Quadratic Forms on Complex Random Matrices and Multi-Antenna Channel Capacity

T. Ratnarajah[*][†] R. Vaillancourt[*][‡]

CRM-2979

March 2004

[*]Department of Mathematics and Statistics, University of Ottawa, 585 King Edward Ave., Ottawa ON K1N 6N5 Canada
[†]t.ratnarajah@ieee.org
[‡]remi@uottawa.ca

Abstract

Quadratic forms on complex random matrices and their joint eigenvalue densities are derived for applications in information theory. These densities are represented by complex hypergeometric functions of matrix arguments, which can be expressed in terms of complex zonal polynomials. The derived densities are used to evaluate the most important information-theoretic measures, the so-called ergodic channel capacity and capacity versus outage of multiple-input multiple-output (MIMO) Rayleigh distributed wireless communication channels. Both (spatially) correlated and uncorrelated channels are considered and the corresponding information-theoretic measure formulas are derived. It is shown how channel correlation degrades the communication system capacity.

To appear in Proc. The Twelfth Annual Workshop on Adaptive Sensor Array Processing, MIT Lincoln Laboratory, Lexington MA

Résumé

On obtient des formes quadratiques sur les matrices aléatoires complexes et la densités de leurs valeurs propres appliquées à la théorie de l'information. On représente ces densités par des fonctions hypergéométriques complexes sur des matrices et au moyen de polynômes zonaux complexes. Ces densités servent à évaluer les deux plus importantes mesures en théorie de l'information, à savoir la capacité ergodique d'un canal et la capacité contre la perte dans un canal de communication sans fil de Rayleigh non corrélés à entrées et sorties multiples (MIMO). On considère des canaux corrélés et non corrélés dans l'espace et on montre que la corrélation dégrade le système.

1 INTRODUCTION

Let an $n \times m$ complex Gaussian random matrix X be distributed as $\mathrm{X} \sim \mathcal{C}N(0, \Sigma_1 \otimes \Sigma_2)$ with mean $\mathcal{E}\{\mathrm{X}\} = 0$ and covariance $\mathrm{cov}\{\mathrm{X}\} = \Sigma_1 \otimes \Sigma_2$, where $\Sigma_1 \in \mathbb{C}^{n\times n}$ and $\Sigma_2 \in \mathbb{C}^{m\times m}$ are positive definite Hermitian matrices. Then the quadratic form on X associated with the positive definite Hermitian matrix A is defined by

$$\mathrm{S} = \mathrm{X}^H A \mathrm{X}.$$

Here, we study the distribution of S, denoted by $\mathcal{C}Q_{n,m}(A, \Sigma_1, \Sigma_2)$, and its application to information theory. We also derive the joint eigenvalue densities of S, which are represented by complex zonal polynomials. Complex zonal polynomials are symmetric polynomials in the eigenvalues of a complex matrix, see [5], and they enable us to represent the derived densities as infinite series. If $A = I_n$, $\Sigma_1 = I_n$ and $\Sigma_2 = \Sigma$, then $\mathrm{S} = \mathrm{X}^H\mathrm{X}$ is said to have a complex Wishart distribution, denoted by $\mathcal{C}W_m(n, \Sigma)$, see [3], [6] and references therein.

The theory of quadratic forms on complex random matrices is used to evaluate the capacity of multiple-input, multiple-output (MIMO) wireless communication systems. Let us denote the number of inputs (or transmitters) and the number of outputs (or receivers) of the MIMO wireless communication system by n_t and n_r, respectively, and assume that the channel coefficients are distributed as complex Gaussian and correlated at both the transmitter and the receiver ends. Then the MIMO channel can be represented by an $n_r \times n_t$ complex random matrix $\mathrm{H} \sim \mathcal{C}N(0, \Sigma_r \otimes \Sigma_t)$, where Σ_r and Σ_t represent the channel correlations at the receiver and transmitter ends, respectively. This means that the covariance matrices of the columns and rows of H are denoted by Σ_r and Σ_t, respectively. If $\Sigma_r = I_{n_r}$ (or $\sigma^2 I_{n_r}$) and $\Sigma_t = \sigma^2 I_{n_t}$ (or I_{n_t}) then the channel is called *uncorrelated Rayleigh distributed channel*. The information processed by this random channel (or mutual information of this random channel) is a random quantity which can be measured in two ways, namely, ergodic capacity (or average mutual information) and capacity versus outage (or x percent outage).

