QUADRATIC FORMS ON COMPLEX RANDOM MATRICES AND CHANNEL CAPACITY

T. Ratnarajah and R. Vaillancourt

University of Ottawa
585 King Edward Ave.
Ottawa ON K1N 6N5 Canada
e-mail: t.ratnarajah@ieee.org

ABSTRACT

Quadratic forms on complex random matrices and their joint eigenvalue densities are derived with the goal of studying the *ergodic channel capacity* of multiple-input, multiple-output (MIMO) Rayleigh distributed wireless communication channels. We consider MIMO channels which are correlated at the transmitter and/or the receiver ends and evaluate the corresponding ergodic capacity formulas. These formulas are expressed in terms of complex zonal polynomials. This study shows how channel correlation degrades the capacity of the communication systems.

1. INTRODUCTION

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

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

Here, we study the distribution of $\mathbf{S}$, denoted by $\mathcal{C}Q_{m,n}(A, \Sigma_1, \Sigma_2)$, and its application to information theory. We also derive the joint eigenvalue density of $\mathbf{S}$, which is represented by complex zonal polynomials (also called Schur polynomials). Complex zonal polynomials are symmetric polynomials in the eigenvalues of a complex matrix [5], and they enable us to represent the derived densities as infinite series.

The theory of quadratic forms on complex random matrices is used to evaluate the capacity of MIMO wireless communication systems. Note that the capacity of a communication channel expresses the maximum rate at which information can be reliably conveyed by the channel [1]. Let us denote the number of inputs (or transmitters) and number of outputs (or receivers) of the MIMO wireless communications 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 $\mathbf{H} \sim \mathcal{C}N(\mathbf{0}, \Sigma_r \otimes \Sigma_t)$, where Σ_r and Σ_t represent the channel correlation at the receiver and transmitter ends, respectively. If $\Sigma_r = \sigma^2 I_{n_r}$ (or I_{n_r}) and $\Sigma_t = I_{n_t}$ (or $\sigma^2 I_{n_t}$) then the channel is 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 different ways, namely, ergodic capacity (or average mutual information) and capacity versus outage (or x percent outage). Note that the x percent outage is defined to be the minimum mutual information that occurs in all but x percent of the instantiations of the channel. In the sequel, we show that the ergodic capacity can be represented by quadratic forms on complex random matrices, $\mathbf{S} = \mathbf{X}\mathbf{X}^H$ (i.e., $A = I_n$), which can be computed by means of the eigenvalue density of $\mathbf{S}$. This is the motivation behind this study.

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 method are given in Sections 3 and 4, respectively.

2. QUADRATIC FORMS ON 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

$$\chi_{[\kappa]}(1) = k! \frac{\prod_{i<j}^{m}(k_i - k_j - i + j)}{\prod_{i=1}^{m}(k_i + m - i)!} \tag{3}$$

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

$$\chi_{[\kappa]}(X) = \frac{\det\left[\left(\lambda_i^{k_j+m-j}\right)\right]}{\det\left[\left(\lambda_i^{m-j}\right)\right]}. \tag{4}$$

Note that both the real and complex zonal polynomials are particular cases of the (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 of Jack polynomials, as was done in equation (2), i.e., $C_\kappa(X) := C_\kappa^{(1)}(X)$. Finally, we have

$$C_\kappa(I_n) = 2^{2k} k! \left[\frac{1}{2}n\right]_\kappa \frac{\prod_{i<j}^{r}(2k_i - 2k_j - i + j)}{\prod_{i=1}^{r}(2k_i + r - i)!}, \tag{5}$$

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 $(a)_k = a(a+1)\cdots(a+k-1)$. The next theorem gives the density of quadratic forms on complex random matrices, $\mathbf{S} = \mathbf{X}A\mathbf{X}^H$.

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

$$\begin{aligned} f(S) &= \frac{1}{\mathcal{C}\Gamma_m(n)(\det\Sigma_1)^n(\det\Sigma_2 A)^m}(\det S)^{n-m} \\ &\quad \times {}_0F_0^{(m)}(B, -\Sigma_1^{-1}S), \end{aligned} \tag{6}$$

where $A \in \mathbb{C}^{n\times n}$ is a positive definite Hermitian matrix and $B = A^{-1/2}\Sigma_2^{-1}A^{-1/2}$.

