

New Properties of Complex Noncentral Quadratic Forms and Bounds on MIMO Mutual Information

Matthew R. McKay

Telecommunications Lab.
School of Elec. and Info. Engineering
University of Sydney, Australia
and also the ICT Centre, CSIRO, Australia
mckay@ee.usyd.edu.au

Peter J. Smith

Dept. of Elec. and Comp. Engineering
University of Canterbury
Christchurch, New Zealand
p.smith@elec.canterbury.ac.nz

Iain B. Collings

Wireless Technologies Lab.
ICT Centre
CSIRO
Sydney, Australia
Iain.Collings@csiro.au

Abstract—This paper presents new statistical properties of complex noncentral matrix-variate quadratic forms. In contrast to previous results, the expressions do not involve infinite sums over partitions, or matrix-variate polynomials, and are easily and efficiently computable. These properties are used to derive new upper and lower bounds on the ergodic mutual information of double-sided correlated Rician MIMO channels with arbitrary-rank channel mean matrices. The bounds are shown to be tighter than previous reported bounds in the literature.

I. INTRODUCTION

Multiple-input multiple-output (MIMO) antenna systems have received considerable research attention since they were initially shown to provide significant channel capacity improvements over single-antenna systems in uncorrelated Rayleigh fading environments [1]. For these channels, the capacity was shown to be equivalent to the ergodic mutual information (MI) with isotropic Gaussian input signalling. Motivated by the results of [1], many bounds and exact ergodic MI expressions (assuming isotropic inputs) have been derived for more practical, spatially-correlated Rayleigh and uncorrelated Rician MIMO channels (e.g. see [2–6] and references therein).

Few analytical MIMO MI results are available for spatially-correlated Rician channels. In [7, 8], upper and lower bounds were presented for *single-sided* correlated Rician channels (i.e. correlation at either the transmitter or receiver, but not both), with rank-1 mean matrices. In [9, 10] these results were extended to mean matrices of arbitrary rank. All of these previous results further restricted the single-sided correlation matrix to occur only at the end of the transmission link with the least number of antennas.

For *double-sided* correlated Rician channels, tight upper and lower bounds on ergodic MI were derived in [9], however those results involved infinite series over partitions of numbers and matrix-variate Hayakawa polynomials, and were not suited to efficient numerical evaluation. A more efficient, but much looser, upper bound was also presented in [9]. Computationally efficient bounds were presented for the special case of rank-1 mean matrices in [11].

The main difficulty in deriving tight analytic bounds on ergodic MI of Rician MIMO channels with double-sided correlation, is that statistical properties of complex noncentral matrix-variate quadratic forms are required (rather than simpler noncentral Wishart matrices, which arise in single-sided correlated scenarios). For these random matrices, most known statistical properties involve infinite series and matrix-variate polynomials (see, for example, [12] and [9, Sect. II], as well as [13, Chapt. 7] for real matrices), and cannot be easily or efficiently computed.

In this paper we derive several new statistical properties of complex noncentral matrix-variate quadratic forms. In contrast to the existing properties in [9, 12, 13], the new results in this paper are finite closed-form expressions which do not involve infinite series or the evaluation of matrix-variate polynomials.

Based on these general statistical results, we then obtain new upper and lower bounds on the ergodic MI of double-sided correlated Rician channels with arbitrary-rank mean matrices (DSC-ARM). Our numerical results show that the upper bound is significantly tighter than previous reported bounds in [9] and [11]. Our lower bound appears to be the only computable DSC-ARM ergodic MI bound in the communications literature, and is shown via simulations to be tight.

II. NOTATION AND DEFINITIONS

A. Notation

Matrices are represented by uppercase boldface symbols, and vectors by lowercase boldface. The superscript $(\cdot)^\dagger$ indicates complex conjugate transpose, and the matrix \mathbf{I}_p denotes a $p \times p$ identity. The trace operation is denoted $\text{tr}(\cdot)$, and $\text{etr}(\cdot)$ is shorthand notation for $\exp(\text{tr}(\cdot))$. The Kronecker product is \otimes , $\text{vec}(\mathbf{A})$ is the operator which stacks the columns of \mathbf{A} into a single vector, and $\mathbf{A} > 0$ denotes positive definiteness. We use $\mathbf{A}_{\mathcal{F}}^{\mathcal{G}}$ or $(\mathbf{A})_{\mathcal{G}}^{\mathcal{F}}$ to denote submatrices of the $p \times q$ matrix \mathbf{A} , formed by taking only the rows indexed by $\mathcal{F} \subseteq \{1, 2, \dots, p\}$ and columns indexed by $\mathcal{G} \subseteq \{1, 2, \dots, q\}$, and $\{\alpha_{n,m}\}$ to denote the set of all $\binom{m}{n}$ ordered length- n subsets of the numbers $\{1, \dots, m\}$. Finally, all logarithms are taken to the base-2 unless otherwise specified.

B. Multivariate Statistics Definitions

Denote $\mathcal{CN}_p(\mu, \Phi)$ to be the p -variate complex Gaussian distribution with mean vector $\mu \in \mathcal{C}^{p \times 1}$ and covariance matrix $\Phi \in \mathcal{C}^{p \times p} > 0$.