Recent studies show that a MIMO uncorrelated Rayleigh distributed channel achieves almost n more bits per hertz for every 3dB increase in signal-to-noise ratio (SNR) compared to a single-input single-output (SISO) system, which achieves only one additional bit per hertz for every 3dB increase in SNR, where $n = \min\{n_t, n_r\}$, see [2] and [12]. In this context, we may mention a MIMO strategy that offers a tremendous potential to increase the information capacity of single user wireless communication systems, namely, the *Bell-Labs Layered Space-Time (BLAST)* architecture, see [8] and references therein. But in many practical situation, channel correlation exists due to poor scattering conditions, which degrades capacity (see [1], [9], [10] and references therein). In [1], the authors studied the correlated channel matrix with large dimension (asymptotic analysis), which is only an approximation to the practical correlated channel matrix with finite dimension. In [10], the effect of channel correlations on MIMO capacity is quantified by employing an abstract scattering model. More recently, in [9] the authors have derived an exact ergodic capacity expression for an uncorrelated Rayleigh MIMO channels, which is different from the work of Telatar [12], and have also derived an upper bound to ergodic capacity for correlated Rayleigh MIMO channels. In this paper, we first derive the densities of quadratic forms on complex random matrices and their joint eigenvalue densities. Then, using these densities we evaluate the capacity degradation for the correlated channel matrices $\mathrm{H} \sim \mathcal{C}N(0, \Sigma_r \otimes \Sigma_t)$ with $n_r > n_t$, $n_r = n_t$ and $n_t > n_r$. This will be done by deriving closed-form ergodic capacity and outage capacity formulas for correlated channels and their numerical evaluation. Note that this work is an extension of our early work [6], where we have extensively studied the complex Wishart matrices and uncorrelated/correlated (only at the transmitter end) MIMO channel capacities.

This paper is organized as follows. Quadratic forms on complex random matrices are studied in Section 2. The capacity of a MIMO channel and the computational methods are given in Sections 3 and 4, respectively.

2 Quadratic forms on complex random matrices

In this section, the densities of quadratic forms on complex random matrices are given and their joint eigenvalue densities are derived. The probability distributions of random matrices are often derived in terms of hypergeometric functions of matrix arguments. In the sequel, we need to use the following complex hypergeometric function of two matrix arguments,

$$ {}_0F_0^{(m)}(X, Y) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{C_\kappa(X) C_\kappa(Y)}{k! C_\kappa(I_n)}, \tag{1}$$

where $X \in \mathbb{C}^{m\times m}$, $Y \in \mathbb{C}^{n\times n}$ and $n \geq m$. Moreover, $\kappa = (k_1, \ldots, k_m)$ denotes a partition of the integer k with $k_1 \geq \cdots \geq k_m \geq 0$ and $k = k_1 + \cdots + k_m$, and $\sum_\kappa$ denotes summation over all partitions κ of k. The complex zonal polynomial, $C_\kappa(X)$, of a complex matrix X defined in [3] is

$$C_\kappa(X) = \chi_{[\kappa]}(1)\chi_{[\kappa]}(X), \tag{2}$$

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

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

and $\chi_{[\kappa]}(X)$ is the character of the representation $[\kappa]$ of the linear group given by

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

as a symmetric function of the eigenvalues, $\lambda_1, \ldots, \lambda_m$, of X. Note that real and complex zonal polynomials are particular cases of (general α) Jack polynomials, $C_\kappa^{(\alpha)}(X)$, where $\alpha = 1$ for complex, and $\alpha = 2$ for real, zonal polynomials, respectively. In this paper we only consider the complex case; therefore, for notational simplicity we drop the superscript, α, of Jack polynomials, as was done in equation (2), i.e., we write $C_\kappa(X) := C_\kappa^{(1)}(X)$. Finally, we have

$$C_\kappa(I_n) = 2^{2k} k! \left(\frac{1}{2}\, n\right)_\kappa \frac{\prod_{i<j}^{r}(2k_i - 2k_j - i + j)}{\prod_{i=1}^{r}(2k_i + r - i)!}, \tag{3}$$

where

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

and the partition κ of k has r nonzero parts. Here $(a)_k = a(a+1)\cdots(a+k-1)$ and $(a)_0 = 1$.

The next theorem gives the density of quadratic forms on complex random matrices $\mathrm{S} = \mathrm{X}^H A \mathrm{X}$, see [7].

