Unifying Analysis of Ergodic MIMO Capacity in Correlated Rayleigh Fading Environments

Mario Kießling, *Student Member, IEEE*

Abstract

We present a novel mathematical approach that for the first time allows for calculating the moment generating function of mutual information of Rayleigh fading MIMO channels with arbitrary fading correlation at transmitter and receiver. Specifically, we make use of certain hypergeometric functions of matrix argument, which are frequently used in multivariate statistics. This allows for a concise general expression for the moment generating function and we specialize the general result to scenarios with one-side correlated and uncorrelated channels. Using the moment generating functions, we derive exact formulas of the ergodic capacity, thus unifying the capacity analysis of correlated Rayleigh fading MIMO channels. It turns out that the ergodic capacity in all cases can be expressed in terms of a sum of determinants with elements that are a combination of polynomials, exponentials, and the exponential integral E_1 solely. The analysis is verified by Monte-Carlo simulations and shows a perfect match.

Index Terms

MIMO, ergodic capacity, fading correlation

I. Introduction

In his seminal paper [1], Telatar calculated the ergodic capacity of a multiple input multiple output (MIMO) link with uncorrelated Rayleigh fading in an additive white Gaussian noise (AWGN) environment in terms of a single scalar integral by integrating over the eigenvalue probability density function (PDF) of certain complex Wishart matrices [2], thereby predicting enormous capacity gains by combined spatial processing at transmitter and receiver and thus

Manuscript received September 2004

M. Kießling is with the University of Stuttgart.

IEEE European Transactions on Telecommunications , Volume 16, Issue 1, Januar 2005, pp. 17-35

initiating immense research activities in this area. Later, Foschini and Gans presented numerical results and bounds on i.i.d. Rayleigh MIMO outage capacity in their fundamental work [3]. Extensions to correlated MIMO channels with fading correlation at one side of the link were given recently in [4] [5] [6]. Bounds on the ergodic capacity of i.i.d. and correlated Rayleigh fading MIMO channels were given in [7] [8] [9] [10] [39] in a mathematically more tractable form than the expressions resulting from the exact analysis.

Generalizing above results, in this work we calculate the exact moment generating function (MGF) of mutual information of MIMO channels with correlated Rayleigh fading at both transmitter and receiver. To this end, we present a novel mathematical approach. A concise mathematical formulation of the MGF is given in terms of a hypergeometric function of matrix arguments [2] [11] and in terms of determinants of scalar hypergeometric functions. In contrast to existing literature [4] [12], our approach is not based on eigenvalue PDFs but uses a direct integration technique. Having available a closed-form MGF, we can derive exact, i.e. non-asymptotic moments of mutual information, including ergodic capacity (the first moment, which is in the focus of this paper) for arbitrary array sizes and arbitrary correlation properties at receiver as well as transmitter, thus unifying and completing existing partial solutions for special propagation scenarios.

Specifically, ergodic MIMO capacity is given in terms of a sum of determinants, whereas the determinant entries are given in terms of polynomials, exponentials, and the only special function that appears is the exponential integral E_1. Due to the space limitation, in this paper we consider the case, where no channel state information (CSI) is available at the transmitter and transmit power is equally distributed to the transmit antennas, i.e. the transmit covariance is a scaled identity matrix [1] (for extensions with statistical CSI at the transmitter see [40] [41]). Ideal short-term channel state information is assumed to be available at the receiver.

We note that the MGF of mutual information given in this paper can be used to derive approximations of the distribution of mutual information by standard distributions like Gaussian [13] or Gamma [14]. These distributions can be characterized by their first and second moments, which can be derived analytically with the help of the MGF.

For completeness, we provide an overview of existing results on the MGF of mutual information of MIMO wireless channels and in particular ergodic capacity. Specifically, by integrating over the eigenvalue PDF of an i.i.d. complex Wishart matrix, in [12] the MGF of the mutual information of an i.i.d. Rayleigh channel is derived. Based on the MGF, numerical Laplace transform inversion techniques are used for calculating outage capacity [15]. A similar MGF approach is taken in [4], where the authors present results for various propagation scenarios including i.i.d. and one-side correlated Rayleigh fading, as well as Ricean fading. Recently, similar results for the case of a Rayleigh fading channel with fading correlation at one side of the MIMO link were obtained in [5] and [6]. Again, a mathematically challenging integration over the eigenvalue PDF of (non-central) Wishart matrices is necessary, which requires the use of certain integration results from statistical literature [16] [17] [18]. However, we emphasize that this eigenvalue based approach until recently prohibited a general solution for the MGF in Rayleigh fading with both receive and transmit correlation, as there were no suitable formulas available for the eigenvalue PDF of certain complex generalized random quadratic forms. This gap was closed in [19], where certain character expansion techniques were used for finding a compact expression for the eigenvalue PDF.

Another approach for characterizing mutual information is an asymptotical analysis, where the number of antenna elements at receiver and transmitter goes to infinity, however has a fixed ratio. In this case, the empirical eigenvalue distribution of the random channel matrix can be obtained in a manageable form [21] [20], such that the capacity expressions simplify and it turns out that the results are a good approximation even for practically relevant systems with a fairly small number of antennas. Asymptotic results can be found for i.i.d. Rayleigh fading in [22] and for the one-side correlated case (i.e. transmit or receive correlation exclusively) in [23]. The same techniques were earlier successfully applied in the analysis of multi user code division multiple access (CDMA) wireless systems [24]. Asymptotical steepest descent integration methods that can frequently be encountered in statistical physics were the key for generalizing the asymptotic results to arbitrarily correlated MIMO channels in [25] (however, we mention that for calculating the asymptotics, a numerical solution of a system of nonlinear equations is necessary).

Finally, we note that a detailed overview of existing literature and results on MIMO capacity is provided in [26]. A powerful survey on random matrix theory is [27], which contains an exhaustive list of references and in particular many results on MIMO capacity.

The remainder of this paper is organized as follows. In Section II we introduce the notation and system parameters that will be used consistently throughout this work. Furthermore, we characterize the statistics of the underlying correlated Rayleigh fading MIMO channel with arbitrary correlation at receiver and transmitter, respectively. A unifying notation of MIMO mutual information for arbitrary system parameters is developed in Section III. The MGF of mutual information is determined in Section IV. We provide concise expressions in terms of hypergeometric functions of matrix arguments and corresponding representations in terms of determinants of scalar hypergeometric functions. Formulas are given for fully correlated channels, channels with one sided fading correlation, and for uncorrelated channels. Based on the MGFs, in Section V we derive closed form ergodic capacity expressions for different propagation scenarios. Numerical results are presented in Section VI, which demonstrate a perfect match of the ergodic capacity analysis and Monte-Carlo simulation results. In the appendix we provide a comprehensive introduction to hypergeometric functions of matrix arguments and derive closed form scalar representations of some special cases that are important in the analysis of MIMO mutual information. Furthermore, we analyze some generalized matrix quadratic forms and present results on related multivariate integrals.

II. Notation and System Model

A. Notation

Vectors are denoted by bold lowercase letters $\mathbf{x}$, matrices by bold uppercase letters $\mathbf{X}$. Transposition is indicated by $\mathbf{X}^{\mathrm{T}}$ and complex conjugate transpose by $\mathbf{X}^{\mathrm{H}}$. By $|\mathbf{X}|$ we denote the determinant of square matrix $\mathbf{X}$, the operator $\mathrm{eig}(\mathbf{X})$ returns a diagonal matrix of eigenvalues of matrix $\mathbf{X}$, and $\mathrm{diag}(x_1, x_2, \ldots, x_n)$ returns a diagonal matrix with elements x_k on the diagonal (with an obvious extension to vectors $\mathbf{x}_k$ and diagonal matrices $\mathbf{X}_k$). $[x_{ij}]$ $(|x_{ij}|)$ is a matrix (determinant) with element x_{ij} in row i and column j. An identity matrix of size $n \times n$ is written as $\mathbf{I}_n$, whereas the index can be omitted for brevity. The matrix variate complex normal distribution with mean $\mathbf{M}$ with m rows and n columns, covariance matrix of column vectors $\mathbf{\Sigma}$, and covariance matrix of row vectors $\mathbf{\Psi}$ is written as $\mathcal{N}_{m,n}(\mathbf{M}, \mathbf{\Sigma}, \mathbf{\Psi})$. By '$\sim$' we denote 'is distributed as', while '$\simeq$' means 'has the same distribution as'. Expected value of a function $f(\mathbf{X})$ with respect to $\mathbf{X}$ reads $E_{\mathbf{X}}[f(\mathbf{X})]$. A definition is indicated by '$\equiv$'. The minimum of the elements x_k is given by $\min(x_1, x_2, \ldots, x_n)$ and $\max(x_1, x_2, \ldots, x_n)$ is the maximum. Furthermore, we make use of the abbreviation $\mathrm{etr}(\mathbf{X}) = e^{\mathrm{tr}(\mathbf{X})}$ with the trace $\mathrm{tr}(\mathbf{X})$ of matrix $\mathbf{X}$. The so-called Pochhammer symbol is defined by

$$[a]_k \equiv a \cdot (a+1) \cdot \ldots \cdot (a+k-1) = \frac{\Gamma(a+k)}{\Gamma(a)}, \tag{1}$$

where $\Gamma(x)$ is the standard Gamma function. Furthermore, let the modified multivariate Gamma function be defined by

$$\Gamma_n(r) \equiv \frac{\tilde{\Gamma}_n(r)}{\pi^{1/2 \cdot n(n-1)}} = \prod_{i=1}^{n} \Gamma(r-i+1) \tag{2}$$

The incomplete Gamma function [33] is written as $\Gamma(a,x)$, the Kummer U function is defined by the integral

$$U(a,b,z) = \frac{1}{\Gamma(a)} \cdot \int_0^\infty e^{-zt} \cdot t^{a-1} \cdot (1+t)^{b-a-1} \mathrm{d}t, \tag{3}$$

and the scalar hypergeometric function has the power series representation

$${}_2F_0(a,b;;z) = \sum_{k=0}^{\infty} \frac{[a]_k [b]_k}{k!} \cdot z^k. \tag{4}$$

The well-known Vandermonde determinant of $n \times n$ diagonal matrix $\mathbf{X} = \mathrm{diag}(x_1, x_2, \ldots, x_n)$ shall be denoted by

$$\alpha_n(\mathbf{X}) = \prod_{i<j} (x_i - x_j) = \begin{vmatrix} x_1^{n-1} & x_1^{n-2} & \ldots & x_1 & 1 \\ x_2^{n-1} & x_2^{n-2} & \ldots & x_2 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_n^{n-1} & x_n^{n-2} & \ldots & x_n & 1 \end{vmatrix}. \tag{5}$$

For brevity we introduce the following complex matrix differential

$$\mathrm{D}_c\mathbf{X} = \prod_{m=1}^{M} \prod_{n=1}^{N} \Re\{\mathrm{d}x_{11}\}\Im\{\mathrm{d}x_{11}\} \cdots \Re\{\mathrm{d}x_{mn}\}\Im\{\mathrm{d}x_{mn}\} \tag{6}$$

for complex $N \times M$ matrix $\mathbf{X} = [\Re\{x_{nm}\} + j \cdot \Im\{x_{nm}\}]$ and write for the multidimensional integral

$$\int f(\mathbf{X})\ \mathrm{D}_c\mathbf{X} = \iint\limits_{\substack{\Re\{x_{11}\} \\ \Im\{x_{11}\}}} \cdots \iint\limits_{\substack{\Re\{x_{nm}\} \\ \Im\{x_{nm}\}}} f(\mathbf{X})\ \mathrm{D}_c\mathbf{X}. \tag{7}$$

The real part of scalar x is given by $\Re\{x\}$, $\Im\{x\}$ is the imaginary part.

B. System Model

We consider a flat fading MIMO link with T transmit and R receive antennas, whereas the $R \times T$ channel matrix is given by $\mathbf{H}$. The transmit symbols are arranged in a $T \times 1$ vector $\mathbf{s}$. On the receiver side we assume additive Gaussian noise modeled by the $R \times 1$ vector $\mathbf{n}$ and the $R \times 1$ noisy received vector is denoted by $\mathbf{y}$. Transmission over the MIMO channel can than be described by

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}, \tag{8}$$

where all time indices have been dropped for brevity (we consider a stationary propagation scenario).

C. Statistics

In this paper, we investigate the transmission over a correlated Rayleigh fading MIMO link, whereas the channel matrix $\mathbf{H}$ is complex Gaussian distributed with

$$\mathbf{H} \sim \mathcal{N}_{R,T}(\mathbf{0}, \mathbf{R}_{\mathrm{RX}}, \mathbf{R}_{\mathrm{TX}}). \tag{9}$$

We note that (9) is the well known MIMO channel model [28] [29] with separable correlation matrices at transmitter $\mathbf{R}_{\mathrm{TX}}$ and receiver $\mathbf{R}_{\mathrm{RX}}$ and

$$\mathbf{H} \simeq \mathbf{A}^{\mathrm{H}} \, \mathbf{H}_w \mathbf{B}, \tag{10}$$

where

$$\mathbf{R}_{\mathrm{RX}} = \mathbf{A}^{\mathrm{H}} \mathbf{A} \tag{11}$$

$$\mathbf{R}_{\mathrm{TX}} = \mathbf{B}^{\mathrm{H}} \mathbf{B} \tag{12}$$

and $\mathbf{H}_w \sim \mathcal{N}_{R,T}(\mathbf{0}, \mathbf{I}_R, \mathbf{I}_T)$ is a $R \times T$ matrix of i.i.d. Gaussian elements. We emphasize that in this paper we assume full rank correlation matrices at transmitter and receiver, respectively (it is a straightforward exercise to reduce a low rank system to an equivalent full rank system with a smaller number of antenna elements). The complex Gaussian PDF of $\mathbf{H}$ is than given by [30]

$$p_{\mathbf{H}}(\mathbf{H}) = \frac{1}{\pi^{RT}|\mathbf{R}_{\mathrm{TX}}|^{R}|\mathbf{R}_{\mathrm{RX}}|^{T}} \cdot \mathrm{etr}\left(-\mathbf{R}_{\mathrm{TX}}^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{R}_{\mathrm{RX}}^{-1}\mathbf{H}\right). \tag{13}$$

Moreover, we assume a transmit symbol vector with covariance

$$\mathbf{R}_{\mathrm{ss}} = E_s \cdot \tilde{\mathbf{R}}_{\mathrm{ss}}, \tag{14}$$

where E_s is the mean energy per transmit symbol and $\tilde{\mathbf{R}}_{\mathrm{ss}}$ is the normalized covariance with $\mathrm{tr}\left(\tilde{\mathbf{R}}_{\mathrm{ss}}\right) = T$. We emphasize that without channel state information at the transmitter, ergodic capacity is achieved by $\mathbf{R}_{\mathrm{ss}} = E_s \cdot \mathbf{I}_T$. However, we allow for an arbitrary signal covariance for generality. Equivalently, we consider additive Gaussian noise with covariance

$$\mathbf{R}_{\mathrm{nn}} = N_0 \cdot \tilde{\mathbf{R}}_{\mathrm{nn}}, \tag{15}$$

where N_0 is the mean noise variance per receive antenna and normalized covariance $\tilde{\mathbf{R}}_{\mathrm{nn}}$ with $\mathrm{tr}\left(\tilde{\mathbf{R}}_{\mathrm{nn}}\right) = R$. Finally, in the following the signal to noise ratio (SNR) is defined by

$$\gamma = \frac{E_s}{N_0}, \tag{16}$$

which follows the standard SNR definition in the scalar case.

