

Random matrix theory of multi-antenna communications: the Ricean channel

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2005 J. Phys. A: Math. Gen. 38 10859

(<http://iopscience.iop.org/0305-4470/38/49/024>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.94

The article was downloaded on 03/06/2010 at 04:04

Please note that [terms and conditions apply](#).

Random matrix theory of multi-antenna communications: the Ricean channel

Aris L Moustakas^{1,3} and Steven H Simon²

¹ Department of Physics, University of Athens, Panepistimiopolis, Athens 15784, Greece

² Bell Labs, Lucent Technologies, 600 Mountain Avenue, Murray Hill, NJ 07974, USA

Received 25 July 2005, in final form 26 July 2005

Published 22 November 2005

Online at stacks.iop.org/JPhysA/38/10859

Abstract

The use of multi-antenna arrays in wireless communications through disordered media promises huge increases in the information transmission rate. It is therefore important to analyse the information capacity of such systems in realistic situations of microwave transmission, where the statistics of the transmission amplitudes (channel) may be coloured. Here, we present an approach that provides analytic expressions for the statistics, i.e. the moments of the distribution, of the mutual information for general Gaussian channel statistics. The mathematical method applies tools developed originally in the context of coherent wave propagation in disordered media, such as random matrix theory and replicas. Although it is valid formally for large antenna numbers, this approach produces extremely accurate results even for arrays with as few as two antennas. We also develop a method to analytically optimize over the input signal distribution, which enables us to calculate analytic capacities when the transmitter has knowledge of the statistics of the channel. The emphasis of this paper is on elucidating the novel mathematical methods used. We do this by analysing a specific case when the channel matrix is a complex Gaussian with arbitrary mean and unit covariance, which is usually called the Ricean channel.

PACS numbers: 84.40.Ba, 02.10.Yn, 07.50.Qx

1. Introduction

Recently, there has been increased interest in using multi-antenna arrays simultaneously in transmission and reception of microwave signals. Recent theoretical work in [1] and [2] has shown that for sufficiently rich scattering environments the Shannon capacity of an n_t -element transmitting and an n_r -element receiving array is roughly proportional to $\min(n_r, n_t)$ for large numbers of antennas. This is significantly more than the usually logarithmic increase in

³ Author to whom correspondence should be addressed.

capacity for increasing antenna numbers when only a single path (plane-wave) connects the transmission and reception arrays (line-of-sight). Intuitively, one can understand this result by observing that if the scattering is rich enough, then there is an independent channel from each transmission antenna to each reception antenna. Therefore, one can send independent signals from each transmission antenna.

Both indoors [3] and outdoors [4] measurements have shown promising throughput gains for such MIMO (multiple input multiple output) technologies. Therefore it is important to be able to assess the capacity gains of MIMO technologies in realistic situations. Such situations include the case where spatial correlations between transmission and/or reception antennas exist, which tend to reduce the effective number of independent channels (paths) between the transmitter and the receiver [5]. Another interesting situation corresponds to the case where the transmitter has partial knowledge of the channel itself. This may arise when the channel feedback to the transmitter from the receiver (who is assumed to always know the channel) may be impaired or noisy. This channel is usually called the Ricean channel in the literature.

For large antenna numbers and Gaussian channel distributions the analysis of MIMO ergodic capacities (expectation value of mutual information over channel realizations) is greatly facilitated by the use of asymptotic techniques of random matrix theory (RMT). These methods were introduced in this context by various authors, starting with Foschini [1] and Telatar [2]. Verdú and Shamai [6] derived the expression of the capacity for infinite antennas with uncorrelated channels in the context of CDMA codes, and more recently Rapajic and Popescu [7] applied it in the context of multi-antenna systems. Lozano and Tulino [8] using methods developed by Tse and Hanley [9] calculated the infinite antenna-number capacity with spatially uncorrelated channels and uncorrelated interferers. Also, Chuah *et al* [10] extended the results of [6] to calculate the mutual information of spatially correlated channels in the infinite antenna limit. In all previous studies, the capacity has been assumed to be asymptotically proportional to $\min(n_t, n_r)$ and only the corresponding proportionality factor was studied as the number of antennas grows indefinitely. However, for finite antenna numbers or for slowly decaying spatial correlations the capacity of arrays with finite antenna numbers cannot be simply described as linear in the antenna number. For example, as shown in figure 2 of [5], for square arrays with $\lambda/2$ spacing the capacity per antenna does not reach the limiting value even at 10 000 antennas per array! In [5], using techniques developed by Sengupta and Mitra [11] in a different context, a method was developed to analytically calculate the capacity of spatially correlated channels for large but finite antenna numbers.

In this paper we extend work done in [5, 12, 13] to provide analytic expressions for the statistics of the mutual information in the case of the Ricean channel. We apply a method from physics, known as the replica approach, to analyse the resulting random matrix problems. The replica approach, first introduced in [14], has been used heavily in physics for understanding random systems [15, 16]. One of the first applications of this method in communication theory was by Sourlas [17] in the field of error correcting codes. More recently, this method has seen increased application in the field of information theory [18–20].

We will use the replica method to average over the channel realizations and obtain moments of the distribution of the mutual information. We find that for large antenna numbers n , the second and third moments of the mutual information is of $\mathcal{O}(1)$ and $\mathcal{O}(1/n)$ respectively, which shows that the mutual information distribution approaches a Gaussian. It should be noted that recently the generating function of the mutual information for the Ricean channel [21, 22], as well as some other channels [23], was calculated in closed form using other methods. Nevertheless, the replica method produces far simpler equations, while the equations resulting from other methods become increasingly difficult to evaluate for large antenna numbers. Surprisingly, the replica method also gives accurate results when applied to arrays with even

few (two or three) antennas. Thus, this analytic approach provides a powerful tool for analysing antenna systems with even a few antennas.

In the remainder of this section we provide some notational definitions used in this paper (section 1.1) and define several quantities of interest (section 1.2). In section 2 we describe the mathematical framework of the methods used to calculate the statistics of the mutual information. In the next section (section 3) we maximize over the input signal distribution to calculate the maximum mutual information (channel capacity). In section 4 we discuss the Gaussian character of the distribution by presenting a few numerical examples. Finally, in the appendix we provide without proof several useful identities regarding complex integrals.

1.1. Notation

1.1.1. Vectors/matrices. Throughout this paper we will use bold-faced upper-case letters to denote matrices, e.g. \mathbf{X} , with elements given by X_{ab} , bold-faced lower-case letters for column vectors, e.g. \mathbf{x} , with elements x_a , and non-bold lower-case letters for scalar quantities. Also the superscripts T and \dagger will indicate transpose and Hermitian conjugate operations and \mathbf{I}_n will represent the n -dimensional identity matrix. In addition, $\mathbf{K} = \mathbf{M} \otimes \mathbf{N}$ will depict the outer product between the $m \times m$ -dimensional matrix \mathbf{M} and the $n \times n$ -dimensional \mathbf{N} .