Theorem 1 *Let* X *be an* $n \times m$ $(n \geq m)$ *complex Gaussian random matrix distributed as* $\mathrm{X} \sim \mathcal{C}N(0, \Sigma_1 \otimes \Sigma_2)$*, where* $\Sigma_1 \in \mathbb{C}^{n\times n}$ *and* $\Sigma_2 \in \mathbb{C}^{m\times m}$ *are positive definite Hermitian matrices. Then the density function of* $\mathrm{S} = \mathrm{X}^H A \mathrm{X}$ *is given by*

$$\begin{aligned} f(S) \quad &= \quad \frac{1}{\mathcal{C}\Gamma_m(n)(\det\Sigma_1)^m\,(\det\Sigma_2 A)^n}(\det S)^{n-m} \\ &\quad \times {}_0F_0^{(m)}(B, -\Sigma_2^{-1}S), \end{aligned} \tag{4}$$

where $A \in \mathbb{C}^{n\times n}$ *is a positive definite Hermitian matrix and* $B = A^{-1/2}\Sigma_1^{-1}A^{-1/2}$.

The distribution of the matrix S is denoted by $\mathcal{C}Q_{n,m}(A, \Sigma_1, \Sigma_2)$. It should be noted that the density given in Theorem 1 is different from the one given in [4]. From this generalized density we can easily derived other well-known densities. Special cases of density (4) are:
(i) If $A = I_n$, then the density of $\mathrm{S} = \mathrm{X}^H\mathrm{X} \sim \mathcal{C}Q_{n,m}(I_n, \Sigma_1, \Sigma_2)$ is given by

$$\begin{aligned} f(S) \quad &= \quad \frac{1}{\mathcal{C}\Gamma_m(n)(\det\Sigma_1)^m(\det\Sigma_2)^n}(\det S)^{n-m} \\ &\quad \times {}_0F_0^{(m)}(\Sigma_1^{-1}, -\Sigma_2^{-1}S). \end{aligned} \tag{5}$$

(ii) If $A = I_n$, $\Sigma_1 = I_n$ and $\Sigma_2 = \Sigma$, then $\mathrm{S} = \mathrm{X}^H\mathrm{X}$ is said to have a complex Wishart distribution, denoted by $\mathcal{C}W_m(n, \Sigma)$, with density

$$f(S) = \frac{1}{\mathcal{C}\Gamma_m(n)(\det\Sigma)^n}\,(\det S)^{n-m}\,\mathrm{etr}(-\Sigma^{-1}S), \tag{6}$$

where etr denotes the exponential of the trace, $\mathrm{etr}(\cdot) = \exp(\mathrm{tr}(\cdot))$.
(iii) If $A = I_n$, $\Sigma_1 = I_n$ and $\Sigma_2 = \sigma^2 I_m$, then $\mathrm{S} = \mathrm{X}^H\mathrm{X} \sim \mathcal{C}W_m(n, \sigma^2 I_m)$ is a complex Wishart matrix with density

$$f(S) = \frac{(\sigma^2)^{-mn}}{\mathcal{C}\Gamma_m(n)}\,(\det S)^{n-m}\,\mathrm{etr}\left(-\frac{1}{\sigma^2}\,S\right). \tag{7}$$

The next theorem gives the joint eigenvalue density of quadratic forms on complex random matrices. This theorem is one of the key contributions of this paper. Moreover, from this theorem we can easily derived other joint eigenvalue densities of complex random matrices, see [7].

Theorem 2 *Consider the $m \times m$ positive definite Hermitian matrix $\mathrm{S} = \mathrm{X}^H A \mathrm{X} \sim \mathcal{C}Q_{n,m}(A, \Sigma_1, \Sigma_2)$, where $A \in \mathbb{C}^{n\times n}$ is a positive definite Hermitian matrix. Then the joint density of the eigenvalues, $\lambda_1 > \lambda_2 > \cdots > \lambda_m > 0$, of S is*

$$f(\Lambda) = \frac{\pi^{m(m-1)}(\det \Sigma_2 A)^{-n}}{\mathcal{C}\Gamma_m(n)\mathcal{C}\Gamma_m(m)(\det \Sigma_1)^m} \prod_{k=1}^{m} \lambda_k^{n-m} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 \times \sum_{k=0}^{\infty} \sum_{\kappa} \frac{C_\kappa(B) C_\kappa(-\Sigma_2^{-1}) C_\kappa(\Lambda)}{k! C_\kappa(I_n) C_\kappa(I_m)}, \tag{8}$$

where $B = A^{-1/2}\Sigma_1^{-1}A^{-1/2}$.