III. MIMO Mutual Information

In this section we introduce a unifying notation of mutual information for an arbitrary number of transmit and receive antennas and for arbitrary correlation properties of the underlying wireless MIMO channel. Furthermore, we calculate a general closed-form expression for the MGF of mutual information by using hypergeometric functions of matrix arguments and certain multivariate integrals. This expression is then reformulated in terms of a determinant of scalar hypergeometric functions. Finally, we specialize the general result to one-side correlated and uncorrelated channels, which requires the calculation of certain multivariate limits.

A. General Expressions

It is well known [1] that the mutual information $I(\mathbf{s},\mathbf{y})$ (in bit per channel use) between transmit vector $\mathbf{s}$ and receive vector $\mathbf{y}$ of the MIMO link in (8) is given by

$$I(\mathbf{s},\mathbf{y}) = \log_2 |\mathbf{I} + \gamma \cdot \tilde{\mathbf{R}}_{\text{ss}} \mathbf{H}^{\text{H}} \tilde{\mathbf{R}}_{\text{nn}}^{-1} \mathbf{H}|. \tag{17}$$

Plugging the channel model with separable correlation matrices according to (10) in (17) we find

$$I(\mathbf{s},\mathbf{y}) = \log_2 \left| \mathbf{I} + \gamma \cdot \tilde{\mathbf{R}}_{\text{ss}} \mathbf{B}^{\text{H}} \mathbf{H}_w \mathbf{A} \tilde{\mathbf{R}}_{\text{nn}}^{-1} \mathbf{A}^{\text{H}} \ \mathbf{H}_w \mathbf{B} \right|. \tag{18}$$

In the following, we reduce (18) to a concise equivalent formulation that allows for a unified analysis of correlated MIMO systems. We can exploit for two matrices $\mathbf{X}$ and $\mathbf{Y}$ of compatible size the relation [31]

$$|\mathbf{I} + \mathbf{XY}| = |\mathbf{I} + \mathbf{YX}|. \tag{19}$$

Furthermore, noticing that the distribution of the i.i.d. complex Gaussian distributed matrix $\mathbf{H}_w$ is invariant to left or right multiplications with unitary matrices $\mathbf{U}$ and $\mathbf{V}$, i.e.

$$\mathbf{U}\mathbf{H}_w\mathbf{V} \simeq \mathbf{H}_w, \tag{20}$$

we find after some simplifications

$$I(\mathbf{s},\mathbf{y}) \simeq \log_2 \left| \mathbf{I} + \gamma \cdot \mathbf{S}\mathbf{H}_w^{\text{H}} \ \mathbf{O}\mathbf{H}_w \right|. \tag{21}$$

Note that for brevity we have introduced the $R \times R$ diagonal matrix of eigenvalues associated to the receive side

$$\mathbf{O} = \text{eig}\left(\tilde{\mathbf{R}}_{\text{nn}}^{-1} \mathbf{R}_{\text{RX}}\right) = \text{diag}\left(o_1, o_2, \ldots, o_R\right), \tag{22}$$

which comprises the effects of receive fading correlation and colored additive Gaussian noise. On the other hand, $\mathbf{S}$ is a diagonal $T \times T$ matrix

$$\mathbf{S} = \text{eig}\left(\tilde{\mathbf{R}}_{\text{ss}} \mathbf{R}_{\text{TX}}\right) = \text{diag}\left(s_1, s_2, \ldots, s_T\right), \tag{23}$$

which takes into account fading correlation at the transmit antenna array and the signal covariance matrix. Note that by exploiting equality (19) we can alternatively formulate (21)

$$I(\mathbf{s},\mathbf{y}) \simeq \log_2 |\mathbf{I} + \gamma \cdot \mathbf{S}\mathbf{H}_w^{\mathrm{H}}\mathbf{O}\mathbf{H}_w| = \log_2 |\mathbf{I} + \gamma \cdot \mathbf{O}\mathbf{H}_w\mathbf{S}\mathbf{H}_w^{\mathrm{H}}|. \tag{24}$$

We rewrite (24) such that the matrix argument $\mathbf{S}\mathbf{H}_w^{\mathrm{H}}\mathbf{O}\mathbf{H}_w$ or $\mathbf{O}\mathbf{H}_w\mathbf{S}\mathbf{H}_w^{\mathrm{H}}$, respectively, of the determinant expression is of full rank, thereby simplifying the subsequent analysis. To this end, we define

$$\mu \equiv \min(R,T) \tag{25}$$

$$\nu \equiv \max(R,T) \tag{26}$$

and the $\mu \times \mu$ diagonal matrix

$$\mathbf{\Sigma} \equiv \begin{cases} \mathbf{S} & R \geq T \\ \mathbf{O} & T > R \end{cases} \tag{27}$$

with elements

$$\mathbf{\Sigma} = \mathrm{diag}\left(\sigma_1, \sigma_2, \ldots, \sigma_\mu\right). \tag{28}$$

Equivalently, we let

$$\mathbf{\Omega} \equiv \begin{cases} \mathbf{O} & R \geq T \\ \mathbf{S} & T > R \end{cases} \tag{29}$$

with elements

$$\mathbf{\Omega} = \mathrm{diag}\left(\omega_1, \omega_2, \ldots, \omega_\nu\right). \tag{30}$$

Furthermore, we introduce a $\nu \times \mu$ matrix of i.i.d. complex Gaussian entries $\mathbf{G}$. With above definitions we can introduce a mathematically convenient and unifying expression for MIMO mutual information, which will serve as a basis for all following derivations

$$I(\mathbf{s},\mathbf{y}) \simeq \log_2 \left|\mathbf{I} + \gamma \cdot \mathbf{\Sigma}\mathbf{G}^{\mathrm{H}}\,\mathbf{\Omega}\mathbf{G}\right|. \tag{31}$$

In addition, it turns out later that it is advantageous in the mathematical derivation to consider MIMO systems of dimension $\nu \times \nu$ first and than generalize the results to arbitrary $\nu \times \mu$ systems. To this end, we introduce an auxiliary $\nu \times \nu$ matrix

$$\tilde{\mathbf{\Sigma}}(\boldsymbol{\varepsilon}) = \mathrm{diag}\left(\sigma_1, \ldots, \sigma_\mu, \boldsymbol{\varepsilon}\right) = \mathrm{diag}\left(\tilde{\sigma}_1, \tilde{\sigma}_2, \ldots, \tilde{\sigma}_\nu\right). \tag{32}$$

with $1 \times (\nu - \mu)$ vector $\boldsymbol{\varepsilon}$

$$\boldsymbol{\varepsilon} = \left\{\varepsilon_0, \ldots, \varepsilon_{\nu-\mu-1}\right\} \tag{33}$$

and an enlarged $\nu \times \nu$ matrix $\tilde{\mathbf{G}}$ of i.i.d. complex Gaussian elements. Without loss of generality we can now reformulate (31) as

$$I(\mathbf{s}, \mathbf{y}) = \lim_{\boldsymbol{\varepsilon} \to \mathbf{0}} \tilde{I}(\mathbf{s}, \mathbf{y}), \tag{34}$$

where $\tilde{I}(\mathbf{s}, \mathbf{y})$ is defined by

$$\tilde{I}(\mathbf{s}, \mathbf{y}) = \log_2 \left|\mathbf{I} + \gamma \cdot \tilde{\mathbf{\Sigma}}(\boldsymbol{\varepsilon}) \tilde{\mathbf{G}}^{\mathrm{H}} \, \mathbf{\Omega} \tilde{\mathbf{G}}\right| = \log_2 \left|\mathbf{I} + \mathbf{Q}\right|. \tag{35}$$

In the following, we will first focus our analysis on the artificial $\nu \times \nu$ system in (35) and than calculate the limit in (34). Note that in (35) we have introduced a so-called $\nu \times \nu$ central complex matrix quadratic form $\mathbf{Q} = \gamma \cdot \tilde{\mathbf{\Sigma}}(\boldsymbol{\varepsilon})^{1/2} \tilde{\mathbf{G}}^{\mathrm{H}} \, \mathbf{\Omega} \tilde{\mathbf{G}} \tilde{\mathbf{\Sigma}}(\boldsymbol{\varepsilon})^{1/2}$ (see [32] and Appendix II).

IV. Moment Generating Function of Mutual Information

The general moment generating function $M_{\tilde{\mathbf{\Sigma}}\mathbf{\Omega}}(s)$ of MIMO mutual information for $\nu \times \nu$ channels with correlation at transmitter and receiver side reads

$$M_{\tilde{\mathbf{\Sigma}}\mathbf{\Omega}}(s) = E_{\tilde{\mathbf{G}}}\left[e^{s \cdot \tilde{I}(\mathbf{s}, \mathbf{y})}\right] = E_{\mathbf{Q}}\left[e^{s \cdot \tilde{I}(\mathbf{s}, \mathbf{y})}\right] = E_{\mathbf{Q}}\left[\left|\mathbf{I}_\nu + \mathbf{Q}\right|^{\tilde{s}}\right], \tag{36}$$

where we have introduced for brevity

$$\tilde{s} = \frac{s}{\ln 2}. \tag{37}$$

The multivariate integral in (36) with respect to the channel statistics can be solved in terms of an hypergeometric function of two matrix arguments.

A. Matrix Representations

Below we present one of the main results of this paper in the following theorem.

Theorem 1: The moment generating function of mutual information of the artificial $\nu \times \nu$ MIMO system according to (35) and (36) is given by

$$M_{\tilde{\mathbf{\Sigma}}\mathbf{\Omega}}(s) = {}_2\tilde{\mathbf{F}}_0^{(\nu,\nu)}\left(-\tilde{s}, \nu;; -\gamma \cdot \tilde{\mathbf{\Sigma}}(\boldsymbol{\varepsilon}), \mathbf{\Omega}\right), \tag{38}$$

where ${}_2\tilde{\mathbf{F}}_0^{(\nu,\nu)}$ is a hypergeometric function of two matrix arguments of equal size [see Appendix I, (106) for a definition and (119) for a scalar representation] and the definitions of $\mathbf{\Omega}$ and $\tilde{\mathbf{\Sigma}}(\boldsymbol{\varepsilon})$ in (29) and (32), respectively.

Proof: We can explicitly calculate the expected value in (36) by noticing

$$M_{\tilde{\mathbf{\Sigma}}\mathbf{\Omega}}(s) = \int {}_1\tilde{\mathbf{F}}_0^{(\nu)}(-\tilde{s};; -\mathbf{Q}) p_{\mathbf{Q}}(\mathbf{Q}) \; \mathrm{D}_c\mathbf{Q}, \tag{39}$$

where we have used (107) in the appendix and the PDF of $\mathbf{Q}$, $p_{\mathbf{Q}}(\mathbf{Q})$. Then applying the multivariate integral (169) yields (38). ■

We note that parts of Theorem 1 were presented in [42].

Starting with Theorem 1, the general MGF can be calculated by a limiting process according to (34)

$$M_{\mathbf{\Sigma}\mathbf{\Omega}}(s) = \lim_{\boldsymbol{\varepsilon} \to \mathbf{0}} M_{\tilde{\mathbf{\Sigma}}\mathbf{\Omega}}(s) \tag{40}$$

and the result of this calculation is given in the following theorem.

Theorem 2: The moment generating function of mutual information of the general $\nu \times \mu$ MIMO system is given by

$$M_{\mathbf{\Sigma}\mathbf{\Omega}}(s) = {}_2\tilde{\mathbf{F}}_0^{(\mu,\nu)}(-\tilde{s}, \nu;; -\gamma \cdot \mathbf{\Sigma}, \mathbf{\Omega}), \tag{41}$$

where the hypergeometric function has two matrix arguments of unequal size [see Appendix I, (137) for a definition as a multivariate limit and Corollary 12 for a scalar representation] and the definitions of $\mathbf{\Omega}$ and $\mathbf{\Sigma}$ in (29) and (27), respectively.

Proof: A detailed derivation of the limiting hypergeometric function of two matrix arguments of unequal size is presented in the proof of Corollary 12 in Appendix I. ■

Calculation of the moment generating function of mutual information, i.e evaluation of the multivariate integral in (36) simplifies in case of vanishing correlation $\mathbf{\Omega} = \mathbf{I}_\nu$ at that side of the MIMO link with the greater number ν of antenna elements.

Corollary 1: The moment generating function of mutual information according to (31) with $\mathbf{\Sigma} \neq \mathbf{I}_\mu, \mathbf{\Omega} = \mathbf{I}_\nu$ is given by

$$M_{\mathbf{\Sigma}}(s) = {}_2\tilde{\mathbf{F}}_0^{(\mu)}\left(-\tilde{s}, \nu;; -\gamma \cdot \mathbf{\Sigma}\right), \tag{42}$$

where the hypergeometric function has one matrix argument [see Appendix I, (101) for a definition and (111) for a scalar representation] and the definition of $\mathbf{\Sigma}$ in (27).

Proof: The corollary can be derived by a limit calculation $\mathbf{\Omega} \to \mathbf{I}_\nu$ from Theorem 2 or we can explicitly calculate the MGF by noticing

$$M_{\mathbf{\Sigma}}(s) = \int {}_1\tilde{\mathbf{F}}_0^{(\mu)}\left(-\tilde{s};; -\mathbf{W}\right) p_{\mathbf{W}}(\mathbf{W}) \ \mathrm{D}_c\mathbf{W} \tag{43}$$

with $\mu \times \mu$ Wishart matrix $\mathbf{W} = \gamma \cdot \mathbf{\Sigma}^{1/2}\mathbf{G}^{\mathrm{H}} \ \mathbf{G}\mathbf{\Sigma}^{1/2}$. Then applying the multivariate integral (175) yields (42). ∎

For completeness, we note that the moment generating function in Theorem 1 reduces to a special hypergeometric function of two matrix arguments of unequal size for the special case $\mathbf{\Sigma} = \mathbf{I}_\mu$.

