

An Upper Bound on the Ergodic Mutual Information of Ricean Fading MIMO Channels

Jari Salo

Helsinki University of Technology
SMARAD/Radio Laboratory
P.O. Box 3000, FI-02015-HUT
Finland
Email: jari.salo@hut.fi

Filip Mikas

Czech Technical University of Prague
Technická 2, Prague 6, CZ-16627
Czech Republic
Email: mikasf@feld.cvut.cz

Pertti Vainikainen

Helsinki University of Technology
SMARAD/Radio Laboratory
P.O. Box 3000, FI-02015-HUT
Finland
Email: pertti.vainikainen@hut.fi

Abstract— We consider Ricean fading Multiple-Input Multiple-Output (MIMO) channels where the transmitted signal has a complex Gaussian distribution, iid across the transmit antennas. Based on expected values of elementary symmetric functions of complex noncentral Wishart matrices, we derive a tight upper bound on the average (ergodic) mutual information for arbitrary SNR, arbitrary rank of the deterministic line-of-sight matrix, and arbitrary number of transmit/receive antennas. The Rayleigh fading signal component is allowed to have spatial correlation at one end of the link. Tight bounds for the cases of rank-1 line-of-sight component and pure Rayleigh fading emerge as special instances of the general result.

I. INTRODUCTION

We consider MIMO systems with n_t transmit and n_r receive antennas with Gaussian transmit signals, independent and identically distributed across the transmit antennas. Recently, a closed-form solution for the ergodic (average) mutual information in the case of semicorrelated Rayleigh fading channel statistics has been presented [1]. However, the average mutual information in the general case of Ricean fading channels is, to our knowledge, still an open issue. In this paper we derive a tight closed-form upper bound for the mutual information of Ricean fading MIMO channels whose random (non-line-of-sight) signal component is semicorrelated¹ Rayleigh fading. The line-of-sight component can have arbitrary rank. A tight bound for the case of Rayleigh fading emerges as a special case of our result. The bounds are simple to compute, for they are given as finite sums of certain functions of the eigenvalues of line-of-sight and spatial correlation matrices.

The paper is organized as follows. In Section II we provide some preliminaries. The main results are in Section III. Numerical examples demonstrating the tightness of the bound are given in Section IV. Section V concludes the paper. Some derivations can be found in the appendix.

II. SYSTEM MODEL

Assume that the distribution of the transmitted signal is complex zero-mean Gaussian with correlation matrix $\frac{P}{n_t} \mathbf{I}_{n_t}$, where P is the total transmitted signal power. Then, for a

given $n_r \times n_t$ channel matrix $\mathbf{H}$, the mutual information is given by

$$C_{\mathbf{H}} = \log_2 |\mathbf{I}_K + \mathbf{W}|, \quad (1)$$

where $K = \min(n_r, n_t)$, the $K \times K$ matrix

$$\mathbf{W} = \begin{cases} \frac{\rho}{n_t} \mathbf{H} \mathbf{H}^H, & \text{if } n_r \leq n_t \\ \frac{\rho}{n_t} \mathbf{H}^H \mathbf{H}, & \text{if } n_r > n_t, \end{cases}$$

and ρ is the average SNR at the output of each of the n_r receiving antennas [3]. The average mutual information $E[C_{\mathbf{H}}]$, which in some cases coincides with channel capacity, has been evaluated for Rayleigh fading channels in an integral form $\mathbf{H}$ [3] [4] and in closed-form [1]. In this paper we derive an upper bound for $E[C_{\mathbf{H}}]$ in the more general case of Ricean fading, which subsumes Rayleigh fading as a special case.