Special cases of Theorem 2 are:
(i) If $A = I_n$, then the joint eigenvalue density of $\mathrm{S} = \mathrm{X}^H\mathrm{X} \sim \mathcal{C}Q_{n,m}(I_n, \Sigma_1, \Sigma_2)$ is given by

$$f(\Lambda) = \frac{\pi^{m(m-1)}(\det \Sigma_2)^{-n}}{\mathcal{C}\Gamma_m(n)\mathcal{C}\Gamma_m(m)(\det \Sigma_1)^m} \prod_{k=1}^{m} \lambda_k^{n-m} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 \times \sum_{k=0}^{\infty} \sum_{\kappa} \frac{C_\kappa(\Sigma_1^{-1}) C_\kappa(-\Sigma_2^{-1}) C_\kappa(\Lambda)}{k! C_\kappa(I_n) C_\kappa(I_m)}. \tag{9}$$

(ii) If $A = I_n$, $\Sigma_1 = I_n$ and $\Sigma_2 = \Sigma$, then the joint eigenvalue density of the complex Wishart matrix $\mathrm{S} = \mathrm{X}^H\mathrm{X} \sim \mathcal{C}W_m(n, \Sigma)$ is given by

$$f(\Lambda) = \frac{\pi^{m(m-1)}(\det \Sigma)^{-n}}{\mathcal{C}\Gamma_m(n)\mathcal{C}\Gamma_m(m)} \prod_{k=1}^{m} \lambda_k^{n-m} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 \times {}_0F_0^{(m)}(-\Sigma^{-1}, \Lambda). \tag{10}$$

(iii) If $A = I_n$, $\Sigma_1 = I_n$ and $\Sigma_2 = \sigma^2 I_m$, then the joint eigenvalue density of the complex Wishart matrix $\mathrm{S} = \mathrm{X}^H\mathrm{X} \sim \mathcal{C}W_m(n, \sigma^2 I_m)$ is given by

$$f(\Lambda) = \frac{\pi^{m(m-1)}(\sigma^2)^{-nm}}{\mathcal{C}\Gamma_m(n)\mathcal{C}\Gamma_m(m)} \prod_{k=1}^{m} \lambda_k^{n-m} \prod_{k<l}^{m} (\lambda_k - \lambda_l)^2 \times \exp\left(-\frac{1}{\sigma^2}\sum_{k=1}^{m} \lambda_k\right). \tag{11}$$

Next, we extend the study of complex central Wishart distribution to the singular case, where $0 < n < m$ and $n, m \in \mathbb{Z}$. Thus the rank of $\mathrm{S} \in \mathbb{C}^{m\times m}$ is n. If $0 < n < m$, then the density does not exist for $\mathrm{S} \sim \mathcal{C}W_m(n, \Sigma)$ on the space of Hermitian $m \times m$ matrices, because S is singular and of rank n almost surely. However, it can be shown that the density does exist on the $(2mn - n^2)$-dimensional manifold, $\mathcal{C}S_{m,n}$, of rank n of positive semidefinite $m \times m$ Hermitian matrices S with n distinct positive eigenvalues. Moreover, the set of all $m \times n$ matrices, E_1, with orthonormal columns is called the *Stiefel manifold*, denoted by $\mathcal{C}V_{n,m}$. Thus,

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

The application of the next two theorems is another key contribution of this paper. The first of these theorems gives the complex singular Wishart density, see [7].

Theorem 3 *The density of $\mathrm{S} = \mathrm{X}^H\mathrm{X} \sim \mathcal{C}W_m(n, \Sigma)$ on the space $\mathcal{C}S_{m,n}$ of rank n $(0 < n < m)$ positive semidefinite $m \times m$ Hermitian matrices is given by*

$$f(S) = \frac{\pi^{n(n-m)}}{\mathcal{C}\Gamma_n(n)(\det \Sigma)^n} \operatorname{etr}\left(-\Sigma^{-1}S\right)(\det \Lambda)^{n-m}, \tag{12}$$

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

The second of these theorems gives the joint eigenvalue density of a complex singular Wishart matrix, see [7].