Corollary 2: The moment generating function of mutual information according to (31) with $\mathbf{\Sigma} = \mathbf{I}_\mu, \mathbf{\Omega} \neq \mathbf{I}_\nu$ is given by

$$M_{\mathbf{\Omega}}(s) = {}_2\tilde{\mathbf{F}}_0^{(\mu,\nu)}\left(-\tilde{s}, \nu;; -\mathbf{I}_\mu, \gamma \cdot \mathbf{\Omega}\right), \tag{44}$$

where the special hypergeometric function has two matrix arguments (one being an identity matrix) of unequal size $\mu \times \mu$ and $\nu \times \nu$ (see Appendix I, Corollary 13) and the definition of $\mathbf{\Omega}$ in (29).

Finally, for an uncorrelated channel the MGF further simplifies and can directly be derived from Corollary 1.

Corollary 3: The moment generating function of mutual information according to (31) for an uncorrelated MIMO channel with $\mathbf{\Sigma} = \mathbf{I}_\mu, \mathbf{\Omega} = \mathbf{I}_\nu$ is given by

$$M_u(s) = {}_2\tilde{\mathbf{F}}_0^{(\mu)}\left(-\tilde{s}, \nu;; -\gamma \cdot \mathbf{I}_\mu\right), \tag{45}$$

where the hypergeometric function has only one scaled identity matrix argument (see Appendix I, Corollary 15).

B. Scalar Representations of General Case

The practical relevance of Theorem 1 with its concise matrix representation of the MGF of mutual information can be established by expressing the hypergeometric function of two matrix arguments in (38) in terms of scalar hypergeometric functions.

We note that if not stated otherwise, we assume that $\mathbf{\Sigma}$ and $\mathbf{\Omega}$, respectively, have pairwise different eigenvalues (with equal eigenvalues, a $\frac{0}{0}$ type of limit has to be calculated).

Theorem 3: The moment generating function of mutual information of the artificial $\nu \times \nu$ MIMO system according to (35) has the scalar representations

$$M_{\tilde{\mathbf{\Sigma}}\mathbf{\Omega}}(s) = \frac{\left|{}_2F_0\left(-\tilde{s}-\nu+1,1;;-\gamma\sigma_i\omega_j\right)\right|}{\prod_{i=1}^{\nu}(\tilde{s}+i-1)^{i-1}\cdot\alpha_\nu\left(\gamma\cdot\tilde{\mathbf{\Sigma}}(\boldsymbol{\varepsilon})\right)\cdot\alpha_\nu(\mathbf{\Omega})} = \tag{46}$$

$$\frac{\left|\frac{1}{\gamma\sigma_i\omega_j}U\left(1,1+\nu+\tilde{s},\frac{1}{\gamma\sigma_i\omega_j}\right)\right|}{\prod_{i=1}^{\nu}(\tilde{s}+i-1)^{i-1}\cdot\alpha_\nu\left(\gamma\cdot\tilde{\mathbf{\Sigma}}(\boldsymbol{\varepsilon})\right)\cdot\alpha_\nu(\mathbf{\Omega})} = \tag{47}$$

$$\frac{\left|\left(\frac{1}{\gamma\sigma_i\omega_j}\right)^{-\tilde{s}}\cdot e^{\frac{1}{\gamma\sigma_i\omega_j}}\cdot\Gamma\left(\nu+\tilde{s},\frac{1}{\gamma\sigma_i\omega_j}\right)\right|}{\prod_{i=1}^{\nu}(\tilde{s}+i-1)^{i-1}\cdot\alpha_\nu\left(\gamma\cdot\tilde{\mathbf{\Sigma}}(\boldsymbol{\varepsilon})\right)\cdot\alpha_\nu(\mathbf{\Omega})}, \tag{48}$$

where i and j run from 1 to ν.

Proof: The theorem follows from direct application of Lemma 2 and its specialization in Corollary 11 in Appendix I. ■

We note that parts of Theorem 3 were given in [43]. The scalar MGF representation in (48) was recently established in [19] via an eigenvalue PDF based approach.

It remains now to establish a scalar expression of the MGF for the general case with $\nu \geq \mu$. To this end, we have to calculate the limit $\lim_{\boldsymbol{\varepsilon}\to\mathbf{0}} M_{\tilde{\mathbf{\Sigma}}\mathbf{\Omega}}(s)$, which is a $0/0$ type of limit that can be solved via application of L'Hospital's rule.

Corollary 4: In the general case with $\nu \geq \mu$ the MGF $M_{\mathbf{\Sigma}\mathbf{\Omega}}(s)$ reads

$$M_{\mathbf{\Sigma\Omega}}(s) = \chi(s) \cdot \frac{\left| \begin{cases} {}_2F_0\left(-\tilde{s}-\nu+1, 1; ; -\gamma\sigma_i\omega_j\right) & i \leq \mu \\ \omega_j^{i-\mu-1} & i > \mu \end{cases} \right|}{|\gamma \cdot \mathbf{\Sigma}|^{\nu-\mu} \cdot \alpha_\mu\left(\gamma \cdot \mathbf{\Sigma}\right) \cdot \alpha_\nu\left(\mathbf{\Omega}\right)}, \tag{49}$$

where i and j run from 1 to ν and

$$\chi(s) = \frac{\prod_{l=1}^{\nu-\mu-1}(\nu+\tilde{s}-j)^{\nu-\mu-j}}{\prod_{i=1}^{\nu}(\tilde{s}+i-1)^{i-1}} \cdot (-1)^{\frac{(\nu-\mu)\cdot(\nu-\mu-1)}{2}}. \tag{50}$$

Again, paralleling Theorem 3, the hypergeometric function in (49) can be rewritten in terms of the Kummer U function or in terms of the incomplete Gamma function. Specifically, we obtain the alternative notation

$$M_{\mathbf{\Sigma\Omega}}(s) = \chi(s) \cdot \frac{|\mathbf{\Psi}_{\mathbf{\Sigma\Omega}}(s)|}{|\gamma \cdot \mathbf{\Sigma}|^{\nu-\mu} \cdot \alpha_\mu\left(\gamma \cdot \mathbf{\Sigma}\right) \cdot \alpha_\nu\left(\mathbf{\Omega}\right)} \tag{51}$$

where we have introduced the $\nu \times \nu$ matrix

$$\mathbf{\Psi}_{\mathbf{\Sigma\Omega}}(s) = \left[\begin{cases} \psi_{ij} & i \leq \mu \\ \omega_j^{i-\mu-1} & i > \mu \end{cases} \right] \tag{52}$$

with elements

$$\psi_{ij} = \sum_{k=\nu-\mu}^{\nu-2} (\gamma\sigma_i\omega_j)^k \cdot [\tilde{s}+\nu-k]_k + (\gamma\sigma_i\omega_j)^{\nu-2} \cdot [\tilde{s}+1]_{\nu-1} \cdot U\left(1, \tilde{s}+2, \frac{1}{\gamma\sigma_i\omega_j}\right). \tag{53}$$

Proof: The first part of the corollary in (49) follows directly from Corollary 12 in the appendix. Equivalently, we can rewrite (49) as

$$M_{\mathbf{\Sigma\Omega}}(s) = \chi(s) \cdot \frac{\left| \begin{cases} \frac{1}{\gamma\sigma_i\omega_j} U\left(1, 1+\nu+\tilde{s}, \frac{1}{\gamma\sigma_i\omega_j}\right) & i \leq \mu \\ \omega_j^{i-\mu-1} & i > \mu \end{cases} \right|}{|\gamma \cdot \mathbf{\Sigma}|^{\nu-\mu} \cdot \alpha_\mu\left(\gamma \cdot \mathbf{\Sigma}\right) \cdot \alpha_\nu\left(\mathbf{\Omega}\right)}. \tag{54}$$

By iteratively applying the recurrence relation for the Kummer U function [33]

$$U(1,b,z) = \frac{1}{z} + \frac{1}{z}(b-2)\cdot U(1,b-1,z) \tag{55}$$

we find the relation

$$\begin{aligned} &\frac{1}{\gamma\sigma_i\omega_j} U\left(1, 1+\nu+\tilde{s}, \frac{1}{\gamma\sigma_i\omega_j}\right) = \\ &\sum_{k=0}^{\nu-2} (\gamma\sigma_i\omega_j)^k \cdot [\tilde{s}+\nu-k]_k + \\ &(\gamma\sigma_i\omega_j)^{\nu-2} \cdot [\tilde{s}+1]_{\nu-1} \cdot U\left(1, \tilde{s}+2, \frac{1}{\gamma\sigma_i\omega_j}\right) \end{aligned} \tag{56}$$

and (51) can be established by subtracting properly scaled rows with index $i > \mu$ from the first μ rows in the determinant. ■

At this point we note that the representation in (51) proves to be useful for finding a concise expression for ergodic MIMO capacity in terms of the exponential integral E_1 (see below).

C. Scalar Representations of Special Cases

In this section, we will specialize the general MGF of mutual information to propagation scenarios with fading correlation at one side of the MIMO link only at one hand and with uncorrelated fading on the other hand. We note that parts of the following results were presented in [44].

Corollary 5: The MGF of MIMO mutual information for a one side correlated channel with $\mathbf{\Sigma} \neq \mathbf{I}_m, \mathbf{\Omega} = \mathbf{I}_n$ is given by

$$M_{\mathbf{\Sigma}}(s) = \frac{\left| U\left(\nu-j+1, \nu-j+2+\tilde{s}, \frac{1}{\gamma\sigma_i}\right)\right|}{|\gamma\cdot\mathbf{\Sigma}|^{\nu-\mu+1}\cdot\alpha_\mu(\gamma\cdot\mathbf{\Sigma})} \tag{57}$$

$$= \frac{\left|\int_0^\infty e^{-t/\sigma_i}\cdot t^{\nu-j}\cdot(1+\gamma t)^{\tilde{s}}\mathrm{d}t\right|}{\prod_{k=1}^{\mu}\Gamma(\nu-k+1)\cdot|\mathbf{\Sigma}|^{\nu-\mu+1}\cdot\alpha_\mu(\mathbf{\Sigma})}, \tag{58}$$

where i and j run from 1 to μ.

Proof: The corollary directly follows from Corollary 1 and the scalar representation of the hypergeometric function of matrix argument in Corollary 14. ■

A different formula results for the case, where that side of the MIMO link with more (ν) antennas is correlated.

Corollary 6: The MGF of MIMO mutual information for a one side correlated channel with $\mathbf{\Sigma} = \mathbf{I}_m, \mathbf{\Omega} \neq \mathbf{I}_n$ is given by

$$M_{\mathbf{\Omega}}(s) = \frac{|\gamma \cdot \mathbf{\Omega}|^{\nu-\mu-1} \cdot \left| \begin{cases} U\left(i, \tilde{s}+i+1, \frac{1}{\gamma\omega_j}\right) & i \leq \mu \\ (\gamma\omega_j)^{-(\nu-i)} & i > \mu \end{cases} \right|}{(-1)^{(\nu-\mu)\cdot\mu} \cdot \alpha_\nu\left(-\gamma\cdot\mathbf{\Omega}\right)} \tag{59}$$

$$= \frac{|\mathbf{\Omega}|^{\nu-\mu-1} \cdot \left| \begin{cases} \int_0^\infty e^{-t/\omega_j} \cdot t^{i-1} \cdot (1+\gamma t)^{\tilde{s}} \mathrm{d}t & i \leq \mu \\ \omega_j^{-(\nu-i)} & i > \mu \end{cases} \right|}{(-1)^{(\nu-\mu)\cdot\mu} \cdot \Gamma_\mu(\mu) \cdot \alpha_\nu\left(-\mathbf{\Omega}\right)}, \tag{60}$$

where i and j run from 1 to ν.

Proof: The corollary directly follows from Corollary 2 and the scalar representation of the hypergeometric function of matrix argument in Corollary 13. ■

Finally, we present results for an uncorrelated channel, which were first established in [12].

Corollary 7: The MGF of MIMO mutual information for an uncorrelated channel with identity covariances $\mathbf{\Sigma} = \mathbf{I}_m, \mathbf{\Omega} = \mathbf{I}_n$ is given by

$$M_u(s) = \frac{\left|\int_0^\infty e^{-t} \cdot t^{\nu-\mu+j+i-2} \cdot (1+\gamma t)^{\tilde{s}} \mathrm{d}t\right|}{\Gamma_\mu(\mu) \cdot \prod_{k=1}^{\mu} \Gamma(\nu-k+1)}, \tag{61}$$

where i and j run from 1 to μ.

Proof: The corollary directly follows from Corollary 3 and the scalar representation of the hypergeometric function of matrix argument in Corollary 13. ■

V. Ergodic Capacity

Based on the moment generating functions derived in the last section, ergodic MIMO capacity (with uninformed transmitter) can simply be derived by the relation

$$C_{\mathbf{\Sigma\Omega}}(\gamma) = E\left[I(\mathbf{s},\mathbf{y})\right] = \frac{\partial}{\partial s} M_{\mathbf{\Sigma\Omega}}(s)\bigg|_{s=0} = \frac{1}{\ln 2} \cdot \frac{\partial}{\partial \tilde{s}} M_{\mathbf{\Sigma\Omega}}(s)\bigg|_{\tilde{s}=0}. \tag{62}$$

We note that higher order moments of mutual information (that could be used e.g. for calculating approximations of the distribution) can be obtained by higher order derivatives.

Again, we have to differentiate between various propagation scenarios with varying correlation properties.

A. General Case

Using (62), we can derive one of the main results of this paper.

Theorem 4: The ergodic capacity of a fully correlated MIMO channel is given by

$$C_{\mathbf{\Sigma}\mathbf{\Omega}}(\gamma) = \chi(0) \cdot \frac{1}{\ln 2 \cdot |\gamma \cdot \mathbf{\Sigma}|^{\nu-\mu} \cdot \alpha_\mu(\gamma \cdot \mathbf{\Sigma}) \cdot \alpha_\nu(\mathbf{\Omega})} \cdot \sum_{l=1}^{\mu} |\Xi_{\mathbf{\Sigma}\mathbf{\Omega}}(l)| \tag{63}$$

with the auxiliary $\nu \times \nu$ matrices

$$\Xi_{\mathbf{\Sigma}\mathbf{\Omega}}(l) = \left[\begin{cases} (\gamma\sigma_i\omega_j)^{\nu-1} \cdot \Gamma(\nu) \cdot e^{\frac{1}{\gamma\sigma_i\omega_j}} \cdot E_1\left(\frac{1}{\gamma\sigma_i\omega_j}\right) & i = l \leq \mu \\ \sum_{k=\nu-\mu}^{\nu-1} (\gamma\sigma_i\omega_j)^k \cdot [\nu-k]_k & i \neq l \leq \mu \\ \omega_j^{i-\mu-1} & i > \mu \end{cases} \right], \tag{64}$$

where i and j run from 1 to ν. The eigenvalues of $\mathbf{\Sigma}$ and $\mathbf{\Omega}$, respectively, have to be pairwise different. This requirement is typically fulfilled in practical applications. Otherwise, a suitable limit has to be calculated.

