)\psi_2^{n_r - 1}} - \frac{\lambda_1^{n_r - 1} e^{-\psi_2\lambda_1}}{\Gamma(n_r)\psi_1^{n_r - 1}} - \frac{\lambda_1^{n_r - 2} e^{-\psi_1\lambda_1}}{\Gamma(n_r - 1)\psi_2^{n_r}} + \frac{\lambda_1^{n_r - 2} e^{-\psi_2\lambda_1}}{\Gamma(n_r - 1)\psi_1^{n_r}} \right\}. \end{aligned} \tag{7.37} $$

Finally, evaluating Equation (7.7) with $f(\lambda_1)$ gives Equation (7.35). □

Table 7.4 shows the capacity in nats[2] for a $n_r \times 2$ Rician fading channel matrix. Note that each column represents different levels of input power or signal to noise ratio (SNR) in dB. Figure 7.6 shows the capacity in nats vs n_r for different levels of input power. From the table and figure we note that the capacity is increasing with increasing n_r and SNR. Moreover, the Rician channel capacity is decreasing compared to an uncorrelated Rayleigh channel capacity, see Table 7.3.

[2] In Equation (7.35), 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 ($e = 2.718\ldots$).

	ρ in dB							
n_r	0 dB	5 dB	10 dB	15 dB	20 dB	25 dB	30 dB	35 dB
2	1.1397	2.0863	3.2236	4.4354	5.6692	6.9088	8.1501	9.3919
4	1.5869	2.6396	3.7924	4.9736	6.1620	7.3524	8.5434	9.7346
6	1.9130	3.0056	4.1603	5.3311	6.5063	7.6827	8.8595	10.0364
8	2.1628	3.2741	4.4287	5.5945	6.7633	7.9332	9.1032	10.2734
10	2.3641	3.4856	4.6399	5.8027	6.9679	8.1338	9.3000	10.4663
12	2.5321	3.6600	4.8140	5.9748	7.1376	8.3010	9.4646	10.6283
14	2.6763	3.8084	4.9621	6.1215	7.2826	8.4443	9.6061	10.7679
16	2.8024	3.9375	5.0909	6.2493	7.4092	8.5695	9.7300	10.8905
18	2.9144	4.0518	5.2050	6.3626	7.5215	8.6808	9.8402	10.9997
20	3.0152	4.1543	5.3074	6.4643	7.6224	8.7809	9.9395	11.0981

Table 7.4: The capacity in nats for a two-input, n_r-output communication system operating over a Rician fading channel, where ρ is signal to noise ratio in dB.

Note that here we assumed $\mathbf{H} \sim \mathcal{CN}(M, I_{n_r} \otimes \Sigma_1)$, where the covariance matrix is $\Sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and the mean matrix is $M = \begin{bmatrix} 0.25 + .25i & 0.25 + 0.25i \\ 0.25 + .25i & 0.25 + 0.25i \end{bmatrix}$.

