Outage Mutual Information of Space-Time MIMO Channels*

Zhengdao Wang
Dept. of Elec. and Comp. Eng.
Iowa State University
2215 Coover Hall
Ames, IA 50011-3060
zhengdao@iastate.edu

Georgios B. Giannakis
Dept. of Elec. and Comp. Eng.
University of Minnesota
200 Union Street SE
Minneapolis, MN 55455
georgios@ece.umn.edu

Abstract

We present analytical expressions for the probability density function (PDF) of the random mutual information between transmitted and received vector signals of a random space-time independent and identically distributed (i.i.d.) multiple-input multiple-output (MIMO) channel, assuming that the transmitted signals from the multiple antennas are Gaussian i.i.d.. We show that this PDF can be well approximated by a Gaussian distribution, and such a Gaussian approximation is based on expressions for the given PDF's mean and variance that we derive. We prove that at high SNR, every 3 dB increase in signal to noise ratio (SNR) leads to an increase in outage rate approximately equal to $\min(M, N)$, where M and N denote the number of transmit- and receive-antennas, respectively. A simple expression for the moment generating function of the mutual information PDF is also provided, based on which we establish normality of the PDF, when both M and N are large, and the SNR is large.

1 Introduction

Employing multiple antennas at the transmitter and/or the receiver end has attracted much attention recently because of the potential for very high spectral efficiencies [4,8,9]. A problem of great interest is therefore the ultimate limit of transmission rate achievable when deploying multiple antennas: the capacity. In wireless links, the fading multiple-input multiple-output (MIMO) channel is modeled as random. When the channel variation is slow, the channel can be estimated relatively accurately at the receiver. By assuming perfect channel state information (CSI) at the receiver but no CSI at the transmitter, it has been shown that the (ergodic) capacity is achieved by sending uncorrelated circularly symmetric zero mean complex Gaussian signals of equal power from all the transmit antennas [8]. When CSI is known also at the transmitter, water-filling type of power allocation achieves the capacity [8]. When CSI is neither known at the transmitter nor the receiver, capacity-achieving signaling matrix has derived in [5].

Let M denote the number of transmit-antennas, and N the number of receive-antennas. Suppose the transmitter does not have CSI, and thus transmits complex

*This work was supported by the NSF Wireless Initiative Grant No. 99-79443, the NSF Grant No. 01-0516, and by the ARL/CTA Grant No. DAAD 19-01-2-011.

Gaussian signals with equal power splitting among the M transmit antennas, while the receiver has perfect CSI. For each realization of the random MIMO channel, there is an associated mutual information between the transmitted and the received signals. This mutual information, conditioned on the channel realization, is a random variable due to the randomness of the channel. The probability that the conditional mutual information is less than a given rate R is termed the *mutual information outage probability.* Finding this outage probability, as a function of R, is therefore equivalent to finding the probability distribution function of the conditional mutual information. The outage probability for the case of one transmit or one receive antenna is relatively easy to derive; see [8]. When there are multiple transmit *and* multiple receive antennas, Monte-Carlo simulations have been presented in [4], together with some lower bound results. Recently, some asymptotic results have been obtained for a large number of transmit and/or receive antennas, where it is shown that the asymptotic mutual information is Gaussian distributed [5].

In this paper, we consider an arbitrary number of transmit and receive antennas, and derive *exact* formulas for the outage probability that can be numerically evaluated. We also derive some asymptotic results for large signal-to-noise ratio (SNR), but for a *finite* number of transmit and receive antennas.

Notation: Bold face lowercase (uppercase) letters denote column vectors (matrices); $()^{\mathcal{H}}$ denotes Hermitian conjugate; log() denotes the base-two logarithm, and ln() denotes the natural logarithm; $\det[a_{ij}]_{ij=0}^{k-1}$ is the determinant of the $k \times k$ matrix whose (i, j)th entry, starting from $(0, 0)$th, is a_{ij}; $\mathrm{E}[\cdot]$ denotes expectation; $tr(\mathbf{A})$ is the trace of a matrix $\mathbf{A}$; and $\Gamma(x) := \int_0^\infty e^{-t} t^{x-1}\, \mathrm{d}t$.