Definition 1 [13]: The random matrix $\mathbf{X} \in \mathcal{C}^{p \times q}$ is said to have a matrix-variate complex Gaussian distribution with mean matrix $\Upsilon \in \mathcal{C}^{p \times q}$ and covariance matrix $\Phi \otimes \Psi$, where $\Phi \in \mathcal{C}^{p \times p} > 0$ and $\Psi \in \mathcal{C}^{q \times q} > 0$, if

$$\text{vec}(\mathbf{X}^\dagger) \sim \mathcal{CN}_{pq}(\text{vec}(\Upsilon^\dagger), \Phi \otimes \Psi). \quad (1)$$

For matrices with matrix-variate complex Gaussian distributions we use the notation $\mathbf{X} \sim \mathcal{CN}_{p,q}(\Upsilon, \Phi \otimes \Psi)$, which has probability density function (p.d.f.)

$$f_G(\mathbf{X}) = \pi^{-pq} \det(\Phi)^{-q} \det(\Psi)^{-p} \times \text{etr}\left(-\Phi^{-1}(\mathbf{X} - \Upsilon)\Psi^{-1}(\mathbf{X} - \Upsilon)^\dagger\right). \quad (2)$$

$$f_{\mathbf{Q}}(\mathbf{Q}) = \frac{\text{etr}(-\Phi^{-1}\Upsilon\Psi^{-1}\Upsilon^\dagger) \text{etr}(-\nu\Phi^{-1}\mathbf{Q}) \det(\mathbf{Q})^{q-p}}{\tilde{\Gamma}_p(q) \det(\Phi)^q \det(\tilde{\mathbf{A}})^p} \sum_{k=0}^{\infty} \sum_{\mathcal{K}} \frac{\tilde{P}_{\mathcal{K}}\left(\Phi^{-\frac{1}{2}}\Upsilon\Psi^{-\frac{1}{2}}(\mathbf{I}_q - \nu\tilde{\mathbf{A}})^{-\frac{1}{2}}, \tilde{\mathbf{A}}^{-1} - \nu\mathbf{I}_q, \Phi^{-\frac{1}{2}}\mathbf{Q}\Phi^{-\frac{1}{2}}\right)}{k!(q)_{\mathcal{K}}} \quad (5)$$

Definition II: Let $\mathbf{X} \sim \mathcal{CN}_{p,q}(\Upsilon, \Phi \otimes \mathbf{I}_q)$, with $p \leq q$. Then $\mathbf{W} = \mathbf{X}\mathbf{X}^\dagger$ has a complex noncentral Wishart distribution $\mathcal{W}_p(q, \Phi, \Theta)$ with p.d.f. [14]

$$f_W(\mathbf{W}) = f_W^c(\mathbf{W}) \text{etr}(-\Theta) {}_0\tilde{F}_1(q; \Theta \Phi^{-1} \mathbf{W}) \quad (3)$$

where $f_W^c(\mathbf{W})$ is the complex central Wishart p.d.f.

$$f_W^c(\mathbf{W}) = \frac{\text{etr}(-\Phi^{-1}\mathbf{W}) \det(\mathbf{W})^{q-p}}{\tilde{\Gamma}_p(q) \det(\Phi)^q} \quad (4)$$

and where $\Theta = \Phi^{-1}\Upsilon\Upsilon^\dagger$ is the *non-centrality* parameter. Also, ${}_0\tilde{F}_1(\cdot)$ is the complex Bessel hypergeometric function (of a matrix argument) [14] and $\tilde{\Gamma}_p(q) = \pi^{\frac{p(p-1)}{2}} \prod_{j=1}^p (q-j)!$ is the complex multivariate gamma function.

Definition III: Let $\mathbf{X} \sim \mathcal{CN}_{p,q}(\Upsilon, \Phi \otimes \Psi)$, with $p \leq q$, and $\mathbf{A} \in \mathcal{C}^{q \times q} > 0$. Then $\mathbf{Q} = \mathbf{X}\mathbf{A}\mathbf{X}^\dagger$ is a noncentral matrix-variate complex quadratic form. The distribution is denoted $\mathcal{Q}_{p,q}(\mathbf{A}, \Phi, \Psi, \Upsilon)$.

The p.d.f. is given by (5) [12], where $\nu \geq 0$ is an arbitrary constant, $\tilde{\mathbf{A}} = \Psi^{1/2}\mathbf{A}\Psi^{1/2}$, and $\mathcal{K} = (k_1, \dots, k_p)$ is a partition of k into p parts, with $(k_1 \geq k_2 \geq \dots \geq k_p \geq 0)$ and $k_1 + k_2 + \dots + k_p = k$. Also, $(\cdot)_{\mathcal{K}}$ is the complex multivariate hypergeometric coefficient [14], and $\tilde{P}_{\mathcal{K}}(\cdot)$ is the complex Hayakawa polynomial with two matrix arguments.

In [9], various general statistical properties were derived for complex matrix-variate noncentral quadratic forms. These properties however, also involved Hayakawa polynomials and infinite summations, and were not suitable for efficient numerical evaluation.

C. Elementary Symmetric Functions

Definition IV: The ℓ^{th} elementary symmetric function (e.s.f.) of the matrix $\mathbf{X} \in \mathcal{C}^{n \times n}$ is defined as [15]

$$\text{tr}_\ell(\mathbf{X}) = \sum_{\{\alpha_{\ell,n}\}} \det(\mathbf{X}_{\alpha_{\ell,n}}^{\alpha_{\ell,n}}) \quad (6)$$

for $\ell = 1, \dots, n$, and $\text{tr}_0(\mathbf{X}) = 1$.