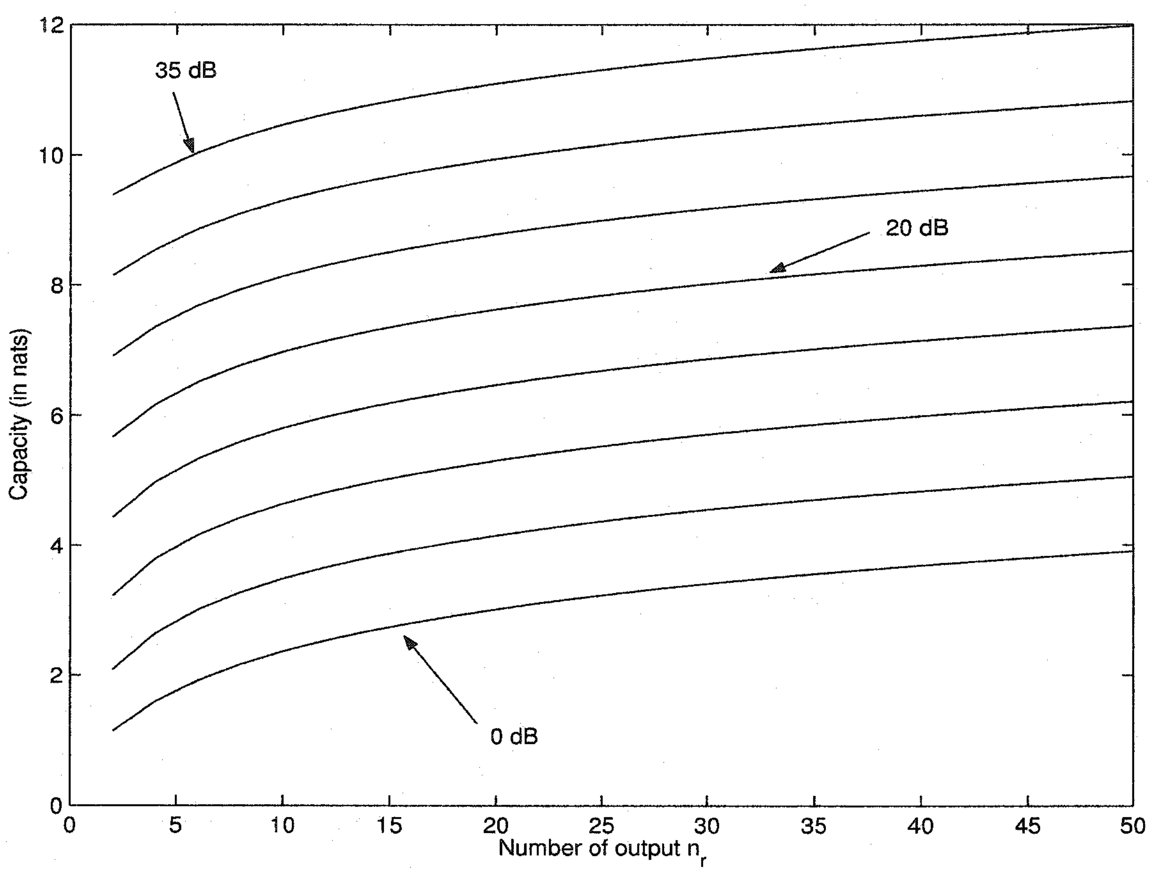

Figure 7.6: Capacity vs number of outputs for SNR=$0, 5, 10, 15, 20, 25, 30, 35$ dB. Note that $\mathbf{H}$ is a $n_r \times 2$ Rician fading channel matrix.

7.4 Error exponents for random channel

For a random matrix channel, $E_0(\alpha, f(x))$ is defined as [45]

$$E_0(\alpha, f(x)) = -\log \int\int \left[\int f(x)p(y, H|x)^{1/(1+\alpha)}\, dx\right]^{1+\alpha} dy\, dH.$$

Note that $\mathbf{H}$ is independent of $\mathbf{x}$, therefore, $p(y, H|x) = g(H)p(y|x, H)$ and

$$E_0(\alpha, f(x)) = -\log \mathcal{E}_{\mathbf{H}} \left[\int \left[\int f(x)p(y|x, \mathbf{H})^{1/(1+\alpha)}\, dx\right]^{1+\alpha} dy\right].$$

The following proposition gives the random coding bound for a known random channel matrix $\mathbf{H}$.