Proof: We start the derivation with the MGF expresssion in (51) and find

$$\begin{aligned} \frac{\partial}{\partial \tilde{s}} M_{\mathbf{\Sigma}\mathbf{\Omega}}(s) = & \frac{1}{|\gamma \cdot \mathbf{\Sigma}|^{\nu-\mu} \cdot \alpha_\mu(\gamma \cdot \mathbf{\Sigma}) \cdot \alpha_\nu(\mathbf{\Omega})} \times \\ & \left[\frac{\partial}{\partial \tilde{s}} \chi(s) \cdot |\Psi_{\mathbf{\Sigma}\mathbf{\Omega}}(s)| + \chi(s) \cdot \frac{\partial}{\partial \tilde{s}} |\Psi_{\mathbf{\Sigma}\mathbf{\Omega}}(s)| \right]. \end{aligned} \tag{65}$$

The derivative of $\chi(s)$ is given by

$$\frac{\partial}{\partial \tilde{s}} \chi(s) = \chi(s) \cdot \lambda, \tag{66}$$

where we have introduced

$$\lambda = \left(\sum_{j=1}^{\nu-\mu-1} \frac{\nu-\mu-j}{\nu-j} - \nu + 1 \right) \tag{67}$$

for brevity. We obtain

$$\frac{\partial}{\partial \tilde{s}} M_{\mathbf{\Sigma\Omega}}(s) = \frac{\chi(s)}{|\gamma \cdot \mathbf{\Sigma}|^{\nu-\mu} \cdot \alpha_\mu(\gamma \cdot \mathbf{\Sigma}) \cdot \alpha_\nu(\mathbf{\Omega})} \times \left[\lambda \cdot |\mathbf{\Psi}_{\mathbf{\Sigma\Omega}}(s)| + \frac{\partial}{\partial \tilde{s}} |\mathbf{\Psi}_{\mathbf{\Sigma\Omega}}(s)|\right] \quad . \tag{68}$$

For finding the derivative of $\frac{\partial}{\partial s} |\mathbf{\Psi}_{\mathbf{\Sigma\Omega}}(s)|$, we can make use of the general formula

$$\frac{\partial}{\partial s} |\mathbf{X}(s)| = \sum_i |\mathbf{X}_i(s)|, \tag{69}$$

where $\mathbf{X}_i(s)$ denotes the matrix $\mathbf{X}$, whereas its ith row (or alternatively column) is differentiated with respect to s. Therefore, we first focus on the derivative of the elements of $\mathbf{\Psi}_{\mathbf{\Sigma\Omega}}(s)$ in (53). With the integral representation of the remaining Kummer U function

$$U\left(1, \tilde{s}+2, \frac{1}{\gamma\sigma_i\omega_j}\right) = \int_0^\infty e^{-\frac{t}{\gamma\sigma_i\omega_j}} \cdot (1+t)^{\tilde{s}} \mathrm{d}t \tag{70}$$

we get [34]

$$U\left(1, \tilde{s}+2, \frac{1}{\gamma\sigma_i\omega_j}\right)\bigg|_{\tilde{s}=0} = \gamma\sigma_i\omega_j \tag{71}$$

and after exchanging the sequence of integration and differentiation (a justification is omitted here) we find [34]

$$\begin{aligned} \frac{\partial}{\partial \tilde{s}} U\left(1, \tilde{s}+2, \frac{1}{\gamma\sigma_i\omega_j}\right)\bigg|_{\tilde{s}=0} &= \int_0^\infty e^{-\frac{t}{\gamma\sigma_i\omega_j}} \cdot \ln(1+t) \mathrm{d}t \\ &= \gamma\sigma_i\omega_j \cdot e^{\frac{1}{\gamma\sigma_i\omega_j}} \cdot E_1\left(\frac{1}{\gamma\sigma_i\omega_j}\right). \end{aligned} \tag{72}$$

Now using (71) in (53) we obtain

$$\psi_{ij}\big|_{\tilde{s}=0} = \sum_{k=0}^{\nu-1} (\gamma\sigma_i\omega_j)^k \cdot [\nu-k]_k \tag{73}$$

and from (52)

$$\boldsymbol{\Psi}_{\boldsymbol{\Sigma\Omega}}(s)|_{\tilde{s}=0} = \left[\begin{cases} \sum_{k=0}^{\nu-1} (\gamma\sigma_i\omega_j)^k \cdot [\nu-k]_k & i \le \mu \\ \omega_j^{i-\mu-1} & i > \mu \end{cases} \right]. \tag{74}$$

Then we introduce the auxiliary function for brevity

$$\xi_k(\tilde{s}) = [\tilde{s}+\nu-k]_k = \prod_{i=1}^{k} (\tilde{s}+\nu-i) \tag{75}$$

with derivative

$$\frac{\partial}{\partial \tilde{s}} \xi_k(\tilde{s}) = \xi_k(\tilde{s}) \cdot \sum_{i=1}^{k} \frac{1}{\tilde{s}+\nu-i} = \xi_k(\tilde{s}) \cdot \kappa_k(\tilde{s}+\nu). \tag{76}$$

and the short-hand notation

$$\kappa_k(x) = \sum_{i=1}^{k} \frac{1}{x-i}. \tag{77}$$

We obtain after some simplifications from (53)

$$\begin{aligned} \frac{\partial}{\partial \tilde{s}} \psi_{ij} \bigg|_{\tilde{s}=0} = \sum_{k=0}^{\nu-1} & (\gamma\sigma_i\omega_j)^k \cdot \xi_k(0) \cdot \kappa_k(\nu) + \\ & (\gamma\sigma_i\omega_j)^{\nu-1} \cdot \xi_{\nu-1}(0) \cdot e^{\frac{1}{\gamma\sigma_i\omega_j}} \cdot E_1\left(\frac{1}{\gamma\sigma_i\omega_j}\right), \end{aligned} \tag{78}$$

which can be written as

$$\begin{aligned} \frac{\partial}{\partial \tilde{s}} \psi_{ij} \bigg|_{\tilde{s}=0} = \sum_{k=0}^{\nu-1} & (\gamma\sigma_i\omega_j)^k \cdot [\nu-k]_k \cdot \kappa_k(\nu) + \\ & (\gamma\sigma_i\omega_j)^{\nu-1} \cdot \Gamma(\nu) \cdot e^{\frac{1}{\gamma\sigma_i\omega_j}} \cdot E_1\left(\frac{1}{\gamma\sigma_i\omega_j}\right). \end{aligned} \tag{79}$$

By applying the differentiation formula (69) to (52) we can derive with the help of (73) and (79) after splitting the determinants

$$\frac{\partial}{\partial \tilde{s}} \left|\boldsymbol{\Psi}_{\boldsymbol{\Sigma\Omega}}(s)\right| \bigg|_{\tilde{s}=0} = \sum_{l=1}^{\mu} |\boldsymbol{\Upsilon}| + \sum_{l=1}^{\mu} |\boldsymbol{\Xi}_{\boldsymbol{\Sigma\Omega}}(l)| \tag{80}$$

with the auxiliary $\nu \times \nu$ matrix

$$\boldsymbol{\Upsilon} = \left[\begin{cases} \sum_{k=0}^{\nu-1} (\gamma\sigma_i\omega_j)^k \cdot \kappa_k(\nu) \cdot [\nu-k]_k & i \leq \mu \\ \omega_j^{i-\mu-1} & i > \mu \end{cases} \right] \tag{81}$$

and the definition of $\Xi_{\boldsymbol{\Sigma\Omega}}(l)$ in (64). Note that by subtracting scaled rows with index $i > \mu$ from the first μ rows of Υ, which does not alter the determinant, we have

$$|\boldsymbol{\Upsilon}| = \left| \begin{cases} \sum_{k=\nu-\mu}^{\nu-1} (\gamma\sigma_i\omega_j)^k \cdot \kappa_k(\nu) \cdot [\nu-k]_k & i \leq \mu \\ \omega_j^{i-\mu-1} & i > \mu \end{cases} \right|. \tag{82}$$

By proper decomposition of the determinants it can be shown that

$$\sum_{l=1}^{\mu} |\boldsymbol{\Upsilon}| = \sum_{k=\nu-\mu}^{\nu-1} \kappa_k(\nu) \cdot |\Psi_{\boldsymbol{\Sigma\Omega}}(s)||_{\tilde{s}=0} \,. \tag{83}$$

For further simplification of (83) note the general formula

$$\sum_{k=0}^{x} \kappa_k(y) = \sum_{k=0}^{x} \sum_{i=1}^{k} \frac{1}{y-i} = \sum_{k=1}^{x} \frac{x-k+1}{y-k} \tag{84}$$

such that

$$\sum_{k=0}^{\nu-1} \kappa_k(\nu) = \sum_{k=0}^{\nu-1} \sum_{i=1}^{k} \frac{1}{\nu-i} = \nu - 1 \tag{85}$$

and finally

$$\sum_{k=\nu-\mu}^{\nu-1} \kappa_k(\nu) = \nu - 1 - \sum_{k=1}^{\nu-\mu-1} \frac{\nu-\mu-k}{\nu-k} = -\lambda \tag{86}$$

with the definition of λ in (67). Plugging all auxiliary results in (80) we arrive at

$$\frac{\partial}{\partial \tilde{s}} \left|\Psi_{\boldsymbol{\Sigma\Omega}}(s)\right| \Big|_{\tilde{s}=0} = -\lambda \cdot |\Psi_{\boldsymbol{\Sigma\Omega}}(s)||_{\tilde{s}=0} + \sum_{l=1}^{\mu} |\Xi_{\boldsymbol{\Sigma\Omega}}(l)| \,. \tag{87}$$

Now inserting (87) in (68) we can establish (63). ■

B. Special Cases

Similar to the proceeding in the last section, we can calculate the ergodic capacity for channels with one sided fading correlation and uncorrelated fading by differentiating the MGF of mutual information. In the following capacity expressions, we make use of the integral

$$\frac{1}{\Gamma(n)}\int_0^\infty e^{-t/c}\cdot t^{n-1}\cdot(1+at)^{\tilde{s}}\mathrm{d}t\Big|_{\tilde{s}=0}=c^n \tag{88}$$

and via integration by parts one can derive

$$\mathscr{I}_1(c,a,n)=\frac{1}{\Gamma(n)}\cdot\int_0^\infty e^{-t/c}\cdot t^{n-1}\cdot\ln(1+at)\mathrm{d}t=a^{-n}\cdot e^{\frac{1}{ac}}\cdot\sum_{k=1}^{n}(ac)^k\cdot\Gamma(-n+k,\frac{1}{ac}) \tag{89}$$

with the special case of the incomplete Gamma function for positive integers m (see [33, equation 6.5.19])

$$\Gamma(-m,1/x)=\frac{(-1)^m}{\Gamma(m+1)}\cdot\left[E_1\left(1/x\right)-e^{-1/x}\cdot x\cdot\sum_{j=0}^{m-1}(-x)^j\cdot\Gamma(j+1)\right]. \tag{90}$$

Using (89), we get

$$\frac{1}{\Gamma(n)}\cdot\frac{\partial}{\partial\tilde{s}}\int_0^\infty e^{-t/c}\cdot t^{n-1}\cdot(1+at)^{\tilde{s}}\mathrm{d}t\Big|_{\tilde{s}=0}=\mathscr{I}_1(c,a,n). \tag{91}$$

We start with one side correlated channels.

Corollary 8: The ergodic capacity of a MIMO channel with $\mathbf{\Sigma}\neq\mathbf{I}_m,\mathbf{\Omega}=\mathbf{I}_n$ is given by

$$C_{\mathbf{\Sigma}}(\gamma)=\frac{1}{\ln 2\cdot|\mathbf{\Sigma}|^{\nu-\mu+1}\cdot\alpha_\mu\left(\mathbf{\Sigma}\right)}\cdot\sum_{l=1}^{\mu}|\Xi_{\mathbf{\Sigma}}(l)| \tag{92}$$

with the auxiliary $\mu\times\mu$ matrices

$$\Xi_{\mathbf{\Sigma}}(l)=\left[\begin{cases}\sigma_i^{\nu-j+1} & i\neq l\\ \mathscr{I}_1(\sigma_i,\gamma,\nu-j+1) & i=l\end{cases}\right] \tag{93}$$

where i and j run from 1 to μ.

Proof: We can start with the scalar representation of the MGF in Corollary 5. For differentiating the MGF we can again use (69). The resulting integrals are given in (88) and (89). ∎

Without proof we state the following corollary. The derivation parallels the derivation of Corollary 8.

Corollary 9: The ergodic capacity of a channel with $\mathbf{\Sigma} = \mathbf{I}_m, \mathbf{\Omega} \neq \mathbf{I}_n$ is given by

$$C_{\mathbf{\Omega}}(\gamma) = \frac{|\mathbf{\Omega}|^{\nu-\mu-1}}{\ln 2 \cdot (-1)^{(\nu-\mu)\cdot\mu} \cdot \alpha_\nu(-\mathbf{\Omega})} \cdot \sum_{l=1}^{\mu} |\Xi_{\mathbf{\Omega}}(l)| \tag{94}$$

with auxiliary $\nu \times \nu$ matrices

$$\Xi_{\mathbf{\Omega}}(l) = \left[\begin{cases} \omega_j^i & i \leq \mu, i \neq l \\ \mathscr{I}_1(\omega_j, \gamma, i) & i \leq \mu, i = l \\ \omega_j^{-(\nu-i)} & i > \mu \end{cases} \right], \tag{95}$$

where i and j run from 1 to ν.

Finally, we present an expression for the ergodic capacity of an uncorrelated Rayleigh fading channel, which was first derived in [1], were it was given in terms of a scalar integral expression.

Corollary 10: The ergodic capacity of a channel with $\mathbf{\Sigma} = \mathbf{I}_m, \mathbf{\Omega} = \mathbf{I}_n$ is given by

$$C_u(\gamma) = \frac{1}{\Gamma_\mu(\mu) \cdot \prod_{k=1}^{\mu} \Gamma(\nu - k + 1)} \cdot \sum_{l=1}^{\mu} |\Xi_u(l)| \tag{96}$$

with auxiliary $\mu \times \mu$ matrices

$$\Xi_u(l) = \left[\begin{cases} \Gamma(\nu - \mu + i + j - 1) & i \neq l \\ \Gamma(\nu - \mu + i + j - 1) \cdot \mathscr{I}_1(1, \gamma, \nu - \mu + i + j - 1) & i = l \end{cases} \right], \tag{97}$$

where i and j run from 1 to μ.