1.1.2. Order of number of antennas $\mathcal{O}(n^k)$. We will be examining quantities in the limit when all n_t, n_r, n_i are large but their ratios are fixed and finite. We will denote collectively the order in an expansion over the antenna numbers as $\mathcal{O}(n), \mathcal{O}(1), \mathcal{O}(1/n)$, etc, irrespective of whether the particular term involves n_t, n_r .

1.1.3. Integral measures. In this paper we will be dealing with two general types of integrals over matrix elements. We will therefore adopt the following notation for their corresponding integration measures. In the first type we will be integrating over the real and imaginary parts of the elements of a complex $m_{\text{rows}} \times m_{\text{cols}}$ matrix \mathbf{X} . The integral measure will be denoted by

$$D\mathbf{X} = \prod_{a=1}^{m_{\text{rows}}} \prod_{\alpha=1}^{m_{\text{cols}}} \frac{d\text{Re } X_{a\alpha} d\text{Im } X_{a\alpha}}{2\pi}. \quad (1)$$

The second type of integration is over pairs of complex square matrices \mathbf{T} and \mathbf{R} . Each element of \mathbf{T} and \mathbf{R} will be integrated over a contour in the complex plane (to be specified). The corresponding measure will be described as

$$d\mu(\mathbf{T}, \mathbf{R}) = \prod_{a=1}^{m_{\text{rows}}} \prod_{\alpha=1}^{m_{\text{cols}}} \frac{dT_{a\alpha} dR_{a\alpha}}{2\pi i} \quad (2)$$

1.2. Definitions

We consider the case of single-user transmission from n_t transmit antennas to n_r receive antennas over a narrow band fading channel. The received n_r -dimensional complex signal vector \mathbf{y} can be written as

$$\mathbf{y} = \sqrt{\frac{\rho}{n_t}} \mathbf{G}\mathbf{x} + \mathbf{z}, \quad (3)$$

where \mathbf{G} is an $n_r \times n_t$ complex matrix with the channel coefficients from the transmitting to the receiving arrays, while \mathbf{x} and \mathbf{z} are n_t - and n_r -dimensional vectors of the transmitted signal and the additive noise, respectively, both assumed to be zero-mean Gaussian. The signal

covariance $\mathbf{Q} = E[\mathbf{x}\mathbf{x}^\dagger]$ is normalized so that $\text{Tr } \mathbf{Q} = n_t$. For simplicity, the noise vector \mathbf{z} is assumed to be white with unit covariance, normalized so that $E[\mathbf{z}\mathbf{z}^\dagger] = \mathbf{I}_{n_r}$. Finally, ρ is the signal-to-noise ratio (SNR). It is assumed that the receiver knows the channel matrix \mathbf{G} and ρ . The transmitter, on the other hand, knows only the statistics of the noise, $p(\mathbf{y}|\{\mathbf{x}, \mathbf{G}\})$, as well as the statistics of the channel $p(\mathbf{G})$.

The associated mutual information, i.e. the reduced entropy of the random variable \mathbf{x} , gives the knowledge of the received signal \mathbf{y} , can be expressed as [1]

$$h(\mathbf{x}|\mathbf{G}) - h(\mathbf{x}|\mathbf{y}, \mathbf{G}) = I(\mathbf{y}; \mathbf{x}|\mathbf{G}) = \log \det \left(\mathbf{I}_{n_r} + \frac{\rho}{n_t} \mathbf{G}\mathbf{Q}\mathbf{G}^\dagger \right). \quad (4)$$

The log above (and throughout the whole paper) represents the natural logarithm and thus I is expressed in nats.

Due to the underlying randomness of \mathbf{G} , I is also a random quantity. The average mutual information can then be expressed as

$$\langle I(\mathbf{y}; \mathbf{x}|\mathbf{G}) \rangle = \left\langle \log \det \left(\mathbf{I}_{n_r} + \frac{\rho}{n_t} \mathbf{G}\mathbf{Q}\mathbf{G}^\dagger \right) \right\rangle. \quad (5)$$

When this is maximized over the signal covariance \mathbf{Q} , one obtains the capacity of the channel. This is the maximum error-free information transmission rate when the channel matrix \mathbf{G} varies through its whole distribution $p(\mathbf{G})$. This is why it is also called the *ergodic* capacity. Another important metric of the information capacity of the channel, the so-called outage capacity [24], is obtained by inverting the expression below with respect to I_{out} ,

$$P_{\text{out}} = \text{Prob}(I < I_{\text{out}}) \quad (6)$$

where $\text{Prob}(I < I_{\text{out}})$ is the probability that the mutual information is less than a given value I_{out} .

In this review we will analyse the case where \mathbf{G} has Gaussian statistics with a non-zero mean and unit covariance, i.e. with

$$\langle G_{\alpha\alpha} \rangle = \mu_{\alpha\alpha} \sqrt{p_c} \quad (7)$$

$$\langle G_{\alpha\alpha} G_{\beta\beta}^* \rangle - \langle G_{\alpha\alpha} \rangle \langle G_{\beta\beta} \rangle = p_d \delta_{\alpha\beta} \delta_{ab}. \quad (8)$$

In the above, p_d signifies the fraction of the power due to diffuse (ergodic) components of the channel, while $p_c = 1 - p_d$ is the stationary part of the channel, which is not averaged over. These are related to the commonly used Ricean factor K , by $K = p_c/p_d$. We will also be using the notation $\rho_d = \rho p_d$ and $\rho_c = \rho p_c$ for the diffuse and stationary fractions of the signal to noise ratio.

The simple form of covariance matrix of (7) has been shown [5, 25] to be the leading term in a controlled approximation for diffusive environments. More complicated correlations between matrix elements have been analysed elsewhere [13]. The non-zero mean of \mathbf{G} can be thought to be present due to a non-diffusive and non-ergodic term in the propagation process, such as a line-of-sight component.

We define the notation $\langle \cdot \rangle$ to denote the ensemble average over channel realizations. Thus for $O(\mathbf{G})$, an arbitrary function of \mathbf{G} , we have

$$\langle O \rangle = \int D\mathbf{G} \exp \left(-\frac{1}{2} \text{Tr} \{ (\mathbf{G} - \boldsymbol{\mu})(\mathbf{G} - \boldsymbol{\mu})^\dagger \} \right) O. \quad (9)$$

2. Mathematical framework

The purpose of this paper is to analyse the statistics of the mutual information I in (4) for Gaussian channels having statistics given by (7). In this section we introduce the mathematical framework necessary for deriving analytic expressions of the cumulant moments of I . This method was introduced in this context in [12, 13]. We first introduce the generating function $g(v)$ of I

$$\begin{aligned} g(v) &= \left\langle \left[\det \left(\mathbf{I}_{n_r} + \frac{\rho}{n_t} \mathbf{G} \mathbf{Q} \mathbf{Q}^\dagger \right) \right]^{-v} \right\rangle = \langle e^{-vI} \rangle \\ &= 1 - v \langle I \rangle + \frac{v^2}{2} \langle I^2 \rangle + \dots \end{aligned} \quad (10)$$