For Ricean fading a channel realization can be written $\mathbf{H} = a\mathbf{H}_d + b\mathbf{H}_s$, where

$$\begin{aligned} a &= \sqrt{\frac{K_r}{K_r + 1}} \\ b &= \sqrt{\frac{1}{K_r + 1}}. \end{aligned}$$

All matrices are $n_r \times n_t$ and $K_r = \frac{a^2}{b^2}$ is the Ricean K -factor. The line-of-sight component $\mathbf{H}_d$ is modelled as deterministic constant, and $\mathbf{H}_s$ is the Rayleigh fading stochastic component whose columns are iid from complex zero-mean Gaussian distribution with $n_r \times n_r$ correlation matrix Σ_s , i.e. $E[\mathbf{H}_s \mathbf{H}_s^H] = n_t \Sigma_s$. Following the convention from [2] we call this semicorrelated fading. We adopt the usual power normalization $E[\|\mathbf{H}\|_F^2] = n_r n_t$. Complex additive white Gaussian noise, iid in all dimensions, is assumed throughout the paper.

Note that, for the Ricean channel, the average of mutual information, $E[C_{\mathbf{H}}]$, does not, in general, coincide with the ergodic channel capacity.

¹By “semicorrelated” we mean that correlation is allowed only at one end of the link [2].

III. THE MAIN RESULTS

A. The basic idea of derivation

The derivation of the upper bound is essentially based on two key components: Jensen's inequality and expected values of elementary symmetric functions of a complex noncentral Wishart matrix. This becomes apparent from the following development:

$$\begin{aligned} E[C_{\mathbf{H}}] &= E[\log_2 |\mathbf{I}_K + \mathbf{W}|] \\ &\leq \log_2 \left[E \left[\prod_{p=1}^K (1 + \lambda_p) \right] \right] \\ &= \log_2 \left[\sum_{p=0}^K E[\text{tr}_p(\mathbf{W})] \right], \end{aligned} \quad (2)$$

where Jensen's inequality and the concavity of $\log_2(\cdot)$ was used in the second line, and $\text{tr}_p(\mathbf{W})$, $p = 0, \dots, K$, is the p th elementary symmetric function² of $\mathbf{W}$ [6]. The function $\text{tr}_p(\mathbf{W})$ depends only on the eigenvalues of $\mathbf{W}$, which are denoted with $\lambda_1, \dots, \lambda_K$. For instance, $\text{tr}_0(\mathbf{W}) = 1$, $\text{tr}_1(\mathbf{W}) = \sum_{p=1}^K \lambda_p = \text{tr}(\mathbf{W})$ and $\text{tr}_K \mathbf{W} = |\mathbf{W}|$. In general, $\text{tr}_p(\mathbf{W}) = \sum \lambda_{i_1} \cdots \lambda_{i_p}$, where the sum is over all $\binom{K}{p}$ combinations of the p indices with $i_1 < \dots < i_p$. Interestingly, the expected values of elementary symmetric functions are known for many different statistics of $\mathbf{W}$ [7]. However, for *complex-valued* distributions results are considerably more scarce. In the case of Ricean fading the distribution of $\mathbf{W}$ is *complex noncentral Wishart* [8]. Therefore, in order to compute the bound (2), we need to evaluate the expected values of the elementary symmetric functions of $\mathbf{W}$ over this distribution. Deferring the details of the derivation to appendix we now jump to the main result of the paper.

B. The upper bound in the general case

Let $L = \max(n_r, n_t)$. Denote $\mathbf{T} = \mathbf{H}_d \mathbf{H}_d^H$, and with $\mathbf{A}^{i,p}$ the i th $p \times p$ principal submatrix [6] of the $K \times K$ matrix $\mathbf{A}$, $i = 1, \dots, \binom{K}{p}$. The Pochhammer symbol is denoted $(a)_p = a(a+1) \cdots (a+p-1)$. We have