VI. Numerical Results

In this section we study systems with white input signals of power E_s (this is the capacity achieving strategy with uninformed transmitter [33]) and additive white Gaussian noise with variance N_0 per receive antenna

$$\mathbf{R}_{\text{ss}} = E_s \cdot \mathbf{I}_T$$
$$\mathbf{R}_{\text{nn}} = N_0 \cdot \mathbf{I}_R. \tag{98}$$

Furthermore, due to their simple structure, in the following we consider exponential correlation matrices [35] at the transmitter and the receiver with

$$\mathbf{R}_{\text{RX}} = \left[r_{RX}^{|i-j|} \right]$$
$$\mathbf{R}_{\text{TX}} = \left[r_{TX}^{|i-j|} \right]. \tag{99}$$

The correlation coefficient at the receiver (transmitter) r_{RX} (r_{TX}) ranges from 0 to 1 and models the correlation between two neighboring receive (transmit) antennas. With the given channel model, correlation between two antenna elements decreases exponentially with their distance. Finally, the SNR in dB is defined by

$$\gamma_{dB} \equiv 10 \cdot \log_{10} \frac{\rho \cdot E_s}{N_0} = 10 \cdot \log_{10}(\rho \cdot \gamma) \quad [\text{dB}], \tag{100}$$

where ρ is the transmit power constraint and we assume in the following numerical results $\rho = T$ in accordance with (98).

Simulation results and theoretical curves according to Theorem 4 closely agree in Fig. 1 for a 4×4 system with correlated MIMO channel. In this particular scenario, the correlation coefficient is the same in each case at the transmitter and the receiver side with $r_{TX} = r_{RX} = r$. As expected, the negative impact of channel correlation on ergodic MIMO capacity with uninformed transmitter can be observed.

In Fig. 2 we consider a receive diversity system with $T = 1$ transmit antenna and $R = 8$ receive antennas. The capacity of this single input multiple output (SIMO) system is very robust with respect to fading correlation at the receive antenna array.

A different behavior can be observed in Fig. 3 for a 2×8 MIMO system. For correlation coefficients greater then approximately 0.7 at transmitter or receiver side, respectively, there is a significant degradation in ergodic capacity.

VII. Conclusion

By using a special hypergeometric function of matrix argument, in this paper we have for the first time presented a concise closed form expression for the moment generating function of mutual information for Rayleigh fading MIMO channels with arbitrary fading correlation at the transmitter and the receiver, respectively. Specifically, in contrast to existing eigenvalue PDF based approaches, we have applied a direct multivariate integration technique. Based on the moment generating function, we have derived exact expressions for ergodic capacity, which require the evaluation of the exponential integral E_1 only. Furthermore, based on the general result, we have demonstrated a unifying derivation of known moment generating functions for special propagation scenarios with one-sided fading correlation and uncorrelated fading.

Appendix I
Hypergeometric Functions of Matrix Argument

In this appendix we provide a brief introduction to hypergeometric functions of matrix arguments. We present their basic definition and analyze some special cases, which are useful for the analysis of MIMO mutual information and capacity, in more detail. In particular, we calculate explicit closed-form scalar representations.

A. General Definition

Paralleling the structure of scalar hypergeometric functions, the complex hypergeometric function of one $n \times n$ matrix argument $\mathbf{X}$ is defined by (see e.g. [2] [11]) the infinite sum of zonal polynomials

$$ {}_p\tilde{\mathbf{F}}_q^{(n)}\left(a_1,\ldots,a_p;b_1,\ldots,b_q;\mathbf{X}\right) \equiv \sum_{k=0}^{\infty}\sum_{\kappa}\frac{[a_1]_\kappa\cdot\ldots\cdot[a_p]_\kappa}{[b_1]_\kappa\cdot\ldots\cdot[b_q]_\kappa}\cdot\frac{\tilde{C}_\kappa(\mathbf{X})}{k!} \tag{101} $$

with parameters $a_1,\ldots,a_p$ and $b_1,\ldots,b_q$. In (101),

$$ \kappa = (k_1,\ldots,k_n) \tag{102} $$

is a partition of k into not more than n parts with

$$k_1 \geq k_2 \geq \ldots \geq k_n \geq 0 \tag{103}$$

and

$$k_1 + k_2 + \ldots + k_n = k. \tag{104}$$

The expression $\tilde{C}_\kappa(\mathbf{X})$ denotes a zonal polynomial and is a symmetric polynomial of degree k in the eigenvalues of matrix $\mathbf{X}$. Finally, the complex multivariate hypergeometric coefficient is defined by

$$[a]_\kappa \equiv \prod_{i=1}^{n} [a-i+1]_{k_i}. \tag{105}$$

Paralleling (101), the complex hypergeometric function of two $n \times n$ matrix arguments $\mathbf{X}$ and $\mathbf{Y}$ is defined by

$$\begin{aligned} {}_p\tilde{\mathbf{F}}_q^{(n,n)}\left(a_1,\ldots,a_p;b_1,\ldots,b_q;\mathbf{X},\mathbf{Y}\right) \equiv \\ \sum_{k=0}^{\infty}\sum_{\kappa} \frac{[a_1]_\kappa \cdot \ldots \cdot [a_p]_\kappa}{[b_1]_\kappa \cdot \ldots \cdot [b_q]_\kappa} \cdot \frac{\tilde{C}_\kappa(\mathbf{X}) \cdot \tilde{C}_\kappa(\mathbf{Y})}{\tilde{C}_\kappa(\mathbf{I}_n) \cdot k!} \quad . \end{aligned} \tag{106}$$

B. Special Case

A key formula (see e.g. [2]) for the analysis of MIMO mutual information and capacity is the following determinant representation with $n \times n$ matrix $\mathbf{X}$

$${}_1\tilde{\mathbf{F}}_0^{(n)}(a;;\mathbf{X}) = |\mathbf{I}_n - \mathbf{X}|^{-a}, \tag{107}$$

which is the matrix analogue to the well-known scalar binomial series [33]

$${}_1F_0(a;;x) = (1-x)^{-a}. \tag{108}$$

C. General Scalar Representations

The practical relevance of hypergeometric functions of matrix argument, which are essentially an infinite sum of zonal polynomials (there are no efficient formulas available for their calculation) can be established by the following lemmas.

Lemma 1: ([38, equation (34)] Let $\mathbf{X}$ be a $n \times n$ matrix with

$$\mathrm{eig}\left(\mathbf{X}\right) = \mathrm{diag}\left(x_1, \ldots, x_n\right) \tag{109}$$

and assume w.l.o.g. that

$$x_1 \geq x_2 \geq \ldots \geq x_n \quad . \tag{110}$$

Then the hypergeometric function of one matrix argument can be expressed in terms of scalar hypergeometric functions

$$ {}_p\tilde{\mathbf{F}}_q^{(n)}\left(a_1, \ldots, a_p; b_1, \ldots, b_q; \mathbf{X}\right) = \frac{|\mathbf{D}|}{\alpha_n(\mathrm{eig}(\mathbf{X}))}, \tag{111}$$

whereas the $n \times n$ matrix $\mathbf{D}$ is given by

$$\mathbf{D} = \left[x_i^{n-j} \cdot {}_pF_q\left(\hat{\mathbf{a}}_j; \hat{\mathbf{b}}_j; x_i\right)\right], \tag{112}$$

where row index i and column index j run from 1 to n and

$$\begin{aligned} \hat{\mathbf{a}}_j &= \left\{a_1 - j + 1, \ldots, a_p - j + 1\right\} \\ \hat{\mathbf{b}}_j &= \left\{b_1 - j + 1, \ldots, b_q - j + 1\right\}. \end{aligned} \tag{113}$$

Proof: See [38] and the references therein for details. The lemma can also be derived by a limiting process from Lemma (2). ■

The scalar representation becomes more involved in case of two matrix arguments.

Lemma 2: ([36, Lemma 3] or [37]) Let $\mathbf{X}$ and $\mathbf{Y}$ be two $n \times n$ matrices with

$$\mathrm{eig}(\mathbf{X}) = \mathrm{diag}(x_1, \ldots, x_n) \tag{114}$$

$$\mathrm{eig}(\mathbf{Y}) = \mathrm{diag}(y_1, \ldots, y_n) \tag{115}$$

and assume w.l.o.g. that

$$x_1 \geq x_2 \geq \ldots \geq x_n \tag{116}$$

$$y_1 \geq y_2 \geq \ldots \geq y_n \quad . \tag{117}$$

Moreover, introduce the auxiliary function with $\mathbf{b} = \{b_1, \ldots, b_q\}$

$$\psi_q^{(n)}(\mathbf{b}) = \prod_{i=1}^{n} \prod_{j=1}^{q} (b_j - i + 1)^{i-1}. \tag{118}$$

Then the hypergeometric function of two matrix arguments can be expressed in terms of scalar hypergeometric functions

$${}_p\tilde{\mathbf{F}}_q^{(n,n)}\left(a_1, \ldots, a_p; b_1, \ldots, b_q; \mathbf{X}, \mathbf{Y}\right) = \frac{\Gamma_n(n) \cdot \psi_q^{(n)}(\mathbf{b}) \cdot |\mathbf{D}|}{\alpha_n(\mathrm{eig}(\mathbf{X})) \cdot \alpha_n(\mathrm{eig}(\mathbf{Y})) \cdot \psi_p^{(n)}(\mathbf{a})}, \tag{119}$$

whereas the $n \times n$ matrix $\mathbf{D}$ is given by

$$\mathbf{D} = \left[{}_pF_q\left(\hat{\mathbf{a}}; \hat{\mathbf{b}}; x_i y_j\right)\right], \tag{120}$$

where row index i and column index j run from 1 to n and

$$\hat{\mathbf{a}} = \left\{a_1 - n + 1, \ldots, a_p - n + 1\right\}$$

$$\hat{\mathbf{b}} = \left\{b_1 - n + 1, \ldots, b_q - n + 1\right\}. \tag{121}$$

Proof: See [36] and [37] for details. ■

D. Special Scalar Representations

In the analysis of the MGF of MIMO mutual information we encounter the special hypergeometric function ${}_2\tilde{\mathbf{F}}_0^{(n,n)}(a_1, n;; -\mathbf{X}, \mathbf{Y})$. Due to its importance, we provide several concise scalar representations.

Corollary 11: The hypergeometric function of two $n \times n$ matrix arguments $\mathbf{X}$ and $\mathbf{Y}$ and parameter a_1 has the scalar representations

$$
\begin{aligned}
&{}_2\tilde{\mathbf{F}}_0^{(n,n)}(a_1, n;; -\mathbf{X}, \mathbf{Y}) = \\
&\frac{\left|{}_2F_0\left(a_1 - n + 1, 1;; -x_i y_j\right)\right|}{\prod_{i=1}^{n}(-a_1 + i - 1)^{i-1} \cdot \alpha_n(\mathrm{eig}(\mathbf{X})) \cdot \alpha_n(\mathrm{eig}(\mathbf{Y}))} = \qquad (122)\\
&\frac{|\mathbf{X}|^{-1} \cdot |\mathbf{Y}|^{-1} \cdot \left|U\left(1, 1 + n - a_1, \frac{1}{x_i y_j}\right)\right|}{\prod_{i=1}^{n}(-a_1 + i - 1)^{i-1} \cdot \alpha_n(\mathrm{eig}(\mathbf{X})) \cdot \alpha_n(\mathrm{eig}(\mathbf{Y}))} = \qquad (123)\\
&\frac{\left|\mathbf{X}^{-1}\right|^{a_1} \cdot \left|\mathbf{Y}^{-1}\right|^{a_1} \cdot \left|e^{\frac{1}{x_i y_j}} \cdot \Gamma(n - a_1, \frac{1}{x_i y_j})\right|}{\prod_{i=1}^{n}(-a_1 + i - 1)^{i-1} \cdot \alpha_n\left(\mathrm{eig}(\mathbf{X})^{-1}\right) \cdot \alpha_n\left(\mathrm{eig}(\mathbf{Y})^{-1}\right)}, \qquad (124)
\end{aligned}
$$

where i and j run from 1 to n.

Proof: From Lemma 2 we get

$$
\begin{aligned}
&{}_2\tilde{\mathbf{F}}_0^{(n,n)}(a_1, n;; -\mathbf{X}, \mathbf{Y}) = \\
&\Gamma_n(n) \cdot \frac{\left|{}_2F_0\left(a_1 - n + 1, 1;; -x_i y_j\right)\right|}{\alpha_n(\mathrm{eig}(-\mathbf{X})) \cdot \alpha_n(\mathrm{eig}(\mathbf{Y})) \cdot \psi_2^{(n)}(a_1, n)}. \qquad (125)
\end{aligned}
$$

Then note that the auxiliary function $\psi_2^{(n)}(a_1, n)$ can be rewritten as

$$
\begin{aligned}
\psi_2^{(n)}(a_1, n) &= \prod_{i=1}^{n}\left[(a_1 - i + 1)^{i-1}(n - i + 1)^{i-1}\right] \\
&= \Gamma_n(n) \cdot \prod_{k=0}^{n-1}[a_1 - n + 1]_k \\
&= (-1)^{\frac{n \cdot (n-1)}{2}} \cdot \Gamma_n(n) \cdot \prod_{i=1}^{n}(-a_1 + i - 1)^{i-1} \qquad (126)
\end{aligned}
$$

and for constant a we have

$$\alpha_n\left(\mathrm{eig}\left(a\cdot\mathbf{X}\right)\right)=a^{\frac{n\cdot(n-1)}{2}}\cdot\alpha_n\left(\mathrm{eig}\left(\mathbf{X}\right)\right). \tag{127}$$

Plugging (126) in (125) and using (127) establishes the first part of Corollary 11 given in (122). Furthermore, we have the notation of the Kummer U function in terms of a hypergeometric function [33]

$$U\left(a,b,z\right)=z^{-a}\cdot{}_2F_0\left(a,1+a-b;;-1/z\right) \tag{128}$$

and the Kummer transformation formula

$$U\left(a,b,z\right)=z^{1-b}\cdot U\left(1+a-b,2-b,z\right). \tag{129}$$

Combining (128) and (129) we find

$${}_2F_0\left(a,1+a-b;;-1/z\right)=z^{1+a-b}\cdot U\left(1+a-b,2-b,z\right). \tag{130}$$

This can be reformulated and we arrive at

$${}_2F_0\left(a,x;;-z\right)=z^{-x}\cdot U\left(x,1-a+x,1/z\right). \tag{131}$$

Using relation (131) in (122) directly yields (123). We emphasize at this point that in contrast to the infinite power series representation of the scalar hypergeometric function (4), the Kummer U function has a mathematically better tractable integral representation (3) with the special case (see also [34])

$$U\left(1,b,z\right)=\frac{e^z\cdot\Gamma(b-1,z)}{z^{b-1}}. \tag{132}$$

With the help of (132) we find from (131)

$${}_2F_0\left(a,1;;-z\right)=z^{-1}\cdot U\left(1,2-a,1/z\right)=z^a\cdot e^{1/z}\cdot\Gamma(1-a,1/z). \tag{133}$$

Now using the Vandermonde determinant formula

$$\alpha_n\left(\mathrm{eig}\left(\mathbf{X}\right)\right)=\left(-1\right)^{\frac{n\cdot(n-1)}{2}}\cdot\alpha_n\left(\mathrm{eig}\left(\mathbf{X}\right)^{-1}\right)\cdot\left|\mathbf{X}^{-1}\right|^{-(n-1)}, \tag{134}$$

which can be obtained by reformulating (5), and (133) in (122), we can directly find the last part of Corollary 11 as given in (124). ■

We now generalize Corollary 11 to allow for two matrix arguments of unequal size. Basically, we apply a limiting process for establishing the general result. For brevity, we assume diagonal matrix arguments (a generalization is straightforward).