III. NEW STATISTICAL PROPERTIES OF COMPLEX NONCENTRAL QUADRATIC FORMS

We now present statistical properties of complex noncentral quadratic forms. These results are new, unless otherwise indicated. The new results are simple closed-form expressions which do not involve any infinite series or matrix-variate polynomials.

A. Moments of the Determinant

Theorem 1: Let $\mathbf{Q} \sim \mathcal{Q}_{p,q}(\mathbf{A}, \Phi, \Psi, \Upsilon)$. Then

$$E[\det(\mathbf{Q})] = \det(\Phi) \sum_{\{\alpha_{p,q}\}} \det(\mathbf{B}_{\alpha_{p,q}}^{\alpha_{p,q}}) (p - L_{\alpha_{p,q}})! \times \frac{\det\left(\left(p - L_{\alpha_{p,q}} + j + \theta_i\right) \theta_i^{j-1}\right)}{\prod_{i < j}^{L_{\alpha_{p,q}}} (\theta_j - \theta_i)} \quad (7)$$

where $\theta_1, \dots, \theta_{L_{\alpha_{p,q}}}$ are the non-zero eigenvalues of

$$\Theta(\alpha_{p,q}) = \left(\mathbf{B}_{\alpha_{p,q}}^{\alpha_{p,q}}\right)^{-1} (\tilde{\Upsilon}_{\alpha_{p,q}})^\dagger \tilde{\Upsilon}_{\alpha_{p,q}} \quad (8)$$

where

$$\mathbf{B} = \mathbf{A}^{1/2} \Psi \mathbf{A}^{1/2} \quad (9)$$

$$\tilde{\Upsilon} = \Phi^{-1/2} \Upsilon \mathbf{A}^{1/2} \quad (10)$$

Proof: We start by writing

$$\begin{aligned} E[\det(\mathbf{Q})] &= E[\det(\mathbf{X}\mathbf{A}\mathbf{X}^\dagger)] \\ &= E\left[\det\left(\Phi^{1/2} \tilde{\mathbf{X}} \tilde{\mathbf{X}}^\dagger \Phi^{1/2}\right)\right] \\ &= \det(\Phi) E[\det(\tilde{\mathbf{X}} \tilde{\mathbf{X}}^\dagger)] \end{aligned} \quad (11)$$

where

$$\tilde{\mathbf{X}} \sim \mathcal{CN}_{p,q}(\tilde{\Upsilon}, \mathbf{I}_p \otimes \mathbf{B}) \quad (12)$$

and where the last line followed from the property

$$\det(\mathbf{C}\mathbf{D}) = \det(\mathbf{C}) \det(\mathbf{D}) \quad (13)$$

for arbitrary square matrices \mathbf{C} and \mathbf{D} .

To calculate the expectation in (11) we expand the determinant by applying the Cauchy-Binet formula for the determinant of a product of two matrices (see [3, 16]), to give

$$\begin{aligned} E[\det(\tilde{\mathbf{X}} \tilde{\mathbf{X}}^\dagger)] &= \sum_{\{\alpha_{p,q}\}} E\left[\det\left(\tilde{\mathbf{X}}_{\alpha_{p,q}} (\tilde{\mathbf{X}}^\dagger)^{\alpha_{p,q}}\right)\right] \\ &= \sum_{\{\alpha_{p,q}\}} E\left[\det\left(\tilde{\mathbf{X}}_{\alpha_{p,q}} (\tilde{\mathbf{X}}_{\alpha_{p,q}})^\dagger\right)\right] \\ &= \sum_{\{\alpha_{p,q}\}} E\left[\det\left((\tilde{\mathbf{X}}_{\alpha_{p,q}})^\dagger \tilde{\mathbf{X}}_{\alpha_{p,q}}\right)\right]. \end{aligned} \quad (14)$$

Now, from [9, Lemma 1] we have

$$(\tilde{\mathbf{X}}_{\alpha_{p,q}})^\dagger \sim \mathcal{CN}_{p,p}\left((\tilde{\Upsilon}_{\alpha_{p,q}})^\dagger, \mathbf{B}_{\alpha_{p,q}}^{\alpha_{p,q}} \otimes \mathbf{I}_p\right) \quad (15)$$

and therefore

$$(\tilde{\mathbf{X}}_{\alpha_{p,q}})^\dagger \tilde{\mathbf{X}}_{\alpha_{p,q}} \sim \mathcal{W}_p\left(p, \mathbf{B}_{\alpha_{p,q}}^{\alpha_{p,q}}, \Theta(\alpha_{p,q})\right). \quad (16)$$

The theorem follows by directly evaluating the expectations in (14) using (16), along with [9, Corr. 1], and substituting the result into (11). \blacksquare

It is important to emphasize that the summation in (7) is a *finite* summation over a collection of $\binom{q}{p}$ subsets, as defined in Section II-A. For example, for $p = 2$ and $q = 3$, the summation involves only 3 terms corresponding to the elements of $\{\alpha_{2,3}\} = \{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$.