Proposition 7.1 *Let $x \sim \mathcal{CN}(0, R_{xx})$ and*

$$p(y|x) = \det(\pi I)\exp\left\{-(y - Hx)^H(y - Hx)\right\}.$$

Then

$$E_0(\alpha, f(x)) = -\log \mathcal{E}_{\mathbf{H}} \left\{\det\left(I + (1+\alpha)^{-1}\mathbf{H}R_{xx}\mathbf{H}^H\right)^{-\alpha}\right\} \tag{7.38}$$

and

$$E_0(\alpha) = \max_{f(x), \mathrm{tr}(R_{xx}) \le \rho} E_0(\alpha, f(x)) = -\log \mathcal{E}_{\mathbf{H}} \left\{\det\left(I + \frac{\rho}{n_t(1+\alpha)}\mathbf{H}\mathbf{H}^H\right)^{-\alpha}\right\}. \tag{7.39}$$

The upper bound for the probability of error is given by

$$P(\text{error}) \le \exp\left\{-n\left[\max_{0\le\alpha\le1}(E_0(\alpha) - \alpha R)\right]\right\}.$$

Proof. The proof is similar to that of Proposition 6.4. □

7.5 Matlab programs

The Matlab programs that have been used in this thesis to compute the channel capacities are given here.

Correlated Rayleigh channel

```
clear all
% This program compute the correlated Rayleigh channel capacity
% Rhh=[1 .4;.4 1]
% eigenvalues of Rhh are v2=.6, v1=1.4; a1=1/v1, a2=1/v2
sig2=1; % noise power or noise variance
a2=1/.6; a1=1/1.4;
SNRdb=0:5:35; % signal power
for rho=1:length(SNRdb)
snr=10^(SNRdb(rho)/10);
k=0;
for n=2:2:50
k=k+1;a=0;i=0;
for x=0:0.1:1000
i=i+1;
a(i)= ...
(((a1^n)*a2)/((a2-a1)*factorial(n-1)) )*(exp(-a1*x)*x^(n-1)*log(1+(snr/2)*x)*0.1)-...
(((a2^n)*a1)/((a2-a1)*factorial(n-1)) )*(exp(-a2*x)*x^(n-1)*log(1+(snr/2)*x)*0.1)-...
(((a1^n))/((a2-a1)*factorial(n-2)) )*(exp(-a1*x)*x^(n-2)*log(1+(snr/2)*x)*0.1)+...
(((a2^n))/((a2-a1)*factorial(n-2)) )*(exp(-a2*x)*x^(n-2)*log(1+(snr/2)*x)*0.1);
end
b(k,rho)=sum(a);
end
end
```

Rician channel

```
clear all
% This program compute the correlated Rician channel capacity
R=[1 0;0 1]; M=(0.25+i*0.25)*ones(2,2);
OM=inv(R)*M'*M;
SNRdb=0:5:35;

for rho=1:length(SNRdb)
snr=(10^(SNRdb(rho)/10))/(M(1,1)*M(1,1)'+1);
k=0;
for n=2:2:50
A=inv(R)-inv(R)*M'*M*inv(R)/n;
aa=sort(eig(A)); a1=aa(1); a2=aa(2);
k=k+1;a=0;m=0;
for x=0:0.1:1000
m=m+1;
Kk=(((det(R))^(-n))*exp(trace(-OM))/(a2-a1) );
a(m)= (Kk/(factorial(n-1)) )*(exp(-a1*x)*(x^(n-1))*log(1+(snr/2)*x)*0.1)-...
      (Kk/(factorial(n-1)) )*(exp(-a2*x)*(x^(n-1))*log(1+(snr/2)*x)*0.1)-...
      (Kk/(factorial(n-2)) )*(exp(-a1*x)*(x^(n-2))*log(1+(snr/2)*x)*0.1)+...
      (Kk/(factorial(n-2)) )*(exp(-a2*x)*(x^(n-2))*log(1+(snr/2)*x)*0.1);

end
b(k,rho)=sum(a);
end
end
```

Bibliography

[1] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, Wiley, 1986.

[2] R. B. Ash, *Information Theory*, Dover, 1965.

[3] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed, Athena Scientific, Belmont, MA, 1999.

[4] E. Biglieri, J. Proakis, S. Shamai (Shitz), "Fading channels: Information-Theoretic and communications aspects," *IEEE Trans. on Information Theory*, Vol. 44, No. 6. pp. 2619–2692, Oct. 1998.