Corollary 12: The hypergeometric function of two matrix arguments ${}_2\tilde{\mathbf{F}}_0^{(m,n)}(a_1, n;; -\mathbf{X}_m, \mathbf{Y}_n)$ with diagonal $m \times m$ matrix $\mathbf{X}_m$ and diagonal $n \times n$ ($n \geq m$) matrix $\mathbf{Y}_n$ has the scalar representation (i and j run from 1 to n)

$$
\begin{aligned}
&{}_2\tilde{\mathbf{F}}_0^{(m,n)}(a_1, n;; -\mathbf{X}_m, \mathbf{Y}_n) = \\
&\xi(a_1) \cdot \frac{\left| \begin{cases} {}_2F_0\left(a_1 - n + 1, 1;; -x_i y_j\right) & i \leq m \\ y_j^{i-m-1} & i > m \end{cases} \right|}{|\mathbf{X}_m|^{n-m} \cdot \alpha_m(\mathbf{X}_m) \cdot \alpha_n(\mathbf{Y}_n)}
\end{aligned} \tag{135}
$$

and

$$
\begin{aligned}
\xi(a_1) &= \frac{\prod_{l=1}^{n-m-1}(n - a_1 - l)^{n-m-l}}{\prod_{i=1}^{n}(-a_1 + i - 1)^{i-1}} \cdot (-1)^{\frac{(n-m)\cdot(n-m-1)}{2}} \\
&= \frac{1}{\prod_{k=n-m}^{n-1}[a_1 - n + 1]_k} \cdot (-1)^{\frac{n\cdot(n-1)}{2}}
\end{aligned} \tag{136}
$$

is a function of the parameter a_1 that has been introduced for notational convenience.

Proof: The hypergeometric function of two matrix arguments of unequal size shall be defined by the limit (using Corollary 11)

$$
\begin{aligned}
&{}_2\tilde{\mathbf{F}}_0^{(m,n)}(a_1, n;; -\mathbf{X}_m, \mathbf{Y}_n) \equiv \\
&\lim_{\boldsymbol{\varepsilon}\to\mathbf{0}} {}_2\tilde{\mathbf{F}}_0^{(n,n)}(a_1, n;; -\mathbf{X}_n(\boldsymbol{\varepsilon}), \mathbf{Y}_n) = \\
&\lim_{\boldsymbol{\varepsilon}\to\mathbf{0}} \frac{\left|{}_2F_0\left(a_1 - n + 1, 1;; -x_i(\boldsymbol{\varepsilon}) y_j\right)\right|}{\prod_{i=1}^{n}(-a_1 + i - 1)^{i-1} \cdot \alpha_n(\mathbf{X}_n(\boldsymbol{\varepsilon})) \cdot \alpha_n(\mathbf{Y}_n)}
\end{aligned} \tag{137}
$$

whereas we have introduced the diagonal auxiliary matrix

$$\mathbf{X}_n(\boldsymbol{\varepsilon}) = \operatorname{diag}\left(\mathbf{X}_m, \boldsymbol{\varepsilon}\right) = \operatorname{diag}\left(x_1(\boldsymbol{\varepsilon}), \ldots, x_n(\boldsymbol{\varepsilon})\right) \tag{138}$$

with a $1 \times (n-m)$ vector $\boldsymbol{\varepsilon}$

$$\boldsymbol{\varepsilon} = \left\{\varepsilon_0, \ldots, \varepsilon_{n-m-1}\right\}. \tag{139}$$

The limit in (137) is a $\frac{0}{0}$ type of limit and we can apply L'Hospital's rule. To this end, we differentiate nominator and denominator k times with respect to ε_k ($k = 0 \ldots n-m-1$), which shall symbolically be denoted by $\frac{\partial}{\partial \boldsymbol{\varepsilon}}$, and then take the limit $\boldsymbol{\varepsilon} \to \mathbf{0}$. Using

$$\frac{\partial^k}{\partial z^k} {}_2F_0\left(a, b;; -cz\right) = (-1)^k \cdot c^k \cdot [a]_k\, [b]_k \cdot {}_2F_0\left(a+k, b+k;; -cz\right) \tag{140}$$

we get

$$\zeta_1 = \frac{\partial}{\partial \boldsymbol{\varepsilon}} \left| {}_2F_0\left(a_1 - n + 1, 1;; -x_i(\boldsymbol{\varepsilon}) y_j\right) \right| \Big|_{\boldsymbol{\varepsilon}=\mathbf{0}} = \left| \begin{cases} {}_2F_0\left(a_1 - n + 1, 1;; -x_i y_j\right) & i \le m \\ (-1)^k \cdot y_j^k \cdot [a_1 - n + 1]_k\, [1]_k & i > m \end{cases} \right| \tag{141}$$

where for brevity $k = i - m - 1$. By row-wise factoring out common factors from the determinant we find

$$\zeta_1 = \prod_{l=0}^{n-m-1} [a_1 - n + 1]_l\, [1]_l\, (-1)^l \times \left| \begin{cases} {}_2F_0\left(a_1 - n + 1, 1;; -x_i y_j\right) & i \le m \\ y_j^{i-m-1} & i > m \end{cases} \right|. \tag{142}$$

Now we can exploit the relations

$$\prod_{l=0}^{n-m-1} [1]_l = \Gamma_{n-m}(n-m) \tag{143}$$

and

$$\prod_{l=0}^{n-m-1} (-1)^l \, [1-x]_l = \prod_{l=0}^{n-m-1} [x-l]_l = \prod_{l=1}^{n-m-1} (x-l)^{n-m-l} \tag{144}$$

for rewriting (142). Furthermore, from (5) we can establish the following identity for the Vandermonde determinant

$$\frac{\partial}{\partial \boldsymbol{\varepsilon}} \alpha_n\left(\mathbf{X}_n(\boldsymbol{\varepsilon})\right)\Big|_{\boldsymbol{\varepsilon}=\mathbf{0}} = \alpha_m\left(\mathbf{X}_m\right) \cdot |\mathbf{X}_m|^{n-m} \cdot \Gamma_{n-m}(n-m) \cdot (-1)^{\frac{(n-m)\cdot(n-m-1)}{2}}. \tag{145}$$

Finally, we can plug the partial results in (137) and after minor modifications we arrive at (135). ■

We further specialize Corollary 12 by an additional limiting process.

Corollary 13: The hypergeometric function of two matrix arguments ${}_2\tilde{\mathbf{F}}_0^{(m,n)}\left(a_1, n;; -\mathbf{I}_m, \mathbf{Y}_n\right)$ with identity $m \times m$ matrix $\mathbf{I}_m$ and diagonal $n \times n$ ($n \geq m$) matrix $\mathbf{Y}_n$ has the scalar representation

$${}_2\tilde{\mathbf{F}}_0^{(m,n)}\left(a_1, n;; -\mathbf{I}_m, \mathbf{Y}_n\right) = \frac{|\mathbf{Y}_n|^{n-m-1} \cdot \left| \begin{cases} U\left(i, -a1+i+1, \frac{1}{y_j}\right) & i \leq m \\ y_j^{-(n-i)} & i > m \end{cases} \right|}{(-1)^{(n-m)\cdot n} \cdot \alpha_n\left(-\mathbf{Y}_n\right)}, \tag{146}$$

where i and j run from 1 to n.

Proof: We have to calculate the limit $\mathbf{X} = \operatorname{diag}\left(x_1, \ldots, x_m\right) \to \mathbf{I}_m$ in Corollary 12. To this end, we differentiate nominator and denominator of (135) $(i-1)$ times with respect to x_i, which is indicated by $\frac{\partial}{\partial \mathbf{x}}$, and then set $x_i = 1$ for all i. We find

$$\frac{\partial}{\partial \mathbf{x}} \left| \begin{cases} {}_2F_0\left(a_1-n+1, 1;; -x_i y_j\right) & i \leq m \\ y_j^{m-i-1} & i > m \end{cases} \right| \Bigg|_{\mathbf{x}=\mathbf{1}} = \left| \begin{cases} (-1)^{i-1} \cdot y_j^{i-1} \cdot [a_1-n+1]_{i-1} \, [1]_{i-1} \cdot {}_2F_0\left(a_1-n+i, i;; -y_j\right) & i \leq m \\ y_j^{m-i-1} & i > m \end{cases} \right| = \prod_{l=1}^{m-1} (-1)^l \cdot [a_1-n+1]_l \, [1]_l \cdot \xi_2 \tag{147}$$

with short-hand notation

$$\xi_2 = \left| \begin{cases} \frac{1}{y_j} \cdot U\left(i, -a_1 + n + 1, \frac{1}{y_j}\right) & i \leq m \\ y_j^{m-i-1} & i > m \end{cases} \right| . \tag{148}$$

Now we can use the formula [33, equation 13.4.18]

$$U(a,b,z) = \frac{1}{z} \cdot \left[(b-a-1) \cdot U(a,b-1,z) + U(a-1,b-1,z)\right] \tag{149}$$

with the special case

$$U(1,b,z) = \frac{1}{z} \cdot \left[(b-2) \cdot U(a,b-1,z) + 1\right] . \tag{150}$$

We can expand the Kummer U functions in the first m rows of the determinant in (148) with the help of (150) and (149), respectively. Then we can subtract a properly scaled multiple of the $(m+1)$th row from the first row. After that, we can iteratively subtract a scaled version of the first row from the second, of the second from the third and so on. If we repeat this process in total $(n-m)$ times, we find

$$\begin{aligned} \xi_2 = & \left| \begin{cases} \xi_{21}(i) \cdot U\left(i, -a_1 + m + 1, \frac{1}{y_j}\right) & i \leq m \\ y_j^{m-i-1} & i > m \end{cases} \right| \\ = & \xi_{22} \cdot |\mathbf{Y}_n|^{n-m-1} \cdot \left| \begin{cases} U\left(i, -a_1 + m + 1, \frac{1}{y_j}\right) & i \leq m \\ y_j^{-(n-i)} & i > m \end{cases} \right| \end{aligned} \tag{151}$$

with

$$\xi_{21}(i) = (-1)^{n-m} \cdot [a_1 - n + i]_{n-m} \cdot {y_j}^{n-m-1} \tag{152}$$

and

$$\xi_{22} = (-1)^{m \cdot (n-m)} \cdot \prod_{k=0}^{m-1} [a_1 - n + 1 + k]_{n-m} . \tag{153}$$

By exploiting the recurrence relation [33, equation (13.4.17)]

$$U(a,b-1,z) = U(a,b,z) - a \cdot U(a+1,b,z) \tag{154}$$

we can iteratively subtract scaled versions of the $(i+1)$th row from the ith column of the determinant in (151) and after repeating this process m times we find

$$\xi_2 = \xi_{22} \cdot |\mathbf{Y}_n|^{n-m-1} \cdot \left| \begin{cases} U\left(i, -a_1+i+1, \frac{1}{y_j}\right) & i \leq m \\ y_j^{-(n-i)} & i > m \end{cases} \right|. \tag{155}$$

On the other hand, it can be shown by the product rule of differentiation that for all k with note $\mathbf{X} = \operatorname{diag}(\mathbf{x})$

$$\frac{\partial}{\partial \mathbf{x}} \left[|\mathbf{X}|^k \cdot \alpha_m(\mathbf{X}) \right] \Big|_{\mathbf{x}=\mathbf{1}} = \frac{\partial}{\partial \mathbf{x}} \alpha_m(\mathbf{X}) \Big|_{\mathbf{x}=\mathbf{1}} = (-1)^{\frac{m\cdot(m-1)}{2}} \cdot \Gamma_m(m). \tag{156}$$

Combining all partial results, we find

$$\begin{aligned} &{}_2\tilde{\mathbf{F}}_0^{(m,n)}(a_1, n;; -\mathbf{I}_m, \mathbf{Y}_n) = \\ &\zeta \cdot \frac{|\mathbf{Y}_n|^{n-m-1} \cdot \left| \begin{cases} U\left(i, -a1+i+1, \frac{1}{y_j}\right) & i \leq m \\ y_j^{-(n-i)} & i > m \end{cases} \right|}{(-1)^{(n-m)\cdot m} \cdot \alpha_n(-\mathbf{Y}_n)} \end{aligned} \tag{157}$$

with

$$\zeta = \frac{\prod_{k=0}^{m-1} [a_1-n+1+k]_{n-m} \cdot \prod_{k=0}^{m-1} [a_1-n+1]_k}{\prod_{k=n-m}^{n-1} [a_1-n+1]_k}. \tag{158}$$

Finally, by writing the Pochhammer expressions in terms of ratios of Gamma functions, it can be shown that ζ reduces to $\zeta = 1$. ■

Similar to Corollary 11, we study a special hypergeometric function of one matrix argument that occurs in the analysis of the MGF of mutual information in more detail.

Corollary 14: The hypergeometric function ${}_2\tilde{\mathbf{F}}_0^{(m)}(a_1, n;; -\mathbf{X})$ of one $m \times m$ matrix argument $\mathbf{X}$ and parameter a_1 has the scalar representations

$$
\begin{aligned}
&{}_2\tilde{\mathbf{F}}_0^{(m)}\left(a_1, n;; -\mathbf{X}\right) = \\
&\frac{\left| x_i^{m-j} \cdot {}_2F_0\left(a_1 - j + 1, n - j + 1;; -x_i\right)\right|}{\alpha_m\left(\mathrm{eig}\left(\mathbf{X}\right)\right)} = && (159)\\
&\frac{|\mathbf{X}|^{-(n-m+1)} \cdot \left| U\left(n - j + 1, 1 + n - a_1, \frac{1}{x_i}\right)\right|}{\alpha_m\left(\mathrm{eig}\left(\mathbf{X}\right)\right)} = && (160)\\
&\frac{|\mathbf{X}|^{-(n-m+1)} \cdot \left| U\left(n - j + 1, n - j + 2 - a_1, \frac{1}{x_i}\right)\right|}{\alpha_m\left(\mathrm{eig}\left(\mathbf{X}\right)\right)}, && (161)
\end{aligned}
$$

where i and j run from 1 to μ.