Result 1: Let be $\mathbf{H}$ be Ricean fading. The average mutual information $E[C_{\mathbf{H}}]$ can be upper bounded as

$$\begin{aligned} E[C_{\mathbf{H}}] &\leq \log_2 \left[1 + \sum_{p=1}^K \left(\frac{\rho b^2}{n_t} \right)^p (L-p+1)_p \text{tr}_p(\Sigma_s) \right. \\ &\quad + \sum_{p=1}^K \left(\frac{\rho b^2}{n_t} \right)^p \sum_{j=1}^p K_r^j (L-p+1)_{(p-j)} \\ &\quad \times \sum_{i=1}^{\binom{K}{p}} |\Sigma_s^{i,p}| \text{tr}_j[(\Sigma_s^{i,p})^{-1} \mathbf{T}^{i,p}] \left. \right]. \end{aligned} \quad (3)$$

Proof: See appendix. ■

In Section IV we compare the upper bound to exact mutual information and show it to be surprisingly tight.

²We remark that the generating function of the elementary symmetric functions of $\mathbf{W}$ is $\sum_{p=0}^K \text{tr}_p(\mathbf{W}) t^p = \prod_{p=1}^K (1 + \lambda_p t)$ [5]. This coincides with $|\mathbf{I}_K + t\mathbf{W}|$ and provides an alternative view on the capacity determinant.

C. The upper bound in special cases

1) *Rayleigh iid $\mathbf{H}_s$:* In this case $\Sigma_s = \mathbf{I}_K$ and Result 1 simplifies to the following form.

Result 2: Let be $\mathbf{H}$ be Ricean fading with $\Sigma_s = \mathbf{I}_K$. Then Result 1 becomes

$$\begin{aligned} E[C_{\mathbf{H}}] &\leq \log_2 \left[\sum_{p=0}^K \left(\frac{\rho b^2}{n_t} \right)^p \sum_{j=0}^p K_r^j (L-p+1)_{(p-j)} \right. \\ &\quad \times \binom{K-j}{p-j} \text{tr}_j(\mathbf{T}) \left. \right]. \end{aligned} \quad (4)$$

Proof: See appendix. ■

2) *Rayleigh iid $\mathbf{H}_s$ and rank-1 $\mathbf{H}_d$:* In practice, the line-of-sight component $\mathbf{H}_d$ often has rank one. In this case we can further simplify Result 2 to yield

Result 3: Let be $\mathbf{H}$ be Ricean fading with $\Sigma_s = \mathbf{I}_K$ and let $\text{rank}(\mathbf{H}_d) = 1$. Then the upper bound (3) reduces to

$$\begin{aligned} E[C_{\mathbf{H}}] &\leq \log_2 \left[1 + \sum_{p=1}^K \sum_{j=0}^1 \left(\frac{\rho b^2}{n_t} \right)^p (K_r K L)^j \right. \\ &\quad \times (L-p+1)_{(p-j)} \binom{K-j}{p-j} \left. \right]. \end{aligned} \quad (5)$$

Proof: See appendix. ■

Note that in this case the average mutual information does not depend on the line-of-sight matrix $\mathbf{H}_d$, since, due to the power normalization, the only nonzero eigenvalue of $\mathbf{T}$ equals KL ; hence $\text{tr}_1(\mathbf{T}) = KL$ and $\text{tr}_j(\mathbf{T}) = 0$ for $j > 1$.

3) *Semicorrelated Rayleigh fading:* This special case results by setting $K_r = 0$ in Result 1.

Result 4: Let $\mathbf{H}$ be Rayleigh fading with correlation matrix Σ_s . Result 1 becomes

$$E[C_{\mathbf{H}}] \leq \log_2 \left[\sum_{p=0}^K \left(\frac{\rho}{n_t} \right)^p (L-p+1)_p \text{tr}_p(\Sigma_s) \right].$$

The case of Rayleigh iid fading results by replacing $\text{tr}_p(\Sigma_s)$ with $\binom{K}{p}$ (since all eigenvalues are one). This bound is particularly convenient (and accurate enough for most practical purposes) for quick capacity estimations.