Assuming that $g(v)$ is analytic at least in the vicinity of $v = 0$, we can express $\log g(v)$ as follows:

$$\log g(v) = -v \langle I \rangle + \sum_{p=2}^{\infty} \frac{(-v)^p}{p!} C_p, \quad (11)$$

where C_p is the p th cumulant moment of I . For example, $C_2 = \text{Var}(I) = \langle (I - \langle I \rangle)^2 \rangle$ is the variance and $C_3 = \text{Sk}(I) = \langle (I - \langle I \rangle)^3 \rangle$ is the skewness of the distribution. It should be noted that $\langle I \rangle$ and C_p for $p > 2$ are implicit functions of \mathbf{Q} and ρ . For notational simplicity we will be suppressing this dependence. Thus to obtain the moments of the mutual information distribution we need to calculate $g(v)$ for v in the vicinity of $v = 0$. This is not necessarily any easier than evaluating the moments C_p directly, which is a notoriously difficult task, since one has to average products of logarithms of random quantities. In contrast, averaging $g(v)$ for integer values of v involves averages over integer powers of determinants of random quantities, in which case some analytic progress can be made. We will therefore make the following assumption:

Assumption 1 (replica method). $g(v)$ evaluated for positive integer values of v can be analytically continued for real v , specifically in the vicinity of $v = 0^+$.

This assumption, used also in [15, 16, 20], alleviates the problem of dealing with averages of logarithms of random quantities, since the logarithm is obtained after calculating $g(v)$. It has seen widespread use in the field of physics for more than 25 years [14], and, in many cases [11] can be shown to produce exactly the same results as systematic series expansions. Thus, the replica method can be seen essentially a bookkeeping tool. Here, we will be using it without any direct proof, although we will be comparing some of our final results to Monte Carlo simulations to demonstrate their validity.

Therefore, in the following analysis we will assume v to be an arbitrary positive integer. Using identity 1 in the appendix we can write $g(v)$ as

$$g(v) = \int \mathbf{D}\mathbf{X} e^{-\frac{1}{2} \text{Tr}\{\mathbf{X}^\dagger \mathbf{X}\}} \left\langle \exp \left(-\frac{\rho}{2n_t} \text{Tr}\{\mathbf{X}^\dagger \mathbf{G} \mathbf{Q} \mathbf{Q}^\dagger \mathbf{X}\} \right) \right\rangle \quad (12)$$

where \mathbf{X} is a complex $n_r \times v$ matrix. The bracketed quantity in (12) can be rewritten, using identity 2 (with $\mathbf{A} = \mathbf{B} = \mathbf{Q}^{1/2} \mathbf{G}^\dagger \mathbf{X} \sqrt{\rho/n_t}$, where $\mathbf{Q}^{1/2}$ is the matrix square root of \mathbf{Q} , which is well defined since \mathbf{Q} is non-negative definite) by introducing an integral over a

complex $n_t \times \nu$ matrix \mathbf{Y} ,

$$\begin{aligned} & \left\langle \exp \left(-\frac{\rho}{2n_t} \text{Tr}\{\mathbf{X}^\dagger \mathbf{G} \mathbf{Q} \mathbf{G}^\dagger \mathbf{X}\} \right) \right\rangle \\ &= \int \mathbf{D}\mathbf{Y} e^{-\frac{1}{2} \text{Tr}\{\mathbf{Y}^\dagger \mathbf{Y}\}} \left\langle \exp \left(-\sqrt{\frac{\rho}{4n_t}} \text{Tr}\{\mathbf{X}^\dagger \mathbf{G} \mathbf{Q}^{1/2} \mathbf{Y} - \mathbf{Y}^\dagger \mathbf{Q}^{1/2} \mathbf{G}^\dagger \mathbf{X}\} \right) \right\rangle. \end{aligned} \quad (13)$$

At this point we can use identity 2 to integrate over \mathbf{G} . Combining the result of (9), (12), (13), $g(\nu)$ can be expressed as

$$\begin{aligned} g(\nu) &= \int \mathbf{D}\mathbf{X} \mathbf{D}\mathbf{Y} \exp \left(-\frac{1}{2} \text{Tr} \left\{ \mathbf{X}^\dagger \mathbf{X} + \mathbf{Y}^\dagger \mathbf{Y} + \frac{\rho_d}{2n_t} \mathbf{Y}^\dagger \mathbf{Q} \mathbf{Y} \mathbf{X}^\dagger \mathbf{X} \right\} \right) \\ &\quad \times \exp \left(-\sqrt{\frac{\rho_c}{4n_t}} \text{Tr}\{\mathbf{X}^\dagger \boldsymbol{\mu} \mathbf{Q}^{1/2} \mathbf{Y} - \mathbf{Y}^\dagger \mathbf{Q}^{1/2} \boldsymbol{\mu}^\dagger \mathbf{X}\} \right). \end{aligned} \quad (14)$$

To make progress we now need to express the term in the exponent proportional to ρ_d , in a quadratic form in terms of \mathbf{X} , \mathbf{Y} . This can be done by using identity 3 and introducing $\nu \times \nu$ matrices \mathcal{R} , \mathcal{T} . Thus the last term in the exponent of (14) becomes

$$\exp \left[\frac{\rho_d}{2n_t} \text{Tr}\{\mathbf{Y}^\dagger \mathbf{Q} \mathbf{Y} \mathbf{X}^\dagger \mathbf{X}\} \right] = \int d\mu(\mathcal{T}, \mathcal{R}) \exp \left[\text{Tr}\{\mathcal{T} \mathcal{R}\} - \sqrt{\frac{\rho_d}{4n_t}} \text{Tr}\{\mathcal{T} \mathbf{Y}^\dagger \mathbf{Q} \mathbf{Y} + \mathbf{X}^\dagger \mathbf{X} \mathcal{R}\} \right] \quad (15)$$

The application of identity 3 and the introduction of matrices \mathcal{R} and \mathcal{T} is a particular form of a Hubbard–Stratonovich transformation. The usefulness of this method is in that it allows the integration of certain quantities (\mathbf{X} , \mathbf{Y}) of limited relevance, and introduces auxiliary quantities (such as \mathcal{R} and \mathcal{T}), which will prove to have particular importance in the final answer. This is in a sense a mean-field theory approach, which will end up being exact in the large- n limit.