Theorem 4 *Let* $\mathrm{S} = \mathrm{X}^H\mathrm{X} \sim \mathcal{C}W_m(n,\Sigma)$ *with* $0 < n < m$ *so that* S *is an* $m \times m$ *positive semidefinite Hermitian matrix with* n *positive eigenvalues. Then the joint density of the eigenvalues,* $\lambda_1, \ldots, \lambda_n$*, of* S *is*

$$\begin{aligned} f(\Lambda) &= \frac{\pi^{n(n-1)}(\det\Sigma)^{-n}}{\mathcal{C}\Gamma_n(n)\mathcal{C}\Gamma_n(m)} \prod_{k=1}^{n} \lambda_k^{m-n} \prod_{k<l}^{n} (\lambda_k - \lambda_l)^2 \\ &\quad \times {}_0F_0^{(n)}(-\Sigma^{-1}, \Lambda), \end{aligned} \tag{13}$$

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

In the next section, we shall use these densities to evaluate the two most important information-theoretic measures, namely, the ergodic channel capacity and the capacity versus outage of MIMO Rayleigh distributed channels.

3 Multiple-Antenna Systems

In recent years, multiple-antenna techniques have become a pervasive idea that promises extremely high spectral efficiency for wireless communications. The basic information theory result reported in the pioneering papers by Foschini and Gans [2] and Telatar [12] showed that enormous spectral efficiency can be achieved through the use of multiple-antenna systems, provided that the complex-valued propagation coefficients between all pairs of transmitter and receiver antennas are statistically independent and known to the receiver antenna array. However, in practical systems the channel coefficients from two different transmitter antennas to a single receiver antenna can be correlated (at the transmitter end with covariance matrix Σ_t) and/or from a single transmitter antenna to two different receiver antennas can be correlated (at the receiver end with covariance matrix Σ_r). Such channel correlation, which degrades capacity (see Chuah et al. [1]), depends on the physical parameters of the MIMO system and the scatterer characteristics. The physical parameters include the antenna arrangement and spacing, the angle spread, the angle of arrival, etc. One of the objectives of this paper is to evaluate this capacity degradation for the correlated channel matrix $\mathrm{H} \sim \mathcal{C}N(0, \Sigma_r \otimes \Sigma_t)$ with $n_r > n_t$, $n_r = n_t$ and $n_t > n_r$.

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

$$y_j = \sum_{i=1}^{n_t} h_{ij} x_i + v_j, \tag{14}$$

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 with unit variance, as shown in Figure 1. The signal vector received at the output can be written as

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

i.e., in vector notation,

$$y = Hx + v, \tag{15}$$

where $y, v \in \mathbb{C}^{n_r}$, $H \in \mathbb{C}^{n_r \times n_t}$, $x \in \mathbb{C}^{n_t}$ and $v \sim \mathcal{C}N(0, I_{n_r})$. The total power of the input is constrained to ρ,

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

We shall deal exclusively with the linear model (15) and derive the capacity of MIMO channel models in this section.

3.1 Ergodic channel capacity

We assume that H is a complex Gaussian random matrix whose realization is known to the receiver, or equivalently, the channel output consists of the pair (y, H). Note that the transmitter does not know the channel and the input power is distributed equally over all transmitting antennas. Moreover, if we assume a block-fading model and coding over many independent fading intervals, then the Shannon or ergodic capacity of the random MIMO channel [12] is given by

$$\begin{aligned} C &= \mathcal{E}_{\mathrm{H}} \left\{ \log\det\left(I_{n_t} + (\rho/n_t)\mathrm{H}^H\mathrm{H} \right) \right\} \\ &= \mathcal{E}_{\mathrm{H}} \left\{ \log\det\left(I_{n_r} + (\rho/n_t)\mathrm{H}\mathrm{H}^H \right) \right\}, \end{aligned} \tag{16}$$

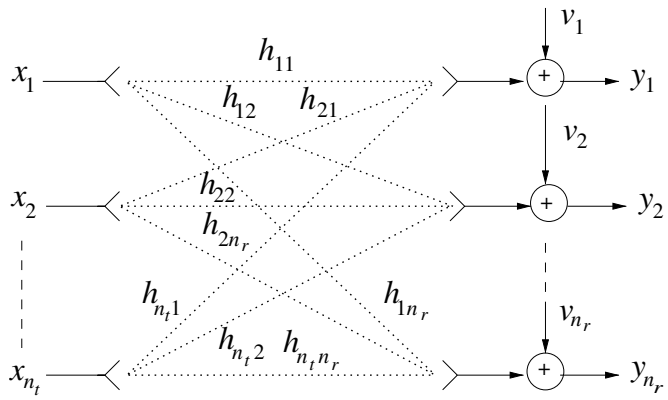

Figure 1: A MIMO communication system.