IV. NUMERICAL EXAMPLES

In both examples we examine a 4×4 MIMO system with rank-1 line-of-sight matrix. Hence, the results in this section hold for arbitrary rank-1 LOS matrix, and we do not need to explicitly define $\mathbf{H}_d$. The exact values for the mutual information have been obtained from Monte Carlo simulation with 10^5 realizations.

A. Effect of SNR and K -factor

In Fig. 1 mutual information has been plotted for different K -factors, denoted with K_r . The stochastic component is Rayleigh iid, i.e. $\Sigma_s = \mathbf{I}_4$. We note that the bound is extremely tight at low SNR. For example, with $K_r = 10$ dB and $\rho = 10$ dB, the error is about 1% (less than one tenth of a bits/s/Hz).

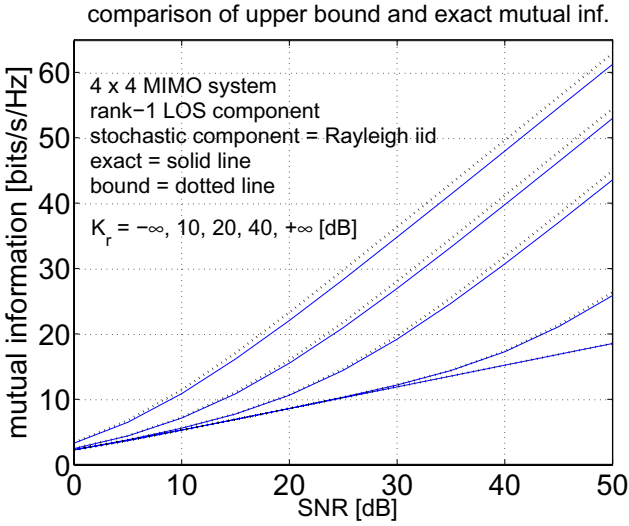


Fig. 1. Exact (empirical) capacity and upper bound for a 4×4 system for different K -factors.

In general, the bound gets tighter as K_r increases and SNR decreases. Even the moderately high K -factor of $K_r = 10$ dB does not dramatically affect mutual information compared to the pure Rayleigh fading case ($K_r = -\infty$). Note that there is slight change in the rate of increase of mutual information at $\rho \approx K_r$; this is where the channel eigenvalues due to the Rayleigh iid component become large enough to have an effect on the slope of the capacity curve.

B. Effect of correlation and K -factor

We consider a simple correlation model where the entries of Σ_s are $\sigma_{ij} = r^{|i-j|}$ with $r \in [0, 1)$. In Fig. 2 the effect of correlation is plotted for varying K_r and $\rho = 30$ dB. The result confirms the intuitively clear fact that as the K -factor increases the effect of correlation becomes less significant; this effect is even more pronounced at low SNR. We have deliberately chosen a fairly high SNR to better illustrate the effect of correlation. The bound is tighter at low SNR.

V. CONCLUSION

We have derived a computationally simple upper bound for ergodic mutual information in Ricean fading MIMO channels. The bound was shown to be tight over all MIMO regimes. The case of Rayleigh fading is a special case of the general result.

APPENDIX I BASIC DEFINITIONS

A. Notation

$$\begin{aligned} (a)_p &= a(a+1)\dots(a+p-1), \quad (a)_0 = 1 \\ \text{etr}(\mathbf{X}) &= \exp[\text{tr}(\mathbf{X})] \\ \Gamma(t) &= \int_0^\infty x^{t-1} e^{-x} dx \\ \Gamma_K(L) &= \pi^{K(K-1)/2} \prod_{p=1}^K \Gamma(L-p+1) \end{aligned}$$