Combining (14), (15) and using identity 1, we can now integrate \mathbf{X} , \mathbf{Y} , resulting in

$$g(\nu) = \int d\mu(\mathcal{T}, \mathcal{R}) e^{-\mathcal{S}} \quad (16)$$

where

$$\begin{aligned} \mathcal{S} &= \log \det \left(\mathbf{I}_{n_r} \otimes \left[\mathbf{I}_\nu + \sqrt{\frac{\rho_d}{n_t}} \mathcal{R} \right] \right) - \text{Tr}\{\mathcal{T} \mathcal{R}\} \\ &\quad + \log \det \left(\mathbf{I}_{n_t} \otimes \mathbf{I}_\nu + \sqrt{\frac{\rho_d}{n_t}} \mathbf{Q} \otimes \mathcal{T} + \frac{\rho_c}{n_t} \mathbf{Q}^{1/2} \boldsymbol{\mu}^\dagger \boldsymbol{\mu} \mathbf{Q}^{1/2} \otimes \left[\mathbf{I}_\nu + \sqrt{\frac{\rho_d}{n_t}} \mathcal{R} \right]^{-1} \right). \end{aligned} \quad (17)$$

In the above, ν is still a positive integer, which should be taken to zero following assumption 1. However, before doing this, we will first take the limit of large antenna numbers $n_t, n_r \gg 1$. In this limit the saddle-point method of evaluating the integral (described below) becomes accurate. Subsequently, we will take the $\nu \rightarrow 0^+$ limit.

Assumption 2 (interchanging 2 limits). *The limits $n \rightarrow \infty$ and $\nu \rightarrow 0^+$ in evaluating $g(\nu)$ in (16) can be interchanged by first taking the former and then the latter without changing the final answer.*

As we shall see below, the two limits of large antenna numbers and small ν are related: higher terms in the expansion in ν involve successively higher terms in a $1/n$ expansion.

We will now use the saddle-point method to calculate the above integral. In the case of (16) and (17), when $n_t, n_r \gg 1$, the exponent \mathcal{S} is nominally of order $\mathcal{O}(\nu n)$. Following assumption 2 and thus keeping ν fixed we may apply the saddle-point approximation to calculate the integral of (16) and then take the $\nu \rightarrow 0^+$ limit. It should be stressed that for a fixed positive integer ν , the saddle-point analysis of \mathcal{S} and $g(\nu)$ is a straightforward exercise in asymptotic analysis. The only additional complexity to the standard textbook treatment of this topic [26] is that \mathcal{S} involves integrals over multiple variables (the elements of \mathcal{T}, \mathcal{R}).

Rather than search for saddle-point solutions over all possible complex \mathcal{T} and \mathcal{R} matrices, we are going to invoke the following hypothesis without proof.

Assumption 3 (replica invariance). *The relevant saddle-point solution for (16) involves matrices \mathcal{T}, \mathcal{R} , which are invariant in replica space and thus are proportional to the identity matrix \mathbf{I}_ν .*

The above hypothesis, used heavily in physics [15, 16], basically states that there is no preferred direction in the space of replicas, and thus if any saddle-point solution is valid, so is any unitary transformation thereof in replica space. Although we are not going to provide a proof here, it should be noted that in [11] it is shown that the results obtained by this method are identical to those using a systematic expansion.

2.1. Saddle-point analysis

The assumed form of \mathcal{T} and \mathcal{R} at the saddle point is $t\sqrt{n_t}\mathbf{I}_\nu$ and $r\sqrt{n_r}\mathbf{I}_\nu$, respectively. The extra factor of $\sqrt{n_t}$ has been included for convenience, as will become evident below. To consider the vicinity around the saddle point, we thus rewrite \mathcal{T}, \mathcal{R} as

$$\begin{aligned}\mathcal{T} &= t\sqrt{n_t}\mathbf{I}_\nu + \delta\mathbf{T} \\ \mathcal{R} &= r\sqrt{n_r}\mathbf{I}_\nu + \delta\mathbf{R}\end{aligned}\quad (18)$$

where $\delta\mathbf{T}, \delta\mathbf{R}$ are $\nu \times \nu$ matrices representing deviations around the saddle point. One can then expand \mathcal{S} of (17) in a Taylor series of increasing powers of $\delta\mathbf{T}, \delta\mathbf{R}$ as follows:

$$\mathcal{S} = \mathcal{S}_0 + \mathcal{S}_1 + \mathcal{S}_2 + \mathcal{S}_3 + \mathcal{S}_4 + \dots \quad (19)$$

with \mathcal{S}_p containing p th-order terms in $\delta\mathbf{T}, \delta\mathbf{R}$. These terms can be obtained explicitly by differentiating (17):

$$\mathcal{S}_0 = \nu \left[n_r \log(1 + \sqrt{\rho_d}r) - n_t r t + \text{Tr} \left\{ \log \left(\mathbf{I}_{n_r} + \sqrt{\rho_d}t\mathbf{Q} + \frac{\rho_c}{n_t} \frac{\mathbf{Q}\mathbf{u}^\dagger\mathbf{u}}{1 + r\sqrt{\rho_d}} \right) \right\} \right] = \nu\Gamma \quad (20)$$

$$\begin{aligned}\mathcal{S}_1 &= \left(\frac{\sqrt{\rho_d}}{1 + r\sqrt{\rho_d}} \left[n_r - \text{Tr} \left\{ \frac{\frac{\rho_c}{n_t}\mathbf{Q}\mathbf{u}^\dagger\mathbf{u}}{(1 + r\sqrt{\rho_d})(\mathbf{I}_{n_r} + \sqrt{\rho_d}t\mathbf{Q}) + \frac{\rho_c}{n_t}\mathbf{Q}\mathbf{u}^\dagger\mathbf{u}} \right\} - n_t t \right] \frac{\text{Tr}\{\delta\mathbf{R}\}}{\sqrt{n_r}} \right. \\ &\quad \left. + \left(\text{Tr} \left\{ \frac{\sqrt{\rho_d}(1 + r\sqrt{\rho_d})\mathbf{Q}}{(1 + r\sqrt{\rho_d})(\mathbf{I}_{n_r} + \sqrt{\rho_d}t\mathbf{Q}) + \frac{\rho_c}{n_t}\mathbf{Q}\mathbf{u}^\dagger\mathbf{u}} \right\} - n_t r \right) \frac{\text{Tr}\{\delta\mathbf{T}\}}{\sqrt{n_t}} \right)\end{aligned}\quad (21)$$

$$\mathcal{S}_2 = \frac{1}{2} \text{Tr} \left\{ \begin{bmatrix} \delta\mathbf{R} \\ \delta\mathbf{T} \end{bmatrix}^T \Sigma \begin{bmatrix} \delta\mathbf{R} \\ \delta\mathbf{T} \end{bmatrix} \right\}, \quad (22)$$

where the 2×2 Hessian Σ is given by

$$\begin{aligned}\Sigma_{11} &= -\frac{\rho_d}{n_t} \left[\frac{n_r - n_t}{(1 + r\sqrt{\rho_d})^2} + \text{Tr} \left\{ \frac{(\mathbf{I}_{n_r} + \sqrt{\rho_d}t\mathbf{Q})^2}{((1 + r\sqrt{\rho_d})(\mathbf{I}_{n_r} + \sqrt{\rho_d}t\mathbf{Q}) + \frac{\rho_c}{n_t}\mathbf{Q}\mathbf{u}^\dagger\mathbf{u})^2} \right\} \right] \\ \Sigma_{22} &= -\frac{\rho_d}{n_t} \text{Tr} \left\{ \frac{(1 + \sqrt{\rho_d}r)^2\mathbf{Q}^2}{((1 + r\sqrt{\rho_d})(\mathbf{I}_{n_r} + \sqrt{\rho_d}t\mathbf{Q}) + \frac{\rho_c}{n_t}\mathbf{Q}\mathbf{u}^\dagger\mathbf{u})^2} \right\} \\ \Sigma_{12} = \Sigma_{21} &= -\left[1 - \frac{\rho_d\rho_c}{n_t^2} \text{Tr} \left\{ \frac{\mathbf{Q}^2\boldsymbol{\mu}^\dagger\boldsymbol{\mu}}{((1 + r\sqrt{\rho_d})(\mathbf{I}_{n_r} + \sqrt{\rho_d}t\mathbf{Q}) + \frac{\rho_c}{n_t}\mathbf{Q}\mathbf{u}^\dagger\mathbf{u})^2} \right\} \right].\end{aligned}\quad (23)$$