Proof. see [7]. □

The distribution of the matrix $\mathbf{S}$ is denoted by $\mathcal{C}Q_{m,n}(A, \Sigma_1, \Sigma_2)$. Special cases of density (6) are:

(i) If $A = I_n$, then the density of $\mathbf{S} = \mathbf{X}\mathbf{X}^H \sim \mathcal{C}Q_{m,n}(I_n, \Sigma_1, \Sigma_2)$ is given by

$$\begin{aligned} f(S) &= \frac{1}{\mathcal{C}\Gamma_m(n)(\det\Sigma_1)^n(\det\Sigma_2)^m}(\det S)^{n-m} \\ &\quad \times {}_0F_0^{(m)}(\Sigma_2^{-1}, -\Sigma_1^{-1}S). \end{aligned} \tag{7}$$

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

$$f(S) = \frac{1}{\mathcal{C}\Gamma_m(n)(\det\Sigma)^n}(\det S)^{n-m}\operatorname{etr}(-\Sigma^{-1}S), \tag{8}$$

where $\operatorname{etr}(\cdot) \equiv e^{\operatorname{tr}(\cdot)} \equiv \exp\operatorname{tr}(\cdot)$. The next theorem gives the joint eigenvalue density of quadratic forms on complex random matrices.

Theorem 2 *Consider the $m \times m$ positive definite Hermitian matrix $\mathbf{S} \sim \mathcal{C}Q(A, \Sigma_1, \Sigma_2)$, 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 $\mathbf{S}$ is*

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

where $B = A^{-1/2}\Sigma_2^{-1}A^{-1/2}$.

Proof. see [7]. □

Special cases of Theorem 2 are:

(i) If $A = I_n$, then the joint eigenvalue density of $\mathbf{S} \sim \mathcal{C}Q_{m,n}(I_n, \Sigma_1, \Sigma_2)$ is given by

$$\begin{aligned} f(\Lambda) &= \frac{\pi^{m(m-1)}(\det\Sigma_2)^{-m}}{\mathcal{C}\Gamma_m(n)\mathcal{C}\Gamma_m(m)(\det\Sigma_1)^n}\prod_{k=1}^{m}\lambda_k^{n-m}\prod_{k<l}^{m}(\lambda_k - \lambda_l)^2 \\ &\quad \times \sum_{k=0}^{\infty}\sum_{\kappa}\frac{C_\kappa(\Sigma_2^{-1})C_\kappa(-\Sigma_1^{-1})C_\kappa(\Lambda)}{k!C_\kappa(I_n)C_\kappa(I_m)}. \end{aligned} \tag{10}$$

(ii) If $A = I_n$, $\Sigma_1 = \Sigma$ and $\Sigma_2 = I_n$, then the joint eigenvalue density of the Wishart matrix $\mathbf{S}$ is given by

$$\begin{aligned} f(\Lambda) &= \frac{\pi^{m(m-1)}(\det\Sigma)^{-n}}{\mathcal{C}\Gamma_m(n)\mathcal{C}\Gamma_m(m)}\prod_{k=1}^{m}\lambda_k^{n-m}\prod_{k<l}^{m}(\lambda_k - \lambda_l)^2 \\ &\quad \times {}_0F_0^{(m)}(-\Sigma^{-1}, \Lambda). \end{aligned} \tag{11}$$

(iii) If $A = I_n$, $\Sigma_1 = \sigma^2 I_m$ and $\Sigma_2 = I_n$, then the joint eigenvalue density of the Wishart matrix $\mathbf{S}$ is given by

$$\begin{aligned} f(\Lambda) &= \frac{\pi^{m(m-1)}(\sigma^2)^{-nm}}{\mathcal{C}\Gamma_m(n)\mathcal{C}\Gamma_m(m)}\prod_{k=1}^{m}\lambda_k^{n-m}\prod_{k<l}^{m}(\lambda_k - \lambda_l)^2 \\ &\quad \times \exp\left(-\frac{1}{\sigma^2}\sum_{k=1}^{m}\lambda_k\right). \end{aligned} \tag{12}$$

In the next section, we use these densities to evaluate the ergodic channel capacity of MIMO Rayleigh distributed channels.