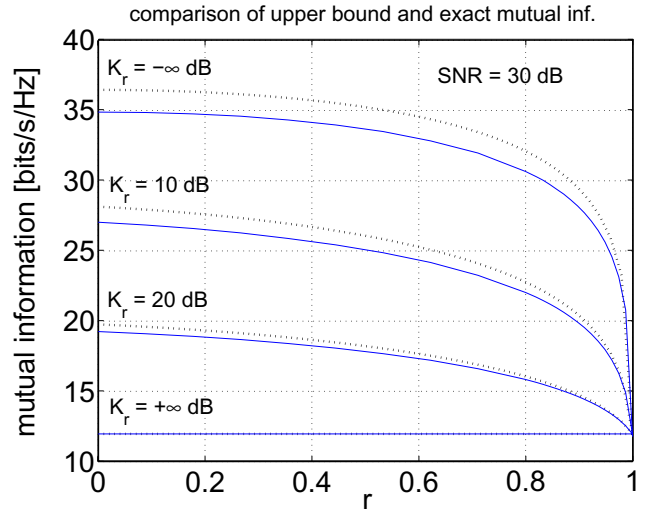


Fig. 2. Exact (empirical) capacity and upper bound for a 4×4 system for different values of correlation coefficient r , $\text{rank}(\mathbf{H}_d) = 1$. Bound = dotted line, exact = solid line.

B. Complex matrix variate Gaussian distribution

The $K \times L$ matrix $\mathbf{X}$ has *complex matrix variate Gaussian* distribution if its pdf is given by [8, Eq. (3.4)]

$$f(\mathbf{X}) = \frac{1}{\pi^{KL} |\Sigma|^L} \text{etr}[-\Sigma^{-1}(\mathbf{X} - \mathbf{M})(\mathbf{X} - \mathbf{M})^H],$$

where $E[\mathbf{X}] = \mathbf{M}$ and $E[(\mathbf{X} - \mathbf{M})(\mathbf{X} - \mathbf{M})^H] = L\Sigma$. This is the distribution of Ricean fading channel matrix, the line-of-sight component being $\mathbf{M}$.

C. Complex noncentral Wishart distribution

Let the $K \times L$ matrix $\mathbf{X}$ (with $L \geq K$) be complex matrix variate Gaussian. Then the $K \times K$ matrix $\mathbf{S} = \mathbf{X}\mathbf{X}^H$ has the *complex noncentral Wishart* distribution with parameters $(L, \Sigma, \mathbf{M})$, and its pdf is [8, Eq. (3.6)]

$$f(\mathbf{S}) = \frac{1}{\Gamma_K(L) |\Sigma|^L} |\mathbf{S}|^{L-K} \text{etr}[-\Sigma^{-1}\mathbf{S}] \quad (6)$$

$$\times \text{etr}[-\Theta] {}_0F_1(L; \Theta \Sigma^{-1} \mathbf{S}), \quad (7)$$

where $\Theta = \Sigma^{-1} \mathbf{M} \mathbf{M}^H$ is the noncentrality parameter and ${}_0F_1(b; \mathbf{A})$ is a hypergeometric function of complex matrix argument [9].

D. Complex zonal polynomials

Let $\kappa = (k_1, k_2, \dots, k_Q)$ be a Q -partition³ of the positive integer p such that $k_1 \geq k_2 \geq \dots \geq k_Q \geq 0$ and $\sum_{i=1}^Q k_i = p$. The *generalized complex hypergeometric coefficient* associated with the Q -partition κ is defined as [8]

$$(a)_\kappa = \prod_{i=1}^Q (a - i + 1)_{k_i}. \quad (8)$$

³For instance, the 2-partitions of $p = 2$ are $\kappa = (2, 0)$ and $\kappa = (1, 1)$.