[5] B. V. Bronk, "Exponential ensembles for random matrices," *J. of Math. Physics*, Vol. 6, pp. 228–237, 1965.

[6] Y. Chikuse, "Distributions of some matrix variates and latent roots in multivariate Behrens-Fisher discriminant analysis," *Ann. of Statist.*, Vol. 9, pp. 401–407, 1981.

[7] Y. Chikuse and A. W. Davis, "Some properties of invariant polynomials with matrix arguments and their applications in econometrics," *Ann. Inst. Statist. Math.*, Vol. 38, pp. 109–122, 1986.

[8] A. G. Constantine, "Some noncentral distribution problems in multivariate analysis," *Ann. Math. Statist.*, Vol. 34, pp. 1270–1285, 1963.

[9] A. G. Constantine, "The distribution of Hotelling's generalized T_0^2," *Ann. Math. Statist.*, Vol. 37, pp. 215–225, 1966.

[10] A. W. Davis, "Invariant polynomials with two matrix arguments extending the zonal polynomials: Applications to multivariate distribution theory," *Ann. Inst. Statist. Math.*, Vol. 31, pp. 465–485, 1979.

[11] A. W. Davis, "Invariant polynomials with two matrix arguments extending the zonal polynomials," *Multivariate Analysis – V* (ed. P. K. Krishnaiah), North-Holland, Amsterdam, pp. 287–299, 1980.

[12] A. Edelman, *Eigenvalues and condition numbers of random matrices*, Ph.D. Dissertation, MIT, Cambridge, MA, May 1989.

[13] R. G. Gallager, *Information Theory and Reliable Communication*, Wiley, 1968.

[14] H. H. Goldstine and J. von Neumann, "Numerical inverting of matrices of high order II," *Amer. Math. Soc. Proc.*, Vol. 2, pp. 188–202, 1951.

[15] A. Gupta and D. Nagar, *Matrix Variate Distributions*, Chapman & Hall, 2000.

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

[17] N. R. Goodman, "The distribution of the determinant of complex Wishart distributed matrix," *Ann. Math. Statist.*, Vol. 34, pp. 178–180, 1963.

[18] R. Gutierrez, J. Rodriguez, and A. J. Saez, "Approximation of hypergeometric functions with matricial argument through their development in series of zonal polynomials," *Electronic Trans. on Numerical Analysis*, Vol. 11, pp. 121-130, 2000.

[19] C. S. Herz, "Bessel functions of matrix argument," *Ann. Math.*, Vol. 61, pp. 474–523, 1955.

[20] A. T. James, "Normal multivariate analysis and the orthogonal group," *Ann. Math. Statist.*, Vol. 25, pp. 40–75, 1954.

[21] A. T. James, "The distribution of the latent roots of the covariance matrix," *Ann. Math. Statist.*, Vol. 31, pp. 151–158, 1960.

[22] A. T. James, "The distribution of noncentral means with known covariance matrix," *Ann. Math. Statist.*, Vol. 32, pp. 874–882, 1961.

[23] A. T. James, "Zonal polynomials of the real positive definite symmetric matrices," *Ann. Math.*, Vol. 74, pp. 456–469, 1961.

[24] A. T. James, "Distributions of matrix variate and latent roots derived from normal samples," *Ann. Math. Statist.*, Vol. 35, pp. 475–501, 1964.

[25] C. G. Khatri, "Distribution of the largest or the smallest characteristic root under null hypothesis concerning complex multivariate normal populations," *Ann. Math. Statist.*, Vol. 35, pp. 1807–1810, 1964.

[26] C. G. Khatri, "Classical statistical analysis based on a certain multivariate complex Gaussian distribution," *Ann. Math. Statist.*, Vol. 36, pp. 98–114, 1965.

[27] C. G. Khatri, "On certain distribution problems based on positive definite quadratic functions in normal vectors," *Ann. Math. Statist.*, Vol. 37, pp. 468–479, 1966.