Proof: The first part of the corollary in (159) directly follows from Lemma 1. By reformulating the hypergeometric function in terms of Kummer U functions we find (160). The last part in (161) can be established by exploiting the recurrence relation (154). To this end, we can iteratively subtract scaled versions of the $(j+1)$th column from the jth column of the determinant (this does not change the result) until finally the structure in (161) can be set up. ■

We further specialize Corollary 14.

Corollary 15: The hypergeometric function ${}_2\tilde{\mathbf{F}}_0^{(m)}\left(a_1, n;; -c \cdot \mathbf{I}_m\right)$ of one $m \times m$ identity matrix argument $c \cdot \mathbf{I}_m$ with constant c and parameter a_1 has the scalar representation

$$
\begin{aligned}
&{}_2\tilde{\mathbf{F}}_0^{(m)}\left(a_1, n;; -c \cdot \mathbf{I}_m\right) = \\
&\frac{\left| \int_0^\infty e^{-t} \cdot t^{n+m-j-i-1} \cdot (1 + ct)^{-a_1} \mathrm{d}t \right|}{\Gamma_m(m) \cdot \prod_{k=1}^m \Gamma(n - k + 1)}, && (162)
\end{aligned}
$$

where i and j run from 1 to m.

Proof: By exploiting the integral representation of the Kummer U function (3) we find

$$
\begin{aligned}
&{}_2\tilde{\mathbf{F}}_0^{(m)}\left(a_1, n;; -\mathbf{X}\right) = \\
&\frac{\left| \frac{1}{\Gamma(n-j+1)} \cdot \int_0^\infty e^{-\frac{t}{x_i}} \cdot t^{n-j} \cdot (1 + t)^{-a_1} \mathrm{d}t \right|}{|\mathbf{X}|^{n-m+1} \cdot \alpha_m\left(\mathrm{eig}\left(\mathbf{X}\right)\right)}. && (163)
\end{aligned}
$$

We then calculate the limit $\mathbf{X} = \mathrm{diag}\left(x_1, \ldots, x_m\right) \to c \cdot \mathbf{I}_m$. To this end, we differentiate nominator and denominator of (163) $(i-1)$ times with respect to x_i, which is indicated by $\frac{\partial}{\partial \mathbf{x}}$, and then set $x_i = c$ for all i. For differentiating the nominator, we can exchange the sequence of differentiation and integration and make use of the derivative

$$\frac{\partial^k}{\partial c^k} e^{-t/c} = \left(\sum_{l=1}^{k-1} \frac{\alpha_l}{c^{l+k}} \cdot t^l + \frac{t^k}{c^{2k}} \right) \cdot e^{-t/c} \tag{164}$$

with constants α_l. Then iteratively adding a properly scaled multiple of the 1st, 2nd, ..., $(i-1)$th row to the ith row in the resulting determinant after the differentiation, we find

$$\begin{aligned} &\frac{\partial}{\partial \mathbf{x}} \left| \frac{1}{\Gamma(n-j+1)} \cdot \int_0^\infty e^{-\frac{t}{x_i}} \cdot t^{n-j} \cdot (1+t)^{-a_1} \mathrm{d}t \right| \Bigg|_{\mathbf{x}=c\cdot\mathbf{1}} = \\ &\left| \frac{1}{\Gamma(n-j+1)} \cdot \int_0^\infty e^{-t/c} \cdot \frac{t^{n-j+i-1}}{c^{2(i-1)}} \cdot (1+t)^{-a_1} \mathrm{d}t \right| = \\ &\frac{c^{m\cdot(n-m+1)}}{\prod_{k=1}^m \Gamma(n-k+1)} \cdot \left| \int_0^\infty e^{-t} \cdot t^{n-m+j+i-2} \cdot (1+ct)^{-a_1} \mathrm{d}t \right|. \end{aligned} \tag{165}$$

On the other hand, it can be shown by the product rule of differentiation that for diagonal $m \times m$ matrix $\mathbf{X}$ and all k

$$\frac{\partial}{\partial \mathbf{x}} \left[|\mathbf{X}|^k \cdot \alpha_m(\mathbf{X}) \right] \Big|_{\mathbf{x}=c\cdot\mathbf{1}} = |\mathbf{X}|^k \cdot \frac{\partial}{\partial \mathbf{x}} \alpha_m(\mathbf{X}) \Big|_{\mathbf{x}=c\cdot\mathbf{1}} = (c)^{mk} \cdot (-1)^{\frac{m\cdot(m-1)}{2}} \cdot \Gamma_m(m). \tag{166}$$

Now using (165) and (166) in (163) we can establish the Corollary. ∎

APPENDIX II
MATRIX QUADRATIC FORMS

A. Probability Distribution Function

The PDF of a complex generalized matrix random quadratic form $\mathbf{S} = \mathbf{M^{1/2} G^H N G M^{1/2}}$ with i.i.d. complex Gaussian distributed $n \times m$ matrix $\mathbf{G} \sim \mathcal{N}_{n,m}(\mathbf{0}, \mathbf{I}_n, \mathbf{I}_m)$, deterministic $n \times n$ matrix $\mathbf{N}$ and deterministic $m \times m$ matrix $\mathbf{M}$ is given by [32]

$$p_\mathbf{S}(\mathbf{S}) = \frac{|\mathbf{S}|^{n-m} \cdot \mathrm{etr}\left(-q^{-1}\mathbf{M}^{-1}\mathbf{S}\right)}{\tilde{\Gamma}_m(n) \cdot |\mathbf{M}|^n \cdot |\mathbf{N}|^m} \cdot {}_0\tilde{\mathbf{F}}_0^{(m,n)}\left(;;-q^{-1}\mathbf{M}^{-1}\mathbf{S}, \mathbf{I}_n - q\mathbf{N}^{-1}\right), \tag{167}$$

with scalar $q > 0$ and $n \geq m$.

B. Related Integrals

For deterministic $m \times m$ matrix $\mathbf{Z}$ we have [32, equation (58)]

$$\int \tilde{C}_\kappa(\mathbf{ZS})\, p_{\mathbf{S}}(\mathbf{S})\ \mathrm{D}_c\mathbf{S} = [n]_\kappa \cdot \frac{\tilde{C}_\kappa(\mathbf{ZM}) \cdot \tilde{C}_\kappa(\mathbf{N})}{\tilde{C}_\kappa(\mathbf{I}_n)}. \tag{168}$$

From (101) we get with the help of (168)

$$\begin{aligned} &\int {}_p\tilde{\mathbf{F}}_q^{(m)}\left(a_1,\ldots,a_p;b_1,\ldots,b_q;\mathbf{ZS}\right) p_{\mathbf{S}}(\mathbf{S})\ \mathrm{D}_c\mathbf{S} = \\ &\sum_{k=0}^{\infty}\sum_\kappa \frac{[a_1]_\kappa\cdot\ldots\cdot[a_p]_\kappa}{[b_1]_\kappa\cdot\ldots\cdot[b_q]_\kappa}\cdot\frac{\int \tilde{C}_\kappa(\mathbf{ZS})p_{\mathbf{S}}(\mathbf{S})\ \mathrm{D}_c\mathbf{S}}{k!} = \\ &\sum_{k=0}^{\infty}\sum_\kappa \frac{[a_1]_\kappa\cdot\ldots\cdot[a_p]_\kappa}{[b_1]_\kappa\cdot\ldots\cdot[b_q]_\kappa}\cdot\frac{[n]_\kappa\cdot\tilde{C}_\kappa(\mathbf{ZM})\cdot\tilde{C}_\kappa(\mathbf{N})}{k!\cdot\tilde{C}_\kappa(\mathbf{I}_n)} = \\ &{}_{p+1}\tilde{\mathbf{F}}_q^{(m,n)}\left(a_1,\ldots,a_p,n;b_1,\ldots,b_q;\mathbf{ZM},\mathbf{N}\right) \quad , \end{aligned} \tag{169}$$

which is a key integral for obtaining the MGF of mutual information in this paper. For completeness, we note that the integral in (169) has independently been solved in [38, equation (50)], where it was given in a similar form for deterministic $n \times n$ matrices $\mathbf{X}$, $\mathbf{Y}$, and complex i.i.d. Gaussian $n \times n$ matrix $\mathbf{G}$

$$\begin{aligned} &\int {}_p\tilde{\mathbf{F}}_q^{(n)}\left(a_1,\ldots,a_p;b_1,\ldots,b_q;\mathbf{G}^{\mathrm{H}}\mathbf{XGY}\right) p_{\mathbf{G}}(\mathbf{G})\ \mathrm{D}_c\mathbf{G} = \\ &{}_{p+1}\tilde{\mathbf{F}}_q^{(n,n)}\left(a_1,\ldots,a_p,n;b_1,\ldots,b_q;\mathbf{X},\mathbf{Y}\right) \end{aligned} \tag{170}$$

with multivariate complex Gaussian PDF

$$p_{\mathbf{G}}(\mathbf{G}) = \frac{1}{\pi^{n^2}} \cdot \mathrm{etr}\left(-\mathbf{G}^{\mathrm{H}}\mathbf{G}\right). \tag{171}$$

C. Wishart Case

With $\mathbf{N} = \mathbf{I}_n$, the complex generalized matrix random quadratic form $\mathbf{S}$ reduces to a so-called Wishart matrix $\mathbf{W} = \mathbf{M}^{1/2}\mathbf{G}^{\mathrm{H}}\mathbf{G}\mathbf{M}^{1/2}$ with i.i.d. complex Gaussian distributed $n \times m$ matrix $\mathbf{G}$, and deterministic $m \times m$ matrix $\mathbf{M}$. Its PDF is given by [2]

$$p_{\mathbf{W}}(\mathbf{W}) = \frac{|\mathbf{W}|^{n-m} \cdot \mathrm{etr}\left(-\mathbf{M}^{-1}\mathbf{W}\right)}{\tilde{\Gamma}_m(n) \cdot |\mathbf{M}|^n}, \tag{172}$$

with $n \geq m$. From (168) we directly get for $m \times m$ matrix $\mathbf{Z}$

$$\int \tilde{C}_\kappa(\mathbf{ZW})\, p_{\mathbf{W}}(\mathbf{W})\ \mathrm{D}_c\mathbf{W} = [n]_\kappa \cdot \tilde{C}_\kappa(\mathbf{ZM})\,. \tag{173}$$

However, for completeness we note that (173) is a direct consequence of the integral [32, equation (53)]

$$\int \tilde{C}_\kappa(\mathbf{ZX}) \cdot \mathrm{etr}(-\mathbf{X}) \cdot |\mathbf{X}|^{n-m}(\mathbf{X})\ \mathrm{D}_c\mathbf{X} = [n]_\kappa \cdot \tilde{\Gamma}_m(n) \cdot \tilde{C}_\kappa(\mathbf{Z})\,. \tag{174}$$

Similar to (169), with the help of (173) we find

$$\begin{aligned} \int {}_p\tilde{\mathbf{F}}_q^{(m)}\left(a_1,\ldots,a_p;b_1,\ldots,b_q;\mathbf{ZW}\right) p_{\mathbf{W}}(\mathbf{W})\ \mathrm{D}_c\mathbf{W} = \\ {}_{p+1}\tilde{\mathbf{F}}_q^{(m)}\left(a_1,\ldots,a_p,n;b_1,\ldots,b_q;\mathbf{ZM}\right) \qquad . \end{aligned} \tag{175}$$

REFERENCES

[1] Telatar I. E., "Capacity of multi-antenna Gaussian channels," *Bell Labs Technical Memorandum*, October 1995

[2] James A. T., "Distributions of matrix variates and latent roots derived from normal samples," *Annals of Mathematical Statistics*, vol. 35, pp. 475-501, 1964

[3] Foschini G. J., Gans M. J., "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, no. 6, pp. 311-335, 1998

[4] Kang M., Alouini M.-S., "On the outage capacity of MIMO channels," *IEEE Transactions on Information Theory*, 2003, submitted

[5] Chiani M., Win M. Z., Zanella A., "On the capacity of spatially correlated MIMO Rayleigh-fading channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2363-2371, Oct. 2003

[6] Smith P. J., Roy S., Shafi M., "Capacity of MIMO systems with semicorrelated flat fading," *IEEE Transactions on Information Theory*, vol. 49, no. 10, Oct. 2003

[7] Grant A., "Rayleigh fading multiple-antenna channels," *EURASIP Journal on Applied Signal Processing, Special Issue on Space-Time Coding (Part I)*, vol. 2002, no. 3, pp. 316-329, March 2002

[8] Ivrlac M. T. , Kurpjuhn T. P., Brunner C., Utschick W., "Efficient use of fading correlations in MIMO systems," *IEEE Vehicular Technology Conference*, VTC, Oct. 2001

[9] Loyka S., Kouki A., "On the use of Jensens inequality for MIMO channel capacity estimation," *Canadian Conference on Electrical and Computer Engineering*, vol. 1, pp. 475-480, 2001

[10] Shiu D., Foschini G. J., Gans M. J., Kahn J. M., "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 502-513, March 2000

[11] Herz C. S., "Bessel functions of matrix argument," *Annals of Mathematics*, vol. 61, no. 3, pp. 474-523, May 1955

[12] Wang Z., Giannakis G. B., "Outage mutual information of space-time MIMO channels," *IEEE Transactions on Information Theory*, 2003, submitted

[13] Smith P. J., Shafi M., ”On a Gaussian approximation to the capacity of wireless MIMO systems,” *IEEE International Conference on Communications*, ICC, May 2002

[14] Kang M., Yang L., Alouini M.-S., ”How accurate are the Gaussian and Gamma approximations to the outage capacity of MIMO channels?,” *6th Baiona Workshop on Signal Processing in Communications*, Sept. 2003

[15] de Hoog F. R., Knight J. H., Stokes A. N., ”An improved method for numerical inversion of Laplace transforms,”*SIAM Journal of Scientific Computation*, vol. 3, no. 3, pp. 357-366, Sept. 1982

[16] Kang M., Alouini M.-S., ”Largest eigenvalue of complex Wishart matrices and performance analysis of MIMO MRC systems,” *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 3, pp. 418-426, April 2003

[17] Khatri C. G., ”Distribution of the largest or smallest characteristic root under null hypothesis concerning complex multivariate normal populations,” *Annals of Mathematical Statistics*, vol. 35, pp. 1807-1810, 1964

[18] Khatri C. G., ”Non-central distributions of ith largest characteristic roots of three matrices concerning complex multivariate normal populations,” *Annals of the Institute of Statistical Mathematics*, vol. 21, pp. 23-32, 1969

[19] Simon S. H., Moustakas A. L., Matinelli L., ”Capacity and character expansions: moment generating function andother exact results for MIMO correlated channels,” *preprint*