For $p > 2$ the expanded terms can be written as

$$\begin{aligned}\mathcal{S}_p &= -\frac{(-1)^p}{p} \left(\frac{\rho_d}{n_t} \right)^{p/2} \left[\frac{n_r - n_t}{(1 + r\sqrt{\rho_d})^p} \text{Tr}\{(\delta\mathbf{R})^p\} \right. \\ &\quad \left. + \text{Tr} \left\{ \frac{((\mathbf{I}_{n_r} + t\sqrt{\rho_d}\mathbf{Q})\delta\mathbf{R} + (1 + r\sqrt{\rho_d})\mathbf{Q}\delta\mathbf{T})^p}{((1 + r\sqrt{\rho_d})(\mathbf{I}_{n_r} + \sqrt{\rho_d}t\mathbf{Q}) + \frac{\rho_c}{n_t}\mathbf{Q}\mathbf{u}^\dagger\mathbf{u})^p} \right\} \right].\end{aligned}\quad (24)$$

The saddle-point solution of (16) and hence the corresponding values of t, r are found by demanding that \mathcal{S} is stationary with respect to variations in \mathcal{T}, \mathcal{R} [26], resulting in $\mathcal{S}_1 = 0$. This produces the following saddle-point equations:

$$\frac{r}{\sqrt{\rho_d}(1 + r\sqrt{\rho_d})} = \frac{1}{n_t} \text{Tr} \left\{ \frac{\mathbf{Q}}{(1 + r\sqrt{\rho_d})(\mathbf{I}_{n_r} + \sqrt{\rho_d}t\mathbf{Q}) + \frac{\rho_c}{n_t}\mathbf{Q}\mathbf{u}^\dagger\mathbf{u}} \right\} \quad (25)$$

$$\frac{t(1 + r\sqrt{\rho_d})}{\sqrt{\rho_d}} = \frac{n_r}{n_t} - \frac{1}{n_t} \text{Tr} \left\{ \frac{\frac{\rho_c}{n_t}\mathbf{Q}\mathbf{u}^\dagger\mathbf{u}}{(1 + r\sqrt{\rho_d})(\mathbf{I}_{n_r} + \sqrt{\rho_d}t\mathbf{Q}) + \frac{\rho_c}{n_t}\mathbf{Q}\mathbf{u}^\dagger\mathbf{u}} \right\}. \quad (26)$$

It is interesting to note that the solutions to the above two equations extremize Γ for real and positive t, r . It is important to note that for generic full rank matrices $\mathbf{Q}, \mathbf{u}^\dagger\mathbf{u}$, both r and t are generally of order unity, $r, t = \mathcal{O}(1)$. Thus the expansion coefficients multiplying the terms $\delta\mathbf{T}^p, \delta\mathbf{R}^p$, etc are generally of order $\mathcal{O}(n^{1-p/2})$, successively decreasing in size for increasing p .

The small parameter controlling this approximation is therefore $n^{-1/2}$, making this saddle-point solution increasingly accurate for large n . Therefore, the aim of this analysis is to calculate successively higher order terms in $n^{-1/2}$ and classify each resulting quantity in terms of their powers of ν to the appropriate cumulant moment \mathcal{C}_p in (11). This matching of powers of ν implicitly assumes that the expansion in (11) is valid for integer ν , as described in assumption 1. Thus, for example, terms of orders $\mathcal{O}(\nu n)$ and $\mathcal{O}(\nu/n)$ will both contribute to $\langle I \rangle$, as we shall see.

2.2. Ergodic mutual information

We start with the leading term of $g(\nu)$ in the saddle-point approximation: $g(\nu) \approx \exp(-\mathcal{S}_0) = \exp(-\nu\Gamma)$, where Γ is evaluated in (21) using (25) and (26). We thus see from (10), (11) that the leading term in the expansion of $\langle I \rangle$ is Γ and note that $\langle I \rangle = \mathcal{O}(n)$.

2.3. Variance of the mutual information

To obtain the $\mathcal{O}(v^2)$ term in the expansion of $\log g(v)$ in (11) we need to include the next non-vanishing term, \mathcal{S}_2 . Thus, for the moment we neglect higher order terms \mathcal{S}_p for $p > 2$. Noting the measure-preserving transformation $\mathcal{T}, \mathcal{R} \rightarrow \delta\mathbf{T}, \delta\mathbf{R}$ of (18) we have

$$\begin{aligned} g(v) &= e^{-S_0} \int d\mu(\delta\mathbf{T}, \delta\mathbf{R}) e^{-S_2} \\ &= e^{-S_0} \int d\mu(\delta\mathbf{T}, \delta\mathbf{R}) \exp\left(-\frac{1}{2} \sum_{a,b=1}^v [\delta T_{1,ab} \delta R_{1,ab}] \Sigma [\delta T_{1,ba} \delta R_{1,ba}]^T\right). \end{aligned} \quad (27)$$

To diagonalize the exponent of the above equation, we rotate each pair $\delta\mathbf{T}, \delta\mathbf{R}$ to a new basis of $v \times v$ -dimensional matrices \mathbf{W}_1 and \mathbf{W}_2 . In particular, $[W_{1,ab} W_{2,ab}]^T = \mathbf{U}[\delta T_{1,ab} \delta R_{1,ab}]^T$, where \mathbf{U} is an orthogonal matrix such that $\mathbf{U}\mathbf{V}\mathbf{U}^T$ is a diagonal matrix with the diagonal given by $\mathbf{v} = [v_1 v_2]^T$, the vector of eigenvalues of Σ . We may now rewrite (27) as

$$g(v) = e^{-S_0} \int d\mu(\mathbf{W}_1, \mathbf{W}_2) \exp\left(-\frac{1}{2} \sum_{a,b=1}^v [v_1 W_{1,ab} W_{1,ba} + v_2 W_{2,ab} W_{2,ba}]\right). \quad (28)$$