Theorem 2: Let $\mathbf{Q} \sim \mathcal{Q}_{p,q}(\mathbf{A}, \Phi, \Psi, \Upsilon)$. Then

$$E[\det(\mathbf{Q})^h] \geq \left(\prod_{i=q-p+1}^q \lambda_i^h\right) \det(\Phi)^h \frac{\tilde{\Gamma}_p(q+h)}{\tilde{\Gamma}_p(q)} \text{etr}(-\Theta) \times \frac{\det\left({}_1F_1(q+h-L+j, q-L+j, \theta_i) \theta_i^{j-1}\right)}{\prod_{i < j}^L (\theta_j - \theta_i)} \quad (17)$$

where $\theta_1, \dots, \theta_L$ are the non-zero eigenvalues of

$$\Theta = \Phi^{-1} \Upsilon \Psi^{-1} \Upsilon^\dagger \quad (18)$$

and $\lambda_1 \geq \dots \geq \lambda_q > 0$ are the eigenvalues of $\Psi^{1/2} \mathbf{A} \Psi^{1/2}$, and ${}_1F_1(\cdot)$ is the scalar confluent hypergeometric function.

Proof: We begin by noting that

$$E \left[\det(\mathbf{Q})^h \right] = E \left[\det \left(\tilde{\mathbf{X}} \Psi^{1/2} \mathbf{A} \Psi^{1/2} \tilde{\mathbf{X}} \right)^h \right] \quad (19)$$

where

$$\tilde{\mathbf{X}} \sim \mathcal{CN}_{p,q} \left(\Upsilon \Psi^{-1/2}, \Phi \otimes \mathbf{I}_q \right). \quad (20)$$

Now applying the inequality [11]

$$\det(\mathbf{C} \mathbf{D} \mathbf{C}^\dagger) \geq \left(\prod_{i=q-p+1}^q d_i \right) \det(\mathbf{C} \mathbf{C}^\dagger) \quad (21)$$

for $\mathbf{D} \in \mathcal{C}^{q \times q} > 0$, with eigenvalues $d_1 \geq \dots \geq d_q > 0$, and arbitrary $\mathbf{C} \in \mathcal{C}^{p \times q}$ (where $q \geq p$), to the right-hand side of (19), and then directly applying [9, Theorem 1] yields the desired result. ■

B. Expected Elementary Symmetric Functions

To derive the main theorem of this subsection we require the following lemma, which is a simple extension of [9, Lemma 3] to allow for the case $\Psi \neq \mathbf{I}_q$.

Lemma 1: Let $\mathbf{Q} \sim \mathcal{Q}_{p,q}(\mathbf{A}, \Phi, \Psi, \Upsilon)$, and let $t \leq p$. Then

$$\mathbf{Q}_{\alpha_{t,p}}^{\alpha_{t,p}} \sim \mathcal{Q}_{t,q} \left(\mathbf{A}, \Phi_{\alpha_{t,p}}^{\alpha_{t,p}}, \Psi, \Upsilon^{\alpha_{t,p}} \right) \quad (22)$$

Proof: Omitted. ■

Theorem 3: Let $\mathbf{Q} \sim \mathcal{Q}_{p,q}(\mathbf{A}, \Phi, \Psi, \Upsilon)$. Then the expected value of the ℓ^{th} e.s.f. of \mathbf{Q} is given by

$$E[\text{tr}_\ell(\mathbf{Q})] = \sum_{\{\alpha_{\ell,p}\}} \det(\Phi_{\alpha_{\ell,p}}^{\alpha_{\ell,p}}) \sum_{\{\alpha_{\ell,q}\}} \det(\mathbf{B}_{\alpha_{\ell,q}}^{\alpha_{\ell,q}}) \times (\ell - L_{\alpha_{\ell,q}})! \frac{\det \left((\ell - L_{\alpha_{\ell,q}} + j + \theta_i) \theta_i^{j-1} \right)}{\prod_{i < j}^{L_{\alpha_{\ell,q}}} (\theta_j - \theta_i)} \quad (23)$$

where \mathbf{B} is defined as in (9), and $\theta_1, \dots, \theta_{L_{\alpha_{\ell,q}}}$ are the non-zero eigenvalues of

$$\Theta(\alpha_{\ell,q}, \alpha_{\ell,p}) = (\mathbf{B}_{\alpha_{\ell,q}}^{\alpha_{\ell,q}})^{-1} \Upsilon(\alpha_{\ell,q}, \alpha_{\ell,p})^\dagger \bar{\Upsilon}(\alpha_{\ell,q}, \alpha_{\ell,p}) \quad (24)$$

where

$$\bar{\Upsilon}(\alpha_{\ell,q}, \alpha_{\ell,p}) = \left((\Phi_{\alpha_{\ell,p}}^{\alpha_{\ell,p}})^{-1/2} \Upsilon^{\alpha_{\ell,p}} \mathbf{A}^{1/2} \right)_{\alpha_{\ell,q}} \quad (25)$$

Proof: The proof follows by first applying Definition IV to the left-hand side of (7), and then directly invoking Lemma 1 and Theorem 1. ■

Note that for the special case $\Phi = \mathbf{I}_p$, $\Psi = \mathbf{I}_q$, $\mathbf{A} = \mathbf{I}_q$, this result can be shown to reduce to an expression reported previously in [17].