where the expectation is evaluated using a complex Gaussian density, i.e., $\mathrm{H} \sim \mathcal{C}N(0, \Sigma_r \otimes \Sigma_t)$. Let $\mathrm{S} = \mathrm{H}^H\mathrm{H}$. Then the channel capacity can be written as

$$C = \mathcal{E}_{\mathrm{S}} \left\{ \log\det\left(I_{n_t} + (\rho/n_t)\mathrm{S}\right) \right\}, \tag{17}$$

where the expectation is evaluated using the density given in (5). Note that if the channel is correlated only at the transmitter end and $n_t > n_r$ then $\mathrm{S} = \mathrm{H}^H\mathrm{H}$ is a complex singular Wishart matrix and its density is given in Theorem 3. Similarly, if the channel is correlated only at the receiver end and $n_r > n_t$ then $\mathrm{S} = \mathrm{H}\mathrm{H}^H$ is a complex singular Wishart matrix. For the other situations, S is full rank and the distributions are chosen according to the spatial correlation types, e.g., (5) – (7) (see Table 1 in [7]).

Let $\lambda_1 > \cdots > \lambda_{n_t}$ be the eigenvalues of S and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_{n_t})$. Then the capacity can also be computed using the joint eigenvalue density, $f(\Lambda)$, i.e.,

$$C = \mathcal{E}_{\Lambda} \left\{ \log\left(\prod_{k=1}^{n_t} \left[1 + \frac{\rho}{n_t}\lambda_k \right] \right) \right\}. \tag{18}$$

Again, the expectation in (18) is evaluated using the density according to the spatial correlation types, e.g., (9) – (11) and (13).

3.2 Capacity versus outage

Another useful information-theoretic measure is the 'x percent outage' which is defined to be the minimum mutual information that occurs in all but x percent of the channel instantiations. In other words, if we measure the mutual information of the channel many times — in many instantiations of the random channel — we would find that a mutual information greater than 5% outage would occur 95% of the time ($x = 5$).

Let us formulate the outage capacity. For a given instantiation of H, if the receiver knows the channel, the mutual information $\mathcal{I}(\mathrm{x}; (\mathrm{y}, \mathrm{H}))$ is given by

$$\begin{aligned} \mathcal{I}(\mathrm{x}; (\mathrm{y}, \mathrm{H})) &= \log\det\left[I_{n_t} + \frac{\rho}{n_t}\mathrm{H}^H\mathrm{H} \right] \\ &= \sum_{k=1}^{n_t} \log\left(1 + \frac{\rho}{n_t}\lambda_k \right), \end{aligned} \tag{19}$$

where $\lambda_1, \ldots, \lambda_{n_t}$ are the eigenvalues of $\mathrm{S} = \mathrm{H}^H\mathrm{H}$. The density $q(\mathcal{I})$ of the mutual information $\mathcal{I}$ in equation (19) over the ensemble of instantiations of H can then be written as

$$q(\mathcal{I}) = \mathcal{E}_{\Lambda} \left\{ \delta\left(\mathcal{I} - \sum_{k=1}^{n_t} \log\left[1 + \frac{\rho}{n_t}\lambda_k \right] \right) \right\}, \tag{20}$$

where δ is the Dirac delta function, see Simon and Moustakas [11]. The following theorem gives this density in terms of the density of a single unordered eigenvalue of $\mathrm{S} = \mathrm{H}^H\mathrm{H}$.

Theorem 5 *Let* H *be an* $n_r \times n_t$ *random matrix channel and* $g(\lambda)$ *be the density of a single unordered eigenvalue of* $\mathrm{S} = \mathrm{H}^H\mathrm{H}$. *Then the density of the mutual information,* $q(\mathcal{I})$*, is given by*

$$q(\mathcal{I}) = \left(\frac{e^{\mathcal{I}/n_t}}{\rho}\right) g\left(\frac{n_t}{\rho}\left[e^{\mathcal{I}/n_t} - 1\right]\right), \tag{21}$$

where ρ *is the total transmitter power.*

We can also express the distribution $Q(\mathcal{I})$ of the mutual information $\mathcal{I}$ as follows:

$$Q(\mathcal{I}) = \int_0^{\mathcal{I}} q(\mathcal{I})\, d\mathcal{I}. \tag{22}$$

The unique inverse function of $Q(\mathcal{I})$ is called the outage mutual information and is denoted by $\mathrm{out}(P_{\mathrm{out}})$ where $P_{\mathrm{out}} = Q(\mathcal{I}_{\mathrm{out}})$. In other words, we define $\mathrm{out}(P_{\mathrm{out}})$ such that

$$\mathcal{I}_{\mathrm{out}} = \mathrm{out}(P_{\mathrm{out}})$$

where

$$P_{\mathrm{out}} = Q(\mathcal{I}_{\mathrm{out}}).$$

Here $\mathcal{I}_{\mathrm{out}} = \mathrm{out}(P_{\mathrm{out}})$ means that, in any instantiation of H from the ensemble, we shall obtain a mutual information $\mathcal{I}$ less than $\mathcal{I}_{\mathrm{out}}$ with probability P_{out}. Using the density of the mutual information we can give another expression for the ergodic capacity as follows:

$$C = \int_0^{\infty} \mathcal{I} q(\mathcal{I})\, d\mathcal{I}. \tag{23}$$

In other words, the mean of the mutual information gives the ergodic capacity.