We now take appropriate paths to integrate $W_{1,2,ab}$, resulting in

$$g(v) = e^{-S_0} |v_1 v_2|^{-\frac{v^2}{2}} = e^{-S_0} |\det \Sigma|^{-\frac{v^2}{2}}. \quad (29)$$

Comparing (11) to (29) and matching order by order the terms of the v -Taylor expansion of the exponent of $g(v)$, we can identify the leading term in the variance of the mutual information to be

$$\langle I^2 \rangle - \langle I \rangle^2 = \mathcal{C}_2 = -\log |\det \Sigma| + \dots \quad (30)$$

We note that since $\Sigma_{ij} = \mathcal{O}(1)$, the variance is also $\mathcal{O}(1)$ in the expansion of $n^{-1/2}$ when both n_t and n_r are of the same order. (However, if n_r is fixed while n_t increases, we find that $\mathcal{C}_2 = \mathcal{O}(n^{-1})$, in agreement with [27].) Also we see that no term proportional to v is produced from \mathcal{S}_2 . Thus no term of $\mathcal{O}(1)$ in the antenna number appears in $\langle I \rangle$, resulting in $\langle I \rangle = \Gamma + o(1)$.

2.4. Higher order terms

To obtain higher order corrections in the small parameter $n^{-1/2}$, we need to take into account the terms \mathcal{S}_p for $p > 2$ in (24). Details of the method can be found elsewhere [13]. For simplicity, here we will briefly describe how to set up the perturbation expansion and discuss its implications to the distribution of the mutual information for large n .

We define an expectation bracket of an arbitrary operator $O(\delta\mathbf{T}, \delta\mathbf{R})$ as

$$\langle\langle O \rangle\rangle = |\det \Sigma|^{v^2/2} \int d\mu(\delta\mathbf{T}, \delta\mathbf{R}) e^{-S_2} O(\delta\mathbf{T}, \delta\mathbf{R}). \quad (31)$$

The integration over $\delta\mathbf{T}, \delta\mathbf{R}$ is performed as described in the previous section. We can obtain the expectations of quadratic terms in $\delta\mathbf{T}, \delta\mathbf{R}$, written here in a compact form as follows:

$$\langle\langle [\delta R_{ab} \delta T_{ab}] [\delta R_{cd} \delta T_{cd}]^T \rangle\rangle = \delta_{ad} \delta_{bc} \Sigma^{-1}. \quad (32)$$

In addition, the expectation of any odd power of $\delta\mathbf{T}, \delta\mathbf{R}$ must vanish by symmetry. As a result, only integer powers of $1/n$ survive in the perturbative expansion.

With this bracket notation we can rewrite

$$g(v) = e^{-S_0} e^{-S_2} e^{-(S_3+S_4+\dots)} \quad (33)$$

$$= e^{-v\Gamma} |\det \Sigma|^{-v^2/2} \left\langle \left(1 - (S_3 + S_4 + \dots) + \frac{1}{2}(S_3 + S_4 + \dots)^2 + \dots \right) \right\rangle \quad (34)$$

At this point, it is interesting to count powers of n in the various terms of the expansion. Using simple power counting arguments, we see that the term $\langle\langle S_p \rangle\rangle$ is of order $n^{-p/2+1}$, but vanishes for p odd. Also, $\langle\langle S_p S_q \rangle\rangle$ is of order $n^{-p/2-q/2+2}$ but is zero for $p+q$ odd, and so forth. By regrouping the terms in the above expansion by their order in $1/n$ we obtain the following expansion:

$$g(v) = e^{-S_0} [|\det \Sigma|^{-v^2/2}] [1 + D_1 + D_2 + D_3 + \dots] \quad (35)$$

where

$$\begin{aligned} D_1 &= \langle\langle S_4 + \frac{1}{2} S_3^2 \rangle\rangle \\ D_2 &= \langle\langle S_6 + \frac{1}{2} S_4^2 + S_3 S_5 + \frac{1}{24} S_3^4 \rangle\rangle \\ &\vdots \end{aligned} \quad (36)$$

Here we have regrouped all terms which are of order $1/n^p$ into the term D_p . Thus, for example, D_1 contains all terms of order $1/n$. As in [13], we can evaluate the averages above using (32) and Wick's theorem. We find that D_1 produces $1/n$ corrections to the mean mutual information and its skewness. All additional terms D_p provide higher order corrections in $1/n$. We thus see that in the large n limit only the first two moments of the distribution are finite. Therefore, asymptotically the mutual information distribution approaches a Gaussian.

2.5. Summary of results

To summarize, we have derived the following results, with Γ given by (20), Σ given by (23), and with the parameters r and t given by (25) and (26),

$$\langle I \rangle = \Gamma + \mathcal{O}(1/n) \quad (37)$$

$$\mathcal{C}_2 = \text{Var}(I) = \langle I^2 \rangle - \langle I \rangle^2 = -\log |\det \Sigma| + \mathcal{O}(1/n^2) \quad (38)$$

$$\mathcal{C}_3 = \text{Sk}(I) = \mathcal{O}(1/n). \quad (39)$$

The expansion can be continued to higher order in $1/n$ straightforwardly. We thus see that for large n the distribution of the mutual information approaches a Gaussian. Similar results have been obtained for other types of statistics of the Gaussian matrix \mathbf{G} [13].

3. Capacity-achieving signal covariance \mathbf{Q}

As discussed in section 1.2, instead of instantaneous channel information, the transmitter has statistical information for the channel \mathbf{G} , namely only $\boldsymbol{\mu}$ and ρ are known. Based on this information, the signal covariance can be optimized to maximize a particular metric of the mutual information distribution. Here we describe how to optimize \mathbf{Q} , in order to maximize the average mutual information, keeping only the $\mathcal{O}(n)$ term in (21), i.e. to find $\max_{\mathbf{Q}} \langle I \rangle$, in the large antenna limit.

We start by observing that Γ depends on \mathbf{Q} through the last term in (21). Expressing the determinant of this term in the eigen-basis of $\boldsymbol{\mu}^\dagger \boldsymbol{\mu}$, it can be written as

$$\det \left(\mathbf{I}_{n_t} + \tilde{\mathbf{Q}} \left(\sqrt{\rho_d} t \mathbf{I}_{n_t} + \frac{\rho_c \tilde{\mathbf{M}}}{n_t (1 + r \sqrt{\rho_d})} \right) \right) \quad (40)$$

where $\tilde{\mathbf{M}}$ is a diagonal matrix with the eigenvalues of $\boldsymbol{\mu}^\dagger \boldsymbol{\mu} / n_t$, given by μ_k^2 for $k = 1, \dots, n_t$ on the diagonals, where μ_k are the singular values of $\boldsymbol{\mu} / \sqrt{n_t}$, and $\tilde{\mathbf{Q}}$ is the original matrix \mathbf{Q} expressed in the eigen-basis of $\boldsymbol{\mu}^\dagger \boldsymbol{\mu}$. We now use the fact that for any non-negative definite matrix \mathbf{A} , $\det(\mathbf{A}) \leq \prod_k A_{kk}$, where A_{kk} are the diagonal elements [2]. Applying this inequality to (40) we get

$$\det \left(\mathbf{I}_{n_t} + \tilde{\mathbf{Q}} \left(\sqrt{\rho_d t} \mathbf{I}_{n_t} + \frac{\rho_c}{n_t(1+r\sqrt{\rho_d})} \tilde{\mathbf{M}} \right) \right) \leq \prod_{k=1}^{n_t} \left(1 + \tilde{Q}_{kk} \left(\sqrt{\rho_d t} + \frac{\rho_c \mu_k^2}{(1+r\sqrt{\rho_d})} \right) \right) \tag{41}$$

with equality when $\tilde{\mathbf{Q}}$ is diagonal. Thus the \mathbf{Q} maximizing Γ is simultaneously diagonalizable with $\boldsymbol{\mu}^\dagger \boldsymbol{\mu}$.