C. Expected Characteristic Polynomial

Theorem 4: Let $\mathbf{Q} \sim \mathcal{Q}_{p,q}(\mathbf{A}, \Phi, \Psi, \Upsilon)$, and μ be an arbitrary real-valued constant. Then the expected characteristic polynomial of \mathbf{Q} is given by

$$E[\det(\mu \mathbf{I}_p - \mathbf{Q})] = \sum_{\ell=0}^p \mu^{p-\ell} (-1)^\ell \sum_{\{\alpha_{\ell,p}\}} \det(\Phi_{\alpha_{\ell,p}}^{\alpha_{\ell,p}}) \times \sum_{\{\alpha_{\ell,q}\}} \det(\mathbf{B}_{\alpha_{\ell,q}}^{\alpha_{\ell,q}}) (\ell - L_{\alpha_{\ell,q}})! \frac{\det \left((\ell - L_{\alpha_{\ell,q}} + j + \theta_i) \theta_i^{j-1} \right)}{\prod_{i < j}^{L_{\alpha_{\ell,q}}} (\theta_j - \theta_i)} \quad (26)$$

where $\theta_1, \dots, \theta_{L_{\alpha_{\ell,q}}}$ and \mathbf{B} are defined as in Theorem 3.

Proof: We begin by writing

$$\begin{aligned} E[\det(\mu \mathbf{I}_p - \mathbf{Q})] &= \mu^p E \left[\det \left(\mathbf{I}_p - \frac{1}{\mu} \mathbf{Q} \right) \right] \\ &= \mu^p E \left[\sum_{\ell=0}^p \text{tr}_\ell \left(-\frac{1}{\mu} \mathbf{Q} \right) \right] \\ &= \sum_{\ell=0}^p \mu^{p-\ell} (-1)^\ell E[\text{tr}_\ell(\mathbf{Q})] \quad (27) \end{aligned}$$

where the second line followed from a determinant expansion given in [16], and the last line followed from Definition IV. The result is now obtained by application of Theorem 3. ■

Note that, for the special case $\Psi = \mathbf{I}_q$ and $\mathbf{A} = \mathbf{I}_q$, it can be easily verified that this generalized result agrees with previous expressions given in [9].

IV. MUTUAL INFORMATION OF DOUBLE-SIDED CORRELATED RICIAN MIMO CHANNELS

Consider a flat-fading $N_t \times N_r$ MIMO link modelled by

$$\mathbf{r} = \mathbf{H} \mathbf{a} + \mathbf{n} \quad (28)$$

where $\mathbf{r} \in \mathcal{C}^{N_r \times 1}$ is the discrete-time received signal vector, $\mathbf{a} \in \mathcal{C}^{N_t \times 1}$ is the transmitted signal vector satisfying the power constraint $E[\mathbf{a}^\dagger \mathbf{a}] \leq P$, and $\mathbf{n} \in \mathcal{C}^{N_r \times 1}$ is a vector of zero-mean additive complex Gaussian noise with $E[\mathbf{n} \mathbf{n}^\dagger] = \sigma_n^2 \mathbf{I}_{N_r}$. Also, $\mathbf{H} \in \mathcal{C}^{N_r \times N_t}$ is the DSC-ARM MIMO channel matrix, with $(i, j)^{\text{th}}$ element containing the complex fading parameter between the j^{th} transmit and i^{th} receive antenna.

The columns of \mathbf{H} are complex Gaussian random vectors, each having the same Hermitian covariance matrix given by the receive correlation matrix \mathbf{R} . We allow the mean vectors of each of the columns of \mathbf{H} to be different. The rows of \mathbf{H} are also modelled as complex Gaussian random vectors (transposed), each with covariance matrix given by the transmit correlation matrix \mathbf{S} . We assume \mathbf{R} and \mathbf{S} are positive definite full rank. Under these assumptions, the channel may be decomposed as

$$\begin{aligned} \mathbf{H} &= \sqrt{a} \mathbf{M} + \sqrt{b} \mathbf{R}^{1/2} \mathbf{H}_w \mathbf{S}^{1/2} \\ &\sim \mathcal{CN}_{N_r, N_t}(\sqrt{a} \mathbf{M}, b \mathbf{R} \otimes \mathbf{S}) \end{aligned} \quad (29)$$

where $\mathbf{R}^{1/2}$ and $\mathbf{S}^{1/2}$ denote the Hermitian square roots of \mathbf{R} and \mathbf{S} respectively, \mathbf{M} is the arbitrary rank mean matrix satisfying $\text{tr}(\mathbf{M} \mathbf{M}^\dagger) = N_r N_t$, a and b are power normalization coefficients¹, and $\mathbf{H}_w \sim \mathcal{CN}_{n,m}(\mathbf{0}, \mathbf{I}_p \otimes \mathbf{I}_q)$. We note that this decomposition has been used extensively

¹In most cases, a and b are chosen to satisfy $a + b = 1$.

in the literature [18, 19], and is also supported by physical measurements [20].

In this paper we focus on the MI of DSC-ARM MIMO channels with isotropic input signalling. In this case it is well known that the ergodic MI is given by

$$\mathcal{I} = E \left[\log \det \left(\mathbf{I}_{N_r} + \frac{\gamma}{N_t} \mathbf{H} \mathbf{H}^\dagger \right) \right] \quad (30)$$

where $\gamma = P/\sigma_n^2$ is the average signal-to-noise ratio (SNR). Define $n = \min(N_r, N_t)$, $m = \max(N_r, N_t)$, and

$$\mathbf{A} \triangleq \begin{cases} \mathbf{R} & \text{for } N_r \leq N_t \\ \mathbf{S} & \text{for } N_r > N_t \end{cases} \quad \mathbf{\Sigma} \triangleq \begin{cases} \mathbf{S} & \text{for } N_r \leq N_t \\ \mathbf{R} & \text{for } N_r > N_t \end{cases} \quad (31)$$

$$\bar{\mathbf{M}} \triangleq \begin{cases} \mathbf{M} & \text{for } N_r \leq N_t \\ \mathbf{M}^\dagger & \text{for } N_r > N_t \end{cases}$$