4 Computation of the capacities

In this section, we evaluate the capacities for both correlated and uncorrelated Rayleigh 2×2 fading channels ($n_r = n_t = 2$). First, we consider a channel with correlation at both transmitter and receiver ends. A typical example of this situation is a communication between two mobile laptops in poor scattering conditions. In addition, due to physical size constraints it is more difficult to space the antennas far apart at the laptops. These constraints lead to channel correlation at both transmitter and receiver ends, i.e., $\mathrm{H} \sim \mathcal{C}N(0, \Sigma_r \otimes \Sigma_t)$ and $\mathrm{S} = \mathrm{H}^H\mathrm{H} \sim \mathcal{C}Q_{n_r,n_t}(I_{n_r}, \Sigma_r, \Sigma_t)$. The ergodic capacity is given by

$$\begin{aligned} C_{cc} \;=\; & \frac{(\det \Sigma_t)^{-2}}{(\det \Sigma_r)^2} \int_0^{\infty}\int_0^{\lambda_1} d\lambda_2\, d\lambda_1 \left(\log\left[1 + \frac{\rho}{2}\lambda_1\right]\right. \\ & \left. + \log\left[1 + \frac{\rho}{2}\lambda_2\right]\right)(\lambda_1 - \lambda_2)^2 \\ & \times \sum_{k=0}^{\infty}\sum_{\kappa} \frac{\chi_{[\kappa]}(\Sigma_r^{-1})\,\chi_{[\kappa]}(-\Sigma_t^{-1})\,\chi_{[\kappa]}(\Lambda)}{(k_1 - k_2 + 1)\,\Gamma(k_1 + 2)\,\Gamma(k_2 + 1)}. \end{aligned}$$

Second, we consider a channel with correlation only at the transmitter end. A typical example of this situation is an uplink communication from a mobile unit to a base station. Here, the antennas at the base station can be spaced sufficiently far apart to achieve uncorrelation at the receiver end but, due to physical size constraints, it is more difficult to space the antennas far apart at the mobile unit (transmitter end), which leads to correlation at the transmitter end, i.e., $\mathrm{H} \sim \mathcal{C}N(0, I_{n_r} \otimes \Sigma_t)$ and $\mathrm{S} = \mathrm{H}^H\mathrm{H} \sim \mathcal{C}W_{n_t}(n_r, \Sigma_t)$ is a complex Wishart matrix. It should be noted that the joint eigenvalue density of a Wishart matrix depends on the population covariance matrix Σ_t only through its eigenvalues, $\upsilon_1, \ldots, \upsilon_{n_t}$, i.e.,

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

where $\Upsilon = \mathrm{diag}(\upsilon_1, \ldots, \upsilon_{n_t})$. Let $n_t = 2$ and $\Upsilon^{-1} = \mathrm{diag}(a_1, a_2)$. Then from [6] we have

$$\begin{aligned} {}_0F_0^{(2)}(-\Upsilon^{-1}, \Lambda) = & \frac{1}{(a_2 - a_1)(\lambda_1 - \lambda_2)} \\ & \times \left[e^{-a_1\lambda_1 - a_2\lambda_2} - e^{-a_1\lambda_2 - a_2\lambda_1}\right]. \end{aligned} \tag{24}$$

The ergodic capacity is given by

$$C_c = \frac{1}{(a_2 - a_1)} \int_0^\infty d\lambda \, \log\left[1 + \frac{\rho}{2}\lambda\right] \times \left[a_1^2 e^{-a_1\lambda}(a_2\lambda - 1) - a_2^2 e^{-a_2\lambda}(a_1\lambda - 1)\right].$$