3. MIMO CHANNEL CAPACITY

Recently, in response to demand for higher bit rates in wireless communications, researchers have exploited the use of multiple-input, multiple-output (MIMO) systems, as shown in Figure 1 below. For example, if $n = \min\{n_t, n_r\}$, these studies show that MIMO uncorrelated Rayleigh distributed channel achieves almost n more bits per hertz for every 3-dB increase in signal-to-noise ratio (SNR) compared to a single-input, single-output (SISO) system, which achieves only one additional bit per hertz for every 3-dB increase in SNR, see Telatar [8]. But 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 [2], 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 $\mathbf{H} \sim \mathcal{C}N(\mathbf{0}, \Sigma_r \otimes \Sigma_t)$ with $n_t \geq n_r$. This will be done by deriving closed-form ergodic capacity formulas for correlated channels and their numerical evaluation.

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

$$y_j = \sum_{i=1}^{n_t} h_{ij} x_i + v_j, \tag{13}$$

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, 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,

$$y = Hx + v, \tag{14}$$

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,

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

We assume that the realization of $\mathbf{H}$ is known only to the

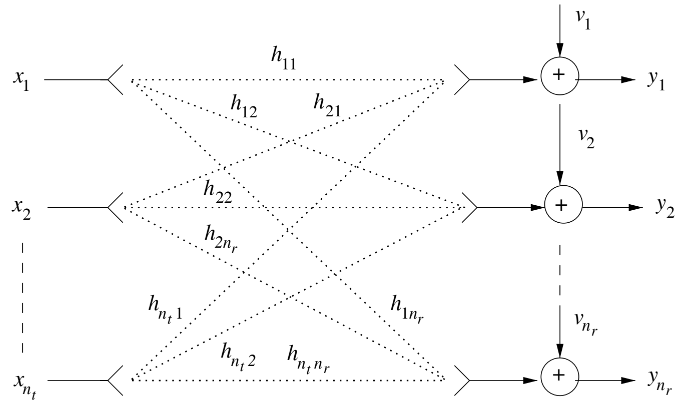

Fig. 1. MIMO communication system.

receiver but not to the transmitter and 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 is given in [8] by

$$\begin{aligned} C &= \mathcal{E}_{\mathbf{H}} \left\{ \log\det \left(I_{n_r} + \frac{\rho}{n_t} \mathbf{H}\mathbf{H}^H \right) \right\} \\ &= \mathcal{E}_{\mathbf{S}} \left\{ \log\det \left(I_{n_r} + \frac{\rho}{n_t} \mathbf{S} \right) \right\}, \end{aligned} \tag{15}$$

where the expectation is evaluated using the density (7), i.e., $\mathbf{S} = \mathbf{H}\mathbf{H}^H \sim \mathcal{C}Q_{n_r,n_t}(I_{n_t}, \Sigma_r, \Sigma_t)$. Let $\lambda_1 > \cdots > \lambda_{n_r}$ be the eigenvalues of $\mathbf{S}$ and $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_{n_r})$. Then the capacity can also be computed using the joint eigenvalue density $f(\Lambda)$ or the single unordered eigenvalue density $f(\lambda)$, i.e.,

$$\begin{aligned} C &= \sum_{k=1}^{n_r} \mathcal{E}_{\lambda_k} \left\{ \log(1 + \frac{\rho}{n_t} \lambda_k) \right\} \\ &= n_r \mathcal{E}_{\lambda} \left\{ \log(1 + (\rho/n_t)\lambda) \right\}. \end{aligned} \tag{16}$$

4. COMPUTATION OF THE CAPACITIES

In this section, we evaluate the capacities for both correlated and uncorrelated 2×2 Rayleigh fading channels (i.e., $n_t = n_r = 2$). First, we consider a channel with correlation at both transmitter and receiver ends, i.e., $\mathbf{H} \sim \mathcal{C}N(\mathbf{0}, \Sigma_r \otimes \Sigma_t)$ and $\mathbf{S} = \mathbf{H}\mathbf{H}^H \sim \mathcal{C}Q_{n_r,n_t}(I_{n_t}, \Sigma_r, \Sigma_t)$. The ergodic capacity is given by