Now substituting (29) into (30), and noting that

$$\det(\mathbf{I} + \mathbf{A}\mathbf{B}) = \det(\mathbf{I} + \mathbf{B}\mathbf{A}) \quad (32)$$

for arbitrary \mathbf{A} and \mathbf{B} , we may write

$$\mathcal{I} = E \left[\log \det \left(\mathbf{I}_n + \frac{\gamma}{N_t} \mathbf{Q} \right) \right] \quad (33)$$

where

$$\mathbf{Q} \sim \mathcal{Q}_{n,m}(\mathbf{I}_m, b\mathbf{A}, \mathbf{\Sigma}, \sqrt{a}\bar{\mathbf{M}}) \quad (34)$$

In this paper we are interested in finding bounds on the ergodic MI (33) using the new statistical properties derived in Section III. These are presented in the following section.

V. BOUNDS ON THE MUTUAL INFORMATION

A. Upper Bound

In this subsection we derive an efficiently computable upper bound on ergodic MI. We note that an upper bound for the general DSC-ARM MIMO channels considered in this paper was derived previously in [9, Theor. 9], and an upper bound for the more restrictive case of rank-1 means was presented in [11, Theor. 8]. We will show in Section VI that the new upper bound derived in this subsection is significantly tighter than these previous results.

We start by noting that $\log \det(\cdot)$ is a concave function on the set of Hermitian positive definite matrices [1], and using Jensen's inequality to upper bound the ergodic MI as follows

$$\mathcal{I} \leq \log E \left[\det \left(\mathbf{I}_n + \frac{\gamma}{N_t} \mathbf{Q} \right) \right] \quad (35)$$

We now evaluate the expectation in (35) using (34) and Theorem 4, and perform some basic algebraic manipulation, to obtain the closed-form upper bound expression

$$\mathcal{I} \leq \log \left(\sum_{\ell=0}^n \left(\frac{b\gamma}{N_t} \right)^\ell \sum_{\{\alpha_{\ell,n}\}} \det(\mathbf{A}_{\alpha_{\ell,n}}^{\alpha_{\ell,n}}) \sum_{\{\alpha_{\ell,m}\}} \det(\mathbf{\Sigma}_{\alpha_{\ell,m}}^{\alpha_{\ell,m}}) \right. \\ \left. \times (\ell - L_{\alpha_{\ell,m}})! \frac{\det \left(\left((\ell - L_{\alpha_{\ell,m}} + j + \tilde{\theta}_i) \tilde{\theta}_i^{j-1} \right) \right)}{\prod_{i < j}^{L_{\alpha_{\ell,m}}} (\tilde{\theta}_j - \tilde{\theta}_i)} \right) \quad (36)$$

where $\tilde{\theta}_1, \dots, \tilde{\theta}_{L_{\alpha_{\ell,m}}}$ are the non-zero eigenvalues of

$$\tilde{\Theta}(\alpha_{\ell,m}, \alpha_{\ell,n}) = \frac{a}{b} (\mathbf{\Sigma}_{\alpha_{\ell,m}}^{\alpha_{\ell,m}})^{-1} (\bar{\mathbf{M}}_{\alpha_{\ell,m}}^{\alpha_{\ell,m}})^\dagger (\mathbf{A}_{\alpha_{\ell,n}}^{\alpha_{\ell,n}})^{-1} (\bar{\mathbf{M}}_{\alpha_{\ell,n}}^{\alpha_{\ell,n}}) \quad (37)$$

At high SNRs, this expression reduces to

$$\mathcal{I} \leq n \log \left(\frac{b\gamma}{N_t} \right) + \log \det(\mathbf{A}) + \log \left(\sum_{\{\alpha_{n,m}\}} \det(\mathbf{\Sigma}_{\alpha_{n,m}}^{\alpha_{n,m}}) \right. \\ \left. \times (n - L_{\alpha_{n,m}})! \frac{\det \left(\left((n - L_{\alpha_{n,m}} + j + \tilde{\theta}_i) \tilde{\theta}_i^{j-1} \right) \right)}{\prod_{i < j}^{L_{\alpha_{n,m}}} (\tilde{\theta}_j - \tilde{\theta}_i)} \right) \quad (38)$$

For the commonly-assumed special case where $\bar{\mathbf{M}}$ is rank-1 (e.g. see [7, 8, 11]), this becomes

$$\mathcal{I} \leq n \log \left(\frac{b\gamma}{N_t} \right) + \log \det(\mathbf{A}) + \log(n-1)! \\ + \log \left(\sum_{\{\alpha_{n,m}\}} \det(\mathbf{\Sigma}_{\alpha_{n,m}}^{\alpha_{n,m}}) \left(n + \text{tr}(\tilde{\Theta}(\alpha_{n,m}, \alpha_{n,n})) \right) \right) \quad (39)$$

We remark that (38) and (39) can also be directly used to obtain tight bounds on the *high SNR power offset* of DSC-ARM MIMO channels, as defined in [21]².

B. Lower Bound

In this subsection we derive the first general, computable lower bound on the ergodic MI of DSC-ARM MIMO channels. Note that computable lower bounds were derived previously only for the special case of $n \times n$ systems [9, Theor. 7], and for the special case of rank-1 mean matrices [11].