Third, we consider an uncorrelated channel at the transmitter and receiver ends, i.e., $\mathrm{H} \sim \mathcal{C}N(0, I_{n_r} \otimes \sigma^2 I_{n_t})$ and $\mathrm{S} = \mathrm{H}^H\mathrm{H} \sim \mathcal{C}W_{n_t}(n_r, \sigma^2 I_{n_t})$. The ergodic capacity is given by

$$C_u = 2\int_0^\infty \log\left[1 + \frac{\rho}{2}\lambda\right] e^{-\lambda/\sigma^2} \left[\frac{\lambda^2}{2\sigma^6} - \frac{\lambda}{\sigma^4} + \frac{1}{\sigma^2}\right] d\lambda.$$

Figure 2 shows the capacity in nats[4] vs signal-to-noise ratio for correlated/uncorrelated Rayleigh 2×2 fading channel matrices. The following parameters are used: $\sigma^2 = 1$ and

$$\Sigma_r = \Sigma_t = \begin{bmatrix} 1 & 0.5 + 0.5i \\ 0.5 - 0.5i & 1 \end{bmatrix}.$$

From this figure we note the following: (i) the capacity is decreasing with channel correlation. Moreover, correlation at the transmitter and receiver ends causes more capacity degradation compare to correlation at the transmitter end only. (ii) the capacity is increasing with SNR.

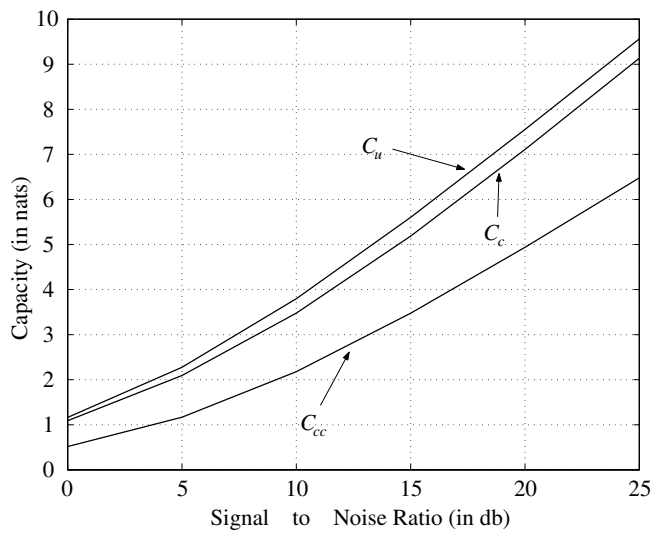

Figure 2: Capacity vs SNR for $n_t = 2$ and $n_r = 2$, i.e., Rayleigh 2×2 channel matrices. Note that C_u, C_c and C_{cc} denote the capacity of the three environments: uncorrelated, correlated at the transmitter end and correlated at both transmitter and receiver ends, respectively.

5 ACKNOWLEDGEMENT

Research partially supported by the Natural Sciences and Engineering Council of Canada and the Centre de recherches mathématiques of the Université de Montréal.

References

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

[2] J. G. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, Vol. 6, pp. 311–335, Mar. 1998.

[4] In equation (18), 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 ($e = 2.718\ldots$).

[3] A. T. James, "Distributions of matrix variate and latent roots derived from normal samples," *Ann. Math. Statist.*, Vol. 35, pp. 475–501, 1964.

[4] 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.

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

[6] T. Ratnarajah, R. Vaillancourt and M. Alvo, "Complex random matrices and Rayleigh channel capacity," *Communications in Information & Systems*, Vol. 3, No. 2, pp. 119–138, Oct. 2003 (http://www.ims.cuhk.edu.hk/~cis/).

[7] T. Ratnarajah and R. Vaillancourt, "Quadratic forms on complex random matrices and multiple-antenna systems," Submitted to *IEEE Trans. on Information Theory*, Jan. 2004.

[8] M. Sellathurai and G. Foschini, "A stratified diagonal layered space-time architecture: Information theoretic and signal processing aspects," *IEEE Trans. on Signal Processing*, Vol. 51, pp. 2943–2954, Nov. 2003.

[9] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering, and keyhole," *IEEE Trans. on Information Theory*, Vol. 49, pp. 2636–2647, Oct. 2003.

[10] D. S. Shiu, G. F. Foschini, M. G. Gans and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. on Communications*, Vol. 48, pp. 502–513, Mar. 2000.

[11] S. H. Simon and A. L. Moustakas, "Optimizing MIMO antenna systems with channel covariance feedback," *IEEE J. Select. Areas Commun.*, Vol. 31, pp. 406–417, Apr. 2003.

[12] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecom*, Vol. 10, pp. 585–595, 1999.