[20] Marcenko V. A., Pastur L. A., ”Distribution of eigenvalues for some sets of random matrices,” *Mathematics of USSR - Sbornik 1*, pp. 457-483, 1967

[21] Silverstein J. W., Bai Z. D, ”On the empirical distribution of eigenvalues of a class of large dimensional random matrices,” *Journal of Multivariate Analysis*, vol. 54, no. 2, pp. 175-192, 1995

[22] Rapajic P. B., Popescu D., ”Information capacity of a random signature multiple-input multiple-output channel,” *IEEE Transactions on Communications*, vol. 48, pp. 1245-48, Aug. 2000

[23] Chuah C.-N., Tse D. N. C., Kahn J. M., Valenzuela R. A., ”Capacity scaling in MIMO wireless systems under correlated fading,” *IEEE Transactions on Information Theory*, vol. 48, no.3, pp. 637-650, March 2002

[24] Verdu S., Shamai S., ”Spectral efficiency of CDMA with random spreading,” *IEEE Transactions on Information Theory*, vol. 45, pp. 622-640, March 1999

[25] Moustakas A. L., Simon S. H., Sengupta A. M., ”MIMO capacity through correlated channels in the presence of correlated interferers and noise: a (not so) large N analysis,” *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2545-2561, Oct. 2003

[26] Goldsmith A., Jafar S. A., Jindal N., Vishwanath S., ”Capacity limits of MIMO channels,” *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, June 2003

[27] Tulino A. M., Verdu S., ”Random matrix theory and wireless communications,” *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, June 2004

[28] Kermoal J. P., Schumacher L., Pedersen K. I., Mogensen P. E., Frederiksen F., ”A stochastic MIMO radio channel model with experimental validation,” *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1211-1226, Aug. 2002

[29] Paulraj A., Nabar R., Gore D., *Introduction to space-time wireless communications*, Cambridge University Press, 2003

[30] Wooding R. A., ”The multivariate distribution of complex normal variables,” *Biometrika*, vol. 43, no. 1/2, pp. 212215, June 1956

[31] Lütkepohl H., *Handbook of matrices*, John Wiley&Sons, 1996

[32] Khatri C. G., ”On certain distribution problems based on positive definite quadratic functions in normal vectors,” *Annals of Mathematical Statistics*, vol. 37, pp. 468-479, 1966

[33] Abramowitz M., Stegun I. A., *Handbook of mathematical functions*, Dover Publications Inc., New York, 1964

[34] Gradsteyn I. S., Ryzhik I. M., *Table of integrals, series, and products*, Academic Press, 6th ed., 2000

[35] Loyka S., ”Channel capacity of MIMO architecture using the exponential correlation matrix,” *IEEE Communication Letters*, vol. 5, no. 9, pp. 369 371, Sep. 2001

[36] Khatri C. G., ”On the moments of traces of two matrices in three situations for complex multivariate normal populations,” *Sankya: The Indian Journal of Statistics, Series A*, vol. 32, part 1, pp. 65-80, 1970

[37] Gross K. I., Richards D. S. P., ”Total positivity, spherical series, and hypergeometric functions of matrix argument,” *Journal on Approximation Theory*, vol. 59, no. 2, pp. 224-246, 1989

[38] Orlov A. Y., ”New solvable matrix integrals”, *Acta Sci. Math.*, vol. 63, pp. 383-395, 1997

[39] Kiessling M., Viering I., Reinhardt M., Speidel J., ”A closed form bound on correlated MIMO channel capacity,” *IEEE Vehicular Technology Conference*, VTC, Sept. 2002

[40] Kiessling M., Speidel J., Reinhardt M., ”Ergodic capacity of MIMO channels with statistical channel state information at the transmitter,” *ITG Workshop on Smart Antennas*, March 2004

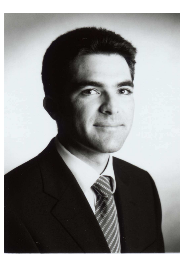
[41] Kiessling M., ”Statistical analysis and transmit processing for MIMO wireless systems in correlated fading environments,” *PhD Thesis*, March 2004

[42] Kiessling M., Speidel J., ”Mutual information of MIMO channels in correlated Rayleigh fading environments - a general solution,” *IEEE International Conference on Communications*, ICC, June 2004

[43] Kiessling M., Speidel J., ”Exact ergodic capacity of MIMO channels in correlated Rayleigh fading environments,” *International Zurich Seminar on Communications*, Feb. 2004

[44] Kiessling M., Speidel J., Reinhardt M., ”Unifying analysis of ergodic MIMO capacity in correlated Rayleigh fading environments,” *European Wireless Conference*, Feb. 2004

[45] Kiessling M., Speidel J., ”Asymptotics of ergodic MIMO capacity in correlated Rayleigh fading environments,” *IEEE Vehicular Technology Conference*, VTC, May 2004



Mario Kießling studied at the University of Stuttgart, Germany, where he received his Dipl. Ing. degree in Electrical Engineering in 2000. Since 2001 he has been with the Institute of Telecommunications at the University of Stuttgart, where he is pursuing a Ph.D. degree in Electrical Engineering, and Siemens Information and Communication Mobile in Ulm, Germany.

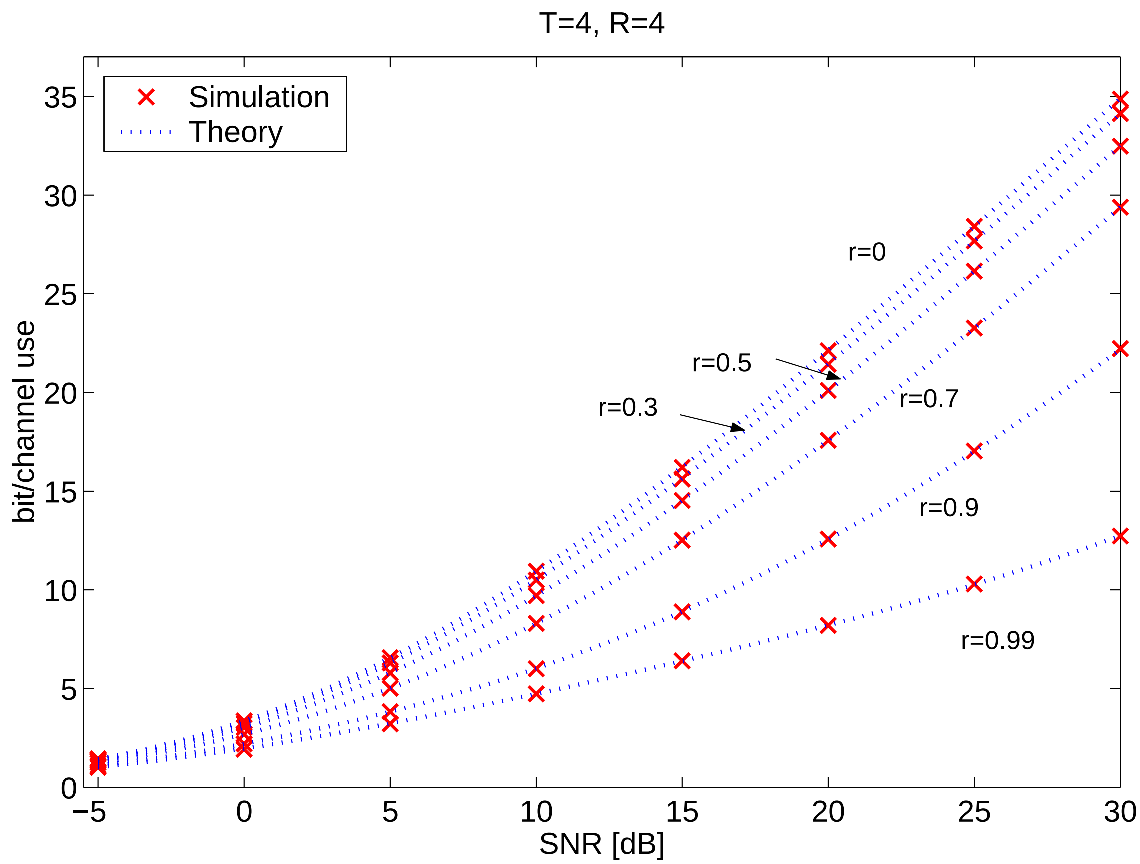


Fig. 1. Ergodic capacity, $r_{TX} = r_{RX} = r$, $T = R = 4$

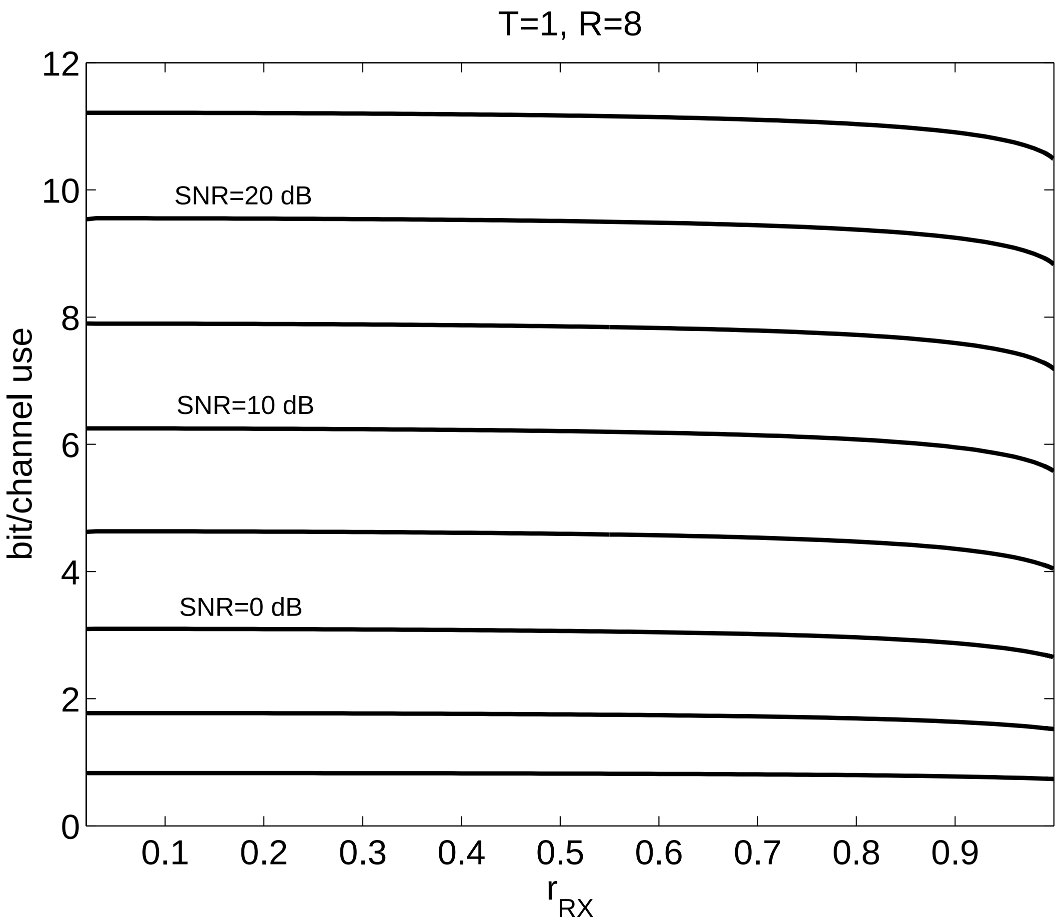


Fig. 2. Ergodic capacity, varying r_{RX}, $T = 1$, $R = 8$

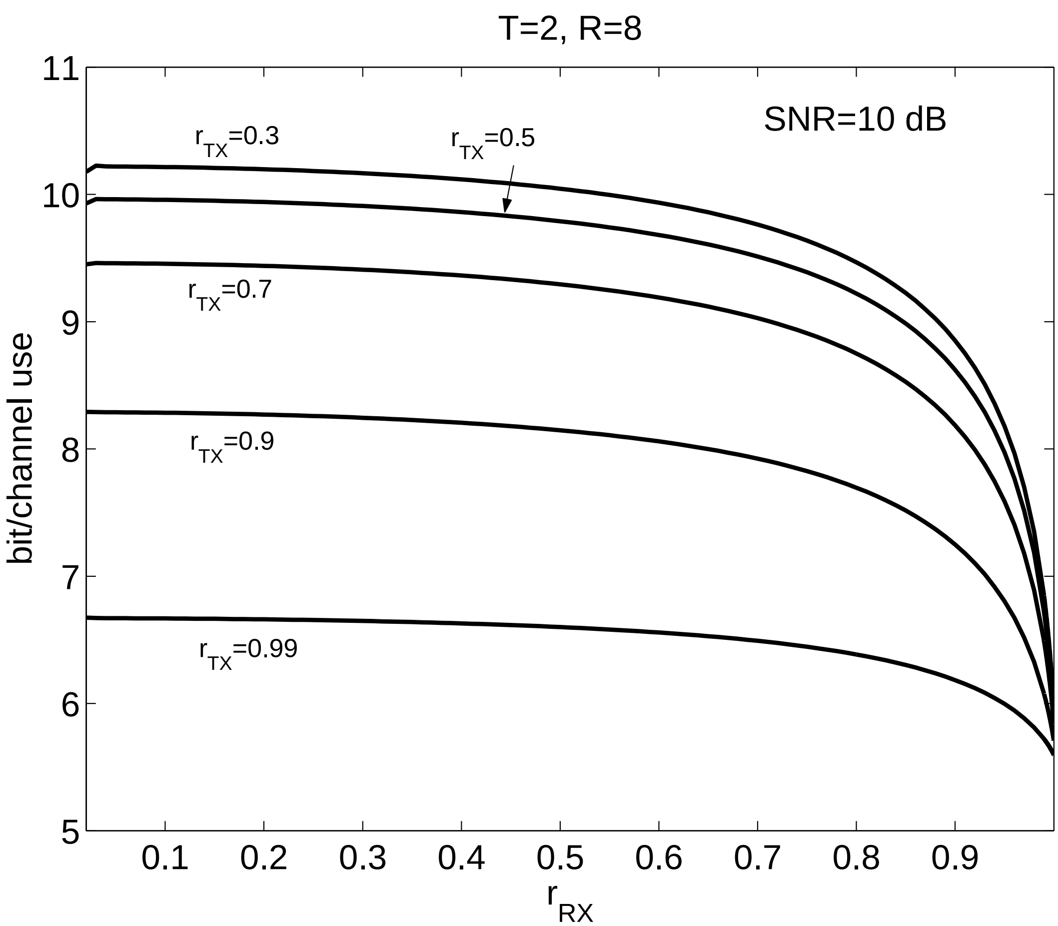


Fig. 3. Ergodic capacity, varying r_{RX}, r_{TX}, fixed $\gamma_{dB} = 10$dB, $T = 2$, $R = 8$