We begin by applying the general lower bounding approach in [6] to (33), which yields

$$\mathcal{I} \geq \log \left(1 + \sum_{\ell=1}^n \left(\frac{\gamma}{N_t} \right)^\ell \sum_{\{\alpha_{\ell,n}\}} \exp(E[\ln \det(\mathbf{Q}_{\alpha_{\ell,n}}^{\alpha_{\ell,n}})]) \right) \quad (40)$$

We now reformulate the remaining expectation using a standard moment generating function technique (e.g. see [9]) as

$$E[\ln \det(\mathbf{Q}_{\alpha_{\ell,n}}^{\alpha_{\ell,n}})] = \frac{d}{ds} \ln E[\det(\mathbf{Q}_{\alpha_{\ell,n}}^{\alpha_{\ell,n}})^s] \Big|_{s=0} \quad (41)$$

Omitting details, we calculate a closed-form lower bound on the expectation on the right-hand side of (41) by applying Lemma 1 and Theorem 2, and evaluate the remaining derivatives using a technique from [10]. Substituting the resulting expression into (40) yields the desired closed-form ergodic MI lower bound given by

$$\mathcal{I} \geq \log \left(1 + \sum_{\ell=1}^n \left(\frac{b\gamma}{N_t} \right)^\ell \left(\prod_{t=m-\ell+1}^m s_t \right) \exp \left(\sum_{t=0}^{\ell-1} \psi(m-t) \right) \right. \\ \left. \times \sum_{\{\alpha_{\ell,n}\}} \det(\mathbf{A}_{\alpha_{\ell,n}}^{\alpha_{\ell,n}}) \exp \left(\frac{\sum_{t=1}^{L_{\alpha_{\ell,n}}} \det(\mathbf{V}_{\alpha_{\ell,n},t})}{\prod_{i < j}^{L_{\alpha_{\ell,n}}} (\tilde{\theta}_j - \tilde{\theta}_i)} \right) \right) \quad (42)$$

where $\psi(\cdot)$ is the digamma function, $s_1 \geq \dots \geq s_m > 0$ are the eigenvalues of $\mathbf{\Sigma}$, and $\tilde{\theta}_1, \dots, \tilde{\theta}_{L_{\alpha_{\ell,n}}}$ are the non-zero eigenvalues of

$$\tilde{\Theta}(\alpha_{\ell,n}) = \frac{a}{b} (\mathbf{A}_{\alpha_{\ell,n}}^{\alpha_{\ell,n}})^{-1} \bar{\mathbf{M}}^{\alpha_{\ell,n}} \mathbf{\Sigma}^{-1} (\bar{\mathbf{M}}^{\alpha_{\ell,n}})^\dagger \quad (43)$$

Also, $\mathbf{V}_{\alpha_{\ell,n},t}$ is an $L_{\alpha_{\ell,n}} \times L_{\alpha_{\ell,n}}$ matrix with (i, j) th element defined in (44), where $\mathcal{V}_{\alpha_{\ell,n},j} = m - L_{\alpha_{\ell,n}} + j$, and $E_1(\cdot)$ is the exponential integral.

²In [21], high SNR power offset expressions were presented for DSC-ARM MIMO channels, but only for the special case of $n \times n$ systems.

$$(\mathbf{V}_{\alpha_{\ell,n,t}})_{i,j} = \begin{cases} \tilde{\theta}_i^{j-1} & \text{for } i \neq t \\ \tilde{\theta}_i^{j-1} \left(\ln(\tilde{\theta}_i) - \text{E}_i(-\tilde{\theta}_i) + \psi(\mathcal{V}_{\alpha_{\ell,n,j}}) + \sum_{k=1}^{\mathcal{V}_{\alpha_{\ell,n,j}-1}} \frac{(k-1)!}{(-\tilde{\theta}_i)^k} \left(e^{-\tilde{\theta}_i} - \binom{\mathcal{V}_{\alpha_{\ell,n,j}-1}{k}} \right) \right) & \text{for } i = t \end{cases} \quad (44)$$

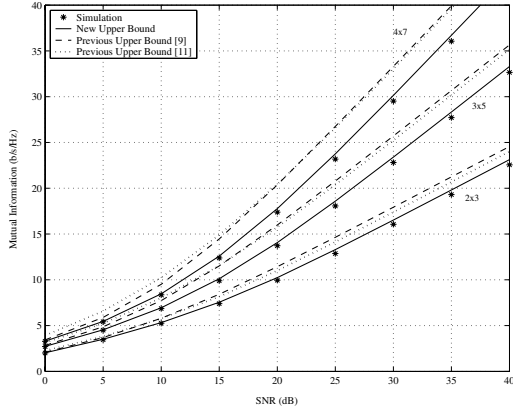


Fig. 1. Upper bounds and simulation results for mutual information of correlated Rician MIMO channels for various antenna configurations, with rank-1 mean matrix and Rician K -factor = 10. Correlation parameters are $\theta_r = \frac{\pi}{2}$, $\theta_t = \frac{\pi}{2}$, $\sigma_r^2 = \frac{\pi}{32}$, $\sigma_t^2 = \frac{\pi}{16}$ and $d_t = d_r = \frac{1}{2}$.