Once the optimal eigen-basis of \mathbf{Q} has been determined to be the same as $\boldsymbol{\mu}^\dagger \boldsymbol{\mu}$, one needs to find its optimal eigenvalues q_k for $k = 1, \dots, n_t$. Γ has to be optimized subject to the power constraint $\text{Tr } \mathbf{Q} = n_t$. This constraint is enforced by adding a Lagrange multiplier to Γ , i.e.

$$\Gamma \rightarrow \Gamma - \Lambda \left(\sum_{k=1}^{n_t} q_k - n_t \right). \tag{42}$$

Incorporating the Lagrange multiplier to (21) and maximizing, it is easy to see that the optimal eigenvalues of \mathbf{Q} are then given by

$$q_k = \left[\frac{1}{\Lambda} - \frac{1+r\sqrt{\rho_d}}{\sqrt{\rho_d t} (1+r\sqrt{\rho_d}) + \rho_c \mu_k^2} \right]_+, \tag{43}$$

where $[x]_+ = \{x + \text{sgn}(x)\}/2$. Here, $\Lambda > 0$ is determined by imposing the power constraint

$$\text{Tr } \mathbf{Q} = \sum_{k=1}^{n_t} q_k = n_t \tag{44}$$

Solving (43), (44) together with (26), (25) allows us to calculate $\langle I \rangle$ and $\text{Var}(I)$ in (37), (38) to obtain the ergodic capacity and the variance of the distribution around it. Note that the optimization over \mathbf{Q} at the transmitter is based on statistical rather than instantaneous information about the channel. Therefore it depends on statistical quantities ($\boldsymbol{\mu}^\dagger \boldsymbol{\mu}$ and ρ rather than \mathbf{G} itself). As a result, the transmitter needs to be updated about the channel information at a relatively slow rate.

4. Validity of Gaussian approximation $\mathcal{N}(\langle I \rangle, \text{Var}(I))$

In section 2.5 we saw that in the limit of large n , the distribution of the mutual information approaches a Gaussian with mean equal to Γ and variance the calculated variance of I , in the sense that all higher moments and corrections tend to zero. Surprisingly, this approximation is valid for even a small number of antennas. We demonstrate this property by comparing numerically the Gaussian distribution $\mathcal{N}(\langle I \rangle, \text{Var}(I))$ calculated using (21) and (38) with the simulated distribution resulting from the generation of a large number of random matrix realizations. This comparison can be seen in figure 1, where indeed the agreement not only to the Gaussian distribution but to its correct means and variances is striking for both small

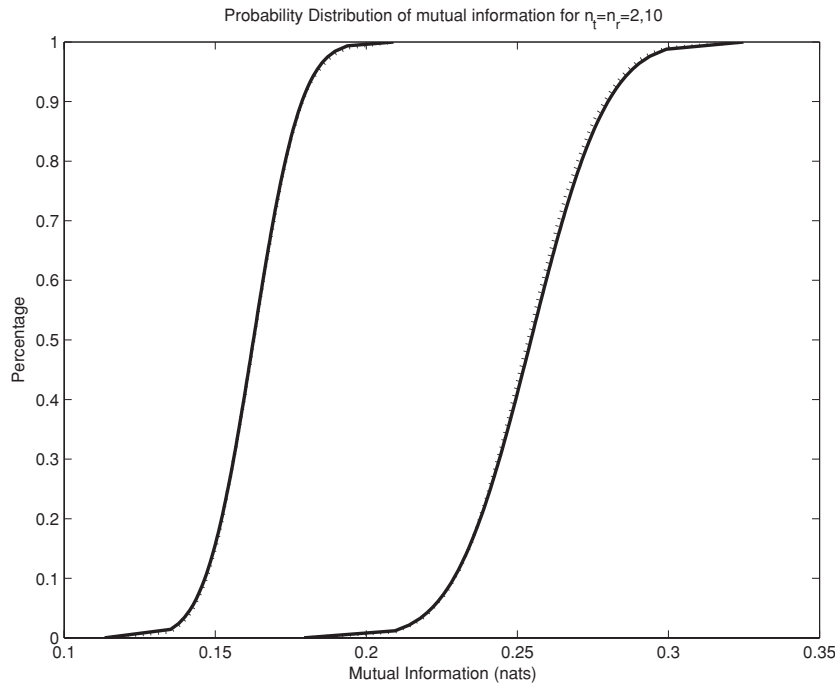


Figure 1. Cumulative distribution (CDF) of mutual information per antenna (I/n_t) for $n_r = n_t = 2, 10$. The signal to noise ratio is set to $\rho = 1$, while the diffuse component is 50% of the power $\rho_d = \rho_c = 0.5$. The dotted lines correspond to the numerically generated curves, while the solid lines are the theoretical ones, generated as Gaussian distributions with mean and variance the ones obtained by applying the methods of this paper. We see that the agreement is very good. We also see that for $n = 10$ the distribution is more narrowly peaked. In both cases the non-zero mean component of the channel (μ) was generated randomly from a Gaussian distribution. Among the pairs corresponding to the same number of antennas, those to the left have the transmission covariance matrix \mathbf{Q} optimized with respect to the known $\mu^\dagger \mu$, while those to the right simply have $\mathbf{Q} = \mathbf{I}_{n_r}$.

$n = 2$ and large $n = 10$ antenna numbers. This allows us to accurately calculate not only the ergodic capacity but also the outage capacity, defined in (6).

5. Conclusion

In conclusion, we have presented an analytic approach to calculate the statistics of the mutual information of MIMO systems for the case of Ricean statistics. To this end we applied tools developed for the analysis of mesoscopic systems, such as replicas random matrix theory. In addition, we have used this method to find the optimal signal covariance \mathbf{Q} and thus analytically calculate the capacity when the statistics of the (Gaussian) channel are known at the transmitter. These methods, although formally valid for large antenna numbers, apply with very high accuracy to arrays with only few number of antennas. This allows us to accurately evaluate the outage capacity for any number of antennas. We demonstrated this by comparing to numerical simulations.