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

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

Second, we consider a channel with only correlation at the receiver end, i.e., $\mathbf{H} \sim \mathcal{C}N(\mathbf{0}, \Sigma_r \otimes I_{n_t})$ and $\mathbf{S} = \mathbf{HH}^H \sim \mathcal{C}W_{n_r}(n_t, \Sigma_r)$ is a Wishart matrix. It should be noted that the joint eigenvalue density of a Wishart matrix depends on the population covariance matrix Σ_r only through its eigenvalues, $\upsilon_1, \ldots, \upsilon_{n_r}$, i.e.,

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

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

$$ {}_0F_0^{(2)}(-\Upsilon^{-1}, \Lambda) = \frac{1}{(a_2 - a_1)(\lambda_1 - \lambda_2)} \times \left[\exp\left\{-(a_1\lambda_1 + a_2\lambda_2)\right\} - \exp\left\{-(a_1\lambda_2 + a_2\lambda_1)\right\}\right]. \quad (17) $$

The ergodic capacity is given by

$$ C_c = \frac{1}{(a_2 - a_1)} \int_0^\infty \log\left[1 + \frac{\rho}{2}\lambda\right] \left[a_1^2 e^{-a_1\lambda}(a_2\lambda - 1) - a_2^2 e^{-a_2\lambda}(a_1\lambda - 1)\right] d\lambda. \quad (18) $$

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

$$ C_u = 2\int_0^\infty \log\left[1 + \frac{\rho}{2}\lambda\right] e^{-\lambda/\sigma^2} \left[\frac{\lambda^2}{2\sigma^6} - \frac{\lambda}{\sigma^4} + \frac{1}{\sigma^2}\right] d\lambda. \quad (19) $$

Figure 2 shows the capacity in nats vs signal-to-noise ratio for a 2×2 correlated/uncorrelated Rayleigh fading channel matrices. In the computation 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, (ii) the capacity is increasing with SNR.

5. CONCLUSION

In this paper, we study quadratic forms on complex random matrices and their application. In particular, we derive the densities of these forms and their joint eigenvalue densities. Using these densities, both correlated and uncorrelated MIMO Rayleigh channel ergodic capacity formulas are obtained. The capacity of 2×2 MIMO Rayleigh channel matrices are computed for both correlated and uncorrelated channels. It is shown how the channel correlation degrades the capacity of the communication systems.

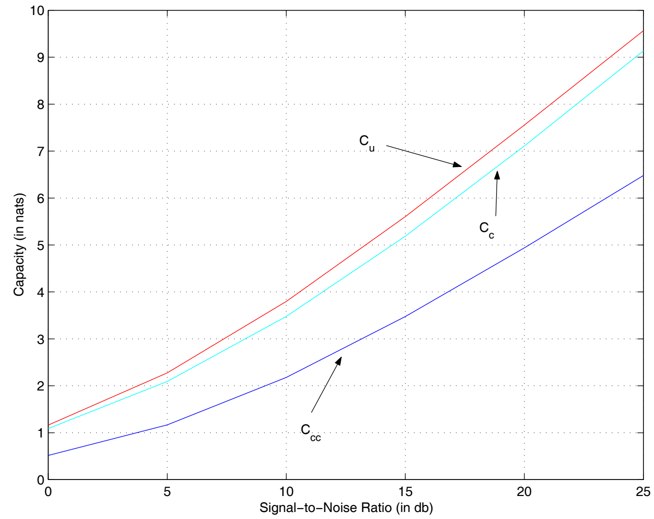

Fig. 2. Capacity vs SNR for $n_t = 2$ and $n_r = 2$, i.e., 2×2 Rayleigh channel matrix. Note that C_u, C_c and C_{cc} denote the capacity of uncorrelated, correlated at the receiver end and correlated at both transmitter and receiver ends, respectively.

6. REFERENCES

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

[2] 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, March 2002.

[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, pp. 119–138, Oct. 2003.

[7] T. Ratnarajah and R. Vaillancourt, "Quadratic forms on complex random matrices and information theory", *Submitted*

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