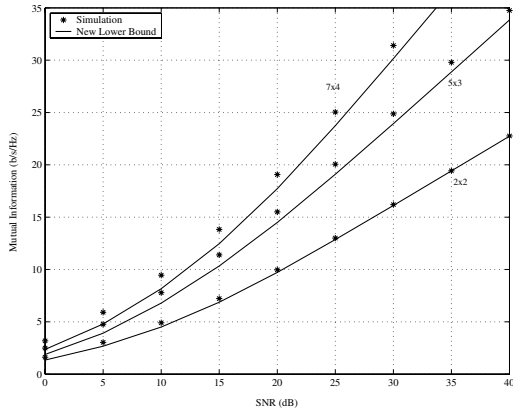


Fig. 2. Lower bounds and simulation results for mutual information of correlated Rician MIMO channels for various antenna configurations, with rank-2 mean matrix and Rician K -factor = 10. Correlation parameters are $\theta_r = \frac{\pi}{2}$, $\theta_t = \frac{\pi}{2}$, $\sigma_r^2 = \frac{\pi}{16}$, $\sigma_t^2 = \frac{\pi}{8}$ and $d_t = d_r = \frac{1}{2}$.

VI. NUMERICAL RESULTS

We consider a common power normalization model with $a = K/(K+1)$ and $b = 1/(K+1)$, where K is the Rician K -factor which is the ratio of the power in the fixed (mean) component with respect to the average power in the fading components. Since $a + b = 1$, for fixed total transmit power the received SNR remains constant for any value of K .

The mean and correlation matrices are generated using the practical channel model from [22].

Fig. 1 gives the upper bound (36) and simulated ergodic MI curves for 2×3 , 3×5 , and 4×7 systems. Results are shown for rank-1 mean matrices with $K = 10$. The previously reported general upper bound in [9, Theor. 9], as well as the more restrictive bound from [11, Theor. 8] (i.e. applying only for the rank-1 mean case), are also shown for comparison. Clearly

the new bound (36) is much tighter than these previous results in all cases.

Fig. 2 gives the lower bound (42) and simulated ergodic MI curves for 2×2 , 5×3 , and 7×4 systems. We see that the lower bound is tight in all cases. It is important to note that there do not appear to be any other lower bounds in the literature which apply for these general DSC-ARM channels.

REFERENCES

- [1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Commun.*, pp. 585–595, Nov-Dec. 1999.
- [2] P. J. Smith, S. Roy, and M. Shaif, "Capacity of MIMO systems with semicorrelated flat fading," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2781–2788, Oct 2003.
- [3] 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.
- [4] Y.-H. Kim and A. Lapidoth, "On the log determinant of non-central Wishart matrices," in *IEEE Int. Symp. on Info. Theory (ISIT)*, Yokohama, Japan, 2003, p. 54.
- [5] S. K. Jayaweera and H. V. Poor, "On the capacity of multiple-antenna systems in Rician fading," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1102–1111, May 2005.
- [6] Q. T. Zhang, X. W. Cui, and X. M. Li, "Very tight capacity bounds for MIMO-correlated Rayleigh-fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 681–688, Mar. 2005.
- [7] X. W. Cui, Q. T. Zhang, and Z. M. Feng, "Generic procedure for tightly bounding the capacity of MIMO correlated Rician fading channels," *IEEE Trans. Commun.*, vol. 53, no. 5, pp. 890–898, May 2005.
- [8] M. R. McKay and I. B. Collings, "Capacity bounds for correlated Rician MIMO channels," in *IEEE Int. Conf. Commun. (ICC)*, Seoul, Korea, May 2005.
- [9] —, "General capacity bounds for spatially correlated Rician MIMO channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 9, pp. 3121–3145, Sept. 2005.
- [10] —, "Improved general lower bound for spatially-correlated Rician MIMO capacity," *IEEE Commun. Lett.*, pp. 162–164, Mar 2006.
- [11] S. Jin, X. Gao, and X. You, "On the ergodic capacity of rank-1 Rician fading MIMO channels," *IEEE Trans. Inform. Theory*, 2005, submitted (revised).
- [12] W. J. Conradie and C. G. Troskie, "The exact non-central distribution of a multivariate complex quadratic form of complex normal variables," *South African Statistical Journal*, vol. 18, pp. 123–134, 1984.
- [13] A. K. Gupta and D. K. Nagar, *Matrix Variate Distributions*. Boca Raton: Chapman & Hall/CRC, 2000.
- [14] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Ann. Math. Statist.*, vol. 35, pp. 475–501, 1964.
- [15] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, 1st ed. New York: Academic Press, 1979.
- [16] A. C. Aitken, *Determinants and Matrices*, 9th ed. Edinburgh: Oliver and Boyd, 1956.
- [17] B. K. Shah and C. G. Khatri, "Proof of conjectures about the expected values of the elementary symmetric functions of a noncentral Wishart matrix," *Ann. of Stat.*, vol. 2, no. 4, pp. 833–836, 1974.
- [18] D. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502–513, Mar 2000.
- [19] C. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 637–650, Mar 2002.
- [20] J. P. Keramoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Select. Areas Commun.*, vol. 20, no. 6, pp. 1211–1226, 2002.
- [21] A. Lozano, A. Tulino, and S. Verdú, "High-SNR power offset in multiantenna communication," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4134–4151, Dec 2005.
- [22] H. Bölcskei, M. Borgmann, and A. J. Paulraj, "Impact of the propagation environment on the performance of space-frequency coded MIMO-OFDM," *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 427–439, Apr 2003.