The *complex zonal polynomial* of a $K \times K$ hermitian positive definite matrix $\mathbf{A}$ can be defined as [4] [9]

$$C_\kappa(\mathbf{A}) = \chi_\kappa(1) \chi_\kappa(\mathbf{A}), \quad (9)$$

where the scaling constant

$$\chi_\kappa(1) = \frac{p! \left[\prod_{m < n}^K (k_m - k_n - m + n) \right]}{\prod_{m=1}^K (k_m + K - m)!},$$

and the polynomial in the eigenvalues of $\mathbf{A}$ is

$$\chi_\kappa(\mathbf{A}) = \frac{\begin{vmatrix} \lambda_1^{k_1+K-1} & \lambda_1^{k_2+K-2} & \dots & \lambda_1^{k_K} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_K^{k_1+K-1} & \lambda_K^{k_2+K-2} & \dots & \lambda_K^{k_K} \end{vmatrix}}{\begin{vmatrix} \lambda_1^{K-1} & \lambda_1^{K-2} & \dots & \lambda_1^0 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_K^{K-1} & \lambda_K^{K-2} & \dots & \lambda_K^0 \end{vmatrix}}.$$

APPENDIX II

EXPECTED VALUES OF ELEMENTARY SYMMETRIC FUNCTIONS OF A COMPLEX NONCENTRAL WISHART MATRIX

In this appendix we outline the derivation of $E[\text{tr}_p(\mathbf{W})]$ required in Result 1. For the real-valued noncentral Wishart the results have been proven in [10]. We were unable to find in literature the corresponding proofs for the complex-valued case, so we sketch the main steps here.

Let $\mathbf{S}$ be complex noncentral Wishart matrix. The expected value of its determinant is

$$\begin{aligned} E[|\mathbf{S}|] &= \int_{\mathbf{S} > 0} |\mathbf{S}| f(\mathbf{S}) d\mathbf{S} \\ &= \frac{\text{etr}[-\boldsymbol{\Theta}]}{\Gamma_K(L) |\boldsymbol{\Sigma}|^L} \int_{\mathbf{S} > 0} |\mathbf{S}|^{L-K+1} \text{etr}[-\boldsymbol{\Sigma}^{-1}\mathbf{S}] \\ &\quad \times {}_0F_1(L; \boldsymbol{\Theta} \boldsymbol{\Sigma}^{-1} \mathbf{S}) d\mathbf{S} \\ &= \frac{\text{etr}[-\boldsymbol{\Theta}] |\boldsymbol{\Sigma}| \Gamma_K(L+1)}{\Gamma_K(L)} {}_1F_1(L+1; L; \boldsymbol{\Theta}), \end{aligned}$$

where ${}_0F_1(b; \mathbf{A})$ and ${}_1F_1(a; b; \mathbf{A})$ are hypergeometric functions of complex matrix argument [9] and we also used the integral [9, p. 369]. From the Kummer relation [11] ${}_1F_1(a_1; b_1; \mathbf{A}) = \text{etr}(\mathbf{A}) {}_1F_1(b_1 - a_1; b_1; -\mathbf{A})$ and $\frac{\Gamma_K(L+1)}{\Gamma_K(L)} = (L - K + 1)_K$ we can write $E[|\mathbf{S}|] = (L - K + 1)_K |\boldsymbol{\Sigma}| {}_1F_1(-1; L; -\boldsymbol{\Theta})$. By definition [9]

$${}_1F_1(-1; L; -\boldsymbol{\Theta}) = \sum_{p=0}^{\infty} \sum_{\text{all } \kappa} \frac{(-1)_\kappa}{(L)_\kappa} \frac{C_\kappa(-\boldsymbol{\Theta})}{p!}. \quad (10)$$

The key observation is that $(-1)_p = 0$ for $p > 1$ and hence we can restrict to partitions $\kappa_1 = (1, 1, \dots, 1)$ with p ones. From (8) it follows that $(-1)_{\kappa_1} = (-1)^p p!$ and $(L)_{\kappa_1} = (L - p + 1)_p$. Hence

$${}_1F_1(-1; L; -\boldsymbol{\Theta}) = \sum_{p=0}^K \frac{(-1)^p}{(L - p + 1)_p} C_{\kappa_1}(-\boldsymbol{\Theta}). \quad (11)$$