[28] C. G. Khatri and K. C. S. Pillai, "On the non-central distributions of two test criteria in multivariate analysis of variance," *Ann. Math. Statist.*, Vol. 39, pp. 215–226, 1968.

[29] C. G. Khatri, "Non-central distributions of ith largest characteristic roots of three matrices concerning complex multivariate normal populations," *Ann. Inst. Statist. Math.*, Vol. 21, pp. 23–32, 1969.

[30] C. G. Khatri, "On the moments of traces of two matrices in three situations for complex multivariate normal populations," *Sankhyā*, Vol. A32, pp. 65–80, 1970.

[31] A. M. Mathai and S. B. Provost, *Quadratic forms in Random Variables: Theory and Applications*, Marcel Dekker, 1992.

[32] M. L Mehta and M. Gaudin, "On the density of eigenvalues of a random matrix," *Nuclear Physics*, Vol. 18, pp. 420–427, 1960.

[33] M. L Mehta, *Random Matrices*, 2nd ed, Academic Press, 1991.

[34] K. S. Miller, *Complex Stochastic Processes: An Introduction to Theory and Application*, Addison-Wesley, 1974.

[35] L. Mirsky, *An Introduction to Linear Algebra*, Oxford University Press, London and New York, 1955.

[36] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, Wiley, 1982.

[37] C. E. Porter, *Statistical Theories of Spectra: Fluctuations*, Academic Press, 1965.

[38] Sabri Al-Ani, "On the distribution of the ith latent root under null hypotheses concerning complex multivariate normal populations," *Canad. Math. Bull.*, Vol. 15 (3), pp. 323–327, 1972.

[39] C. E. Shannon, "A mathematical theory of communication," *Bell System Tech. J.* Vol. 27, pp. 623–656, 1948.

[40] S. Smale, "On the efficiency of algorithms of analysis," *Bull. Amer. Math. Soc.*, Vol. 13, pp. 87–121, 1985.

[41] T. Sugiyama, "On the distribution of the largest latent root of the covariance matrix," *Ann. Math. Statist.*, Vol. 38, pp. 1148–1151, 1967.

[42] T. Sugiyama, "Joint distribution of the extreme roots of a covariance matrix," *Ann. Math. Statist.*, Vol. 41, pp. 655–657, 1970.

[43] T. Sugiyama, "Distributions of the largest latent root of the multivariate complex Gaussian distribution," *Ann. Inst. Statist. Math.*, Vol. 24, pp. 87–94, 1972.

[44] A. Takemura, *Zonal Polynomials*, Institute of Mathematical Statistics, Hayward, CA, Vol. 4, 1984.

[45] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," AT & T Bell-Lab Technical Memorandum 1995. Later published in *Eur. Trans. Telecom*, Vol. 10, pp. 585–595, Nov. 1999.

[46] J. von Neumann and H. H. Goldstine, "Numerical inverting of matrices of high order," *Bull. Amer. Math. Soc.*, Vol. 53, pp. 1021–1099, 1947.

[47] V. B. Waikar and F. J. Schuurmann, "Exact joint density of the largest and smallest roots of the Wishart and MANOVA matrices," *Utilitas Mathematica*, Vol. 4, pp. 253–260, 1973.

[48] V. B. Waikar, "On the joint distributions of the largest and smallest latent roots of two random matrices (noncentral case)," *South African Statistical Journal*, Vol. 7, pp. 103–108, 1973.

[49] E. Wigner, "Distribution laws for roots of a random Hermitian matrix," *Statistical theories of spectra: Fluctuations* (ed. C. E. Porter), Academic Press, pp. 446–461, 1965.

[50] E. Wigner, "Random matrices in physics," *SIAM Review*, Vol. 9, pp. 1–23, 1967.

[51] J. H. Wilkinson, "Error analysis revisited," *IMA Bulletin*, Vol. 22, pp. 192–200, 1986.

[52] S. S. Wilks, *Mathematical Statistics*, John Wiley & Sons, 1962.