This analytic approach provides the framework and a simple tool to accurately analyse the statistics of throughput of even small arrays. It is a simple example where mesoscopic physics methods can have technological applications.

Acknowledgments

We wish to acknowledge enlightening discussions with Harold Baranger and Anirvan Sengupta. After this work was completed we became aware of a work [28] that using different methods provides analytic results only for the average of the mutual information in the infinite antenna limit.

Appendix. Complex integrals

Identity 1. Let \mathbf{M} be a Hermitian positive-definite square $m \times m$ matrix, and let \mathbf{X} be a complex $m \times n$ matrix. Then

$$(\det \mathbf{M})^{-n} = \int \mathbf{D}\mathbf{X} e^{-\frac{1}{2} \text{Tr}\{\mathbf{X}^\dagger \mathbf{M} \mathbf{X}\}} \quad (\text{A.1})$$

where the integration measure $\mathbf{D}\mathbf{X}$ is given by (1).

Identity 2. Let \mathbf{X} , \mathbf{A} , \mathbf{B} be $m \times n$ complex matrices. Then, the following equality holds:

$$\int \mathbf{D}\mathbf{X} \exp\left(-\frac{1}{2} \text{Tr}\{\mathbf{X}^\dagger \mathbf{X} + \mathbf{A}^\dagger \mathbf{X} - \mathbf{X}^\dagger \mathbf{B}\}\right) = \exp\left(-\frac{1}{2} \text{Tr}\{\mathbf{A}^\dagger \mathbf{B}\}\right). \quad (\text{A.2})$$

Identity 3 (Hubbard–Stratonovich transformation). Let \mathbf{U} , \mathbf{V} be arbitrary complex $v \times v$ matrices, where v is assumed to be an arbitrary positive integer. Then the following identity holds:

$$\exp[-\text{Tr}\{\mathbf{U}\mathbf{V}\}] = \int d\mu(\mathbf{T}, \mathbf{R}) \exp[\text{Tr}\{\mathbf{R}\mathbf{T} - \mathbf{U}\mathbf{T} - \mathbf{R}\mathbf{V}\}]. \quad (\text{A.3})$$

In the above equation, the auxiliary matrices \mathbf{T} , \mathbf{R} are general complex matrices $v \times v$ and their integration measure is given by (2). The integration of the elements of \mathbf{R} and \mathbf{T} is along contours in complex space parallel to the real and imaginary axis.

References

- [1] Foschini G J and Gans M J 1998 On limits of wireless communications in a fading environment when using multiple antennas *Wirel. Pers. Commun.* **6** 311–35
- [2] Telatar I E 1999 Capacity of multi-antenna gaussian channels *Eur. Trans. Telecommun. Relat. Technol.* **10** 585–96
- [3] Wolniansky P W, Foschini G J, Golden G D and Valenzuela R A 1998 V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel *URSI Int. Symp. Signals Syst. Electronics* pp 295–300
- [4] Ling J, Chizhik D, Wolniansky P and Valenzuela R 2001 Multiple transmit multiple receive (MTMR) capacity survey in Manhattan *IEEE Electronics Lett.* **37** 1041–2
- [5] Moustakas A L, Baranger H U, Balents L, Sengupta A M and Simon S H 2000 Communication through a diffusive medium: coherence and capacity *Science* **287** 287–90 (Preprint cond-mat/0009097)
- [6] Verdú S and Shamai S 1999 Spectral efficiency of CDMA with random spreading *IEEE Trans. Inform. Theory* **45** 622–40
- [7] Rapajic P B and Popescu D 2000 Information capacity of a random signature multiple-input multiple-output channel *IEEE Trans. Commun.* **48** 1245
- [8] Lozano A and Tulino A M 2002 Capacity of multiple-transmit multiple-receive antenna architectures *IEEE Trans. Inform. Theory* **48** 3117–28
- [9] Tse D N and Hanly S V 1999 Linear multiuser receivers: effective interference, effective bandwidth and user capacity *IEEE Trans. Inform. Theory* **45** 641
- [10] Chuah C N, Tse D, Kahn J and Valenzuela R A 2002 Capacity scaling in MIMO wireless systems under correlated fading *IEEE Trans. Inform. Theory* **48** 637

- [11] Sengupta A M and Mitra P P 1999 Distributions of singular values for some random matrices *Phys. Rev. E* **60** 3389–92
- [12] Sengupta A M and Mitra P P 2000 Capacity of multivariate channels with multiplicative noise: I. Random matrix techniques and large- n expansions for full transfer matrices *Preprint physics/0010081*
- [13] Moustakas A L, Simon S H and Sengupta A M 2003 MIMO capacity through correlated channels in the presence of correlated interferers and noise: a (not so) large N analysis *IEEE Trans. Inform. Theory* **45** 2545–61
- [14] Edwards S F and Anderson P W 1975 Theory of spin glasses *J. Phys. F: Met. Phys.* **5** 965–74
- [15] Mézard M, Parisi G and Virasoro M A 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)
- [16] Itzykson C and Drouffe J-M 1989 *Statistical Field Theory* (Cambridge: Cambridge University Press)
- [17] Surlas N 1989 Spin glass models as error correcting codes *Nature* **339** 693–5
- [18] Montanari A and Surlas N 2000 Statistical mechanics of turbo codes *Eur. Phys. J. B* **18** 107–19
- [19] Montanari A 2000 Turbo codes: the phase transition *Eur. Phys. J. B* **18** 121–36
- [20] Tanaka T 2002 A statistical-mechanics approach to large-system analysis of CDMA multiuser detectors *IEEE Trans. Inform. Theory* **48** 2888–910
- [21] Kang M and Alouini M-S 2002 Capacity of MIMO Rician channels *Proc. 40th Annual Conference on Communication, Control, and Computing (Monticello, IL)*
- [22] Simon S H, Moustakas A L and Marinelli L 2004 Capacity and character expansions: moment generating function and other exact results for MIMO correlated channels (submitted) (*Preprint cs.IT/0509080*)
- [23] Simon S H and Moustakas A L 2004 Eigenvalue density of correlated random Wishart matrices *Phys. Rev. E* **69** (*Preprint math-ph/0401038*)
- [24] Ozarow L H, Shamai S and Wyner A D 1994 Information theoretic considerations for cellular mobile radio *IEEE Trans. Veh. Technol.* **43** 359–78
- [25] Simon S H, Moustakas A L, Stoytchev M and Safar H 2001 Communication in a disordered world *Phys. Today* September 38–43
- [26] Bender C M and Orszag S A 1978 *Advanced Mathematical Methods for Scientists and Engineers* (New York: McGraw-Hill)
- [27] Hochwald B M, Marzetta T L and Tarokh V 2004 Multi-antenna channel hardening and its implications for rate feedback and scheduling *IEEE Trans. Inform. Theory* **50** 1893–909
- [28] Dumont J, Loubaton P, Lasaulce S and Debbah M 2005 On the asymptotic performance of MIMO correlated Rician channels *IEEE Int. Conf. Acoustics, Speech and Signal Processing*