From (9) it can be shown that the zonal polynomial $C_{\kappa_1}(-\boldsymbol{\Theta}) = (-1)^p \text{tr}_p(\boldsymbol{\Theta})$, i.e. the p th elementary symmetric function of $\boldsymbol{\Theta}$ (here, due to space limitations, we omit some details). With $\frac{(L-K+1)_K}{(L-p+1)_p} = (L - K + 1)_{(K-p)}$ we have

$$E[|\mathbf{S}|] = |\boldsymbol{\Sigma}| \sum_{p=0}^K (L - K + 1)_{(K-p)} \text{tr}_p(\boldsymbol{\Theta}). \quad (12)$$

The p th elementary symmetric function, $(1 \leq p \leq K)$, can be written [6]

$$\text{tr}_p(\mathbf{A}) = \sum_{\text{all } i} [\text{ith } p \times p \text{ principal minor of } \mathbf{A}].$$

Denote with $\mathbf{A}^{i,p}$ the i th $p \times p$ principal submatrix of $\mathbf{A}$. There are $\binom{K}{p}$ such principal submatrices. The submatrix $\mathbf{S}^{i,p}$ is distributed as complex noncentral Wishart with parameters $\{L, \boldsymbol{\Sigma}^{i,p}, (\boldsymbol{\Sigma}^{i,p})^{-1} \mathbf{T}^{i,p}\}$, where $\mathbf{T} = \mathbf{M} \mathbf{M}^H$. We can use (12) and write

$$\begin{aligned} E[\text{tr}_p(\mathbf{S})] &= (L - p + 1)_p \text{tr}_p(\boldsymbol{\Sigma}) + \sum_{j=1}^p (L - p + 1)_{(p-j)} \\ &\quad \times \sum_{i=1}^{\binom{K}{p}} |\boldsymbol{\Sigma}^{i,p}| \text{tr}_j[(\boldsymbol{\Sigma}^{i,p})^{-1} \mathbf{T}^{i,p}]. \end{aligned} \quad (13)$$

Result 1 follows from (13). The special cases (Results 2–4) follow straightforwardly from (13), but, due to space limitations, we omit details.

ACKNOWLEDGMENT

The first author would like to thank the Foundation of Commercial and Technical Sciences and the Nokia foundation for financial support.

REFERENCES

- [1] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels: spatial fading correlation, double scattering, and keyhole," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2636–2647, Oct. 2003.
- [2] P. J. Smith, S. Roy, and M. Shafi, "Capacity of MIMO systems with semicorrelated flat fading," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2781–2788, Oct. 2003.
- [3] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–596, November 1999.
- [4] T. Ratnarajah, R. Vaillancourt, and M. Alvo, "Complex random matrices and Rayleigh channel capacity," *Communications in Information and Systems*, vol. 3, no. 2, pp. 119–138, Oct. 2003.
- [5] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, 2nd ed. Oxford University Press, 1995.
- [6] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.
- [7] A. Gupta and D. Nagar, *Matrix Variate Distributions*. Chapman and Hall/CRC, 2000.
- [8] P. Krishnaiah, "Some recent developments on complex multivariate distributions," *Journal of Multivariate Analysis*, vol. 6, pp. 1–30, 1976.
- [9] A. Mathai, *Jacobians of Matrix Transformations and Functions of Matrix Argument*. World Scientific Publishing, 1997.
- [10] B. Shah and C. Khatri, "Proof of conjectures about the expected values of the elementary symmetric functions of a noncentral Wishart matrix," *The Annals of Statistics*, vol. 2, no. 4, pp. 833–836, Jul. 1974.
- [11] A. Constantine, "Some non-central distribution problems in multivariate analysis," *The Annals of Mathematical Statistics*, vol. 34, no. 4, pp. 1270–1285, Dec. 1964.