2 System Model and Existing Results

Consider a MIMO channel with M inputs and N outputs. Let $\mathbf{s}$ be an $M \times 1$ vector denoting the transmitted symbols, $\mathbf{x}$ an $N \times 1$ vector denoting the received symbols, $\mathbf{n}$ an $N \times 1$ noise vector with independent zero-mean complex Gaussian noise entries with variance $N_0/2$ per real component, and $\mathbf{H}$ an $N \times M$ complex mixing matrix, known to the receiver. The MIMO channel can be represented as

$$\mathbf{x} = \mathbf{H}\mathbf{s} + \mathbf{n}. \tag{1}$$

Assuming zero-mean complex Gaussian $\mathbf{s}$ with autocorrelation matrix $\mathbf{R}_{ss} = \mathrm{E}[\mathbf{s}\mathbf{s}^{\mathcal{H}}]$, the mutual information between $\mathbf{s}$ and $\mathbf{x}$, conditioned on $\mathbf{H}$ is given by[1] [8]

$$I(\mathbf{H}, \mathbf{R}_{ss}) := I(\mathbf{s}, \mathbf{x}|\mathbf{H}) = \log\det(\mathbf{I}_N + \frac{1}{N_0}\mathbf{H}\mathbf{R}_{ss}\mathbf{H}^{\mathcal{H}}) = \log\det(\mathbf{I}_M + \frac{1}{N_0}\mathbf{R}_{ss}\mathbf{H}^{\mathcal{H}}\mathbf{H}). \tag{2}$$

The (average or ergodic) capacity C of the channel is defined as

$$C = \max_{\mathbf{R}_{ss}} \mathrm{E}[I(\mathbf{H}, \mathbf{R}_{ss})], \quad \text{subject to } \mathrm{tr}(\mathbf{R}_{ss}) = \mathcal{P}, \tag{3}$$

where $\mathcal{P}$ is the total transmit power, and the expectation is with respect to the channel $\mathbf{H}$. From now on, we assume that:
(AS1) $\mathbf{H}$ *has i.i.d. zero-mean complex Gaussian entries of unit variance.*

[1] We distinguish between the concept of mutual information and capacity: we use the term capacity only in the Shannon sense. Our terminology is the same as that of [8]; it is therefore different from, e.g., [1,4,5], where $I(\mathbf{H}, \mathbf{R}_{ss})$ with $\mathbf{R}_{ss} = (E_s/M)\mathbf{I}_M$ is referred to as the capacity for the MIMO channel.

It is shown in [8] that the ergodic capacity for an $\mathbf{H}$ obeying AS1) is achieved by $\mathbf{R}_{ss} = (P/M)\mathbf{I}_M$. The capacity is also evaluated as an integral in [8].

When the transmitter does not have CSI and assumes $\mathbf{R}_{ss} = (\mathcal{P}/M)\mathbf{I}_M$, we use $\mathcal{I}$ to denote the mutual information between $\mathbf{s}$ and $\mathbf{x}$:

$$\mathcal{I} = \log\det(\mathbf{I}_N + \frac{\rho}{M}\mathbf{H}^{\mathcal{H}}\mathbf{H}) = \sum_{i=0}^{k-1}\log(1+\lambda_i\rho/M), \tag{4}$$

where λ_i is the ith unordered eigenvalue of $\mathbf{H}\mathbf{H}^{\mathcal{H}}$, starting from index $i = 0$. As a function of the random channel $\mathbf{H}$, $\mathcal{I}$ is a random variable. Let $\rho := \mathcal{P}/N_0$ denote the SNR per receive antenna. The *mutual information outage probability* or simply outage probability $P_{\text{out}}(R; M, N, \rho)$ is defined as

$$P_{\text{out}}(R; M, N, \rho) := \Pr[\mathcal{I} < R], \tag{5}$$

from which it is clear that the outage probability is nothing but the probability distribution function of the random variable $\mathcal{I}$. Using the matrix identity $\det(\mathbf{I}_N + \mathbf{A}\mathbf{B}) = \det(\mathbf{I}_M + \mathbf{B}\mathbf{A})$, where $\mathbf{A}$ and $\mathbf{B}$ are $N \times M$ and $M \times N$ matrices, respectively, it is easy to verify that under AS1)

$$P_{\text{out}}(R; M, N, \rho) = P_{\text{out}}(R; N, M, \frac{N}{M}\rho), \tag{6}$$

which shows the asymmetry between M and N. The reason is that the total receive SNR increases when N increases, but remains the same when M increases.

Some existing results about the mutual information, when both M and N are greater than 1, are presented below (when $M = 1$ or $N = 1$, the derivation is simple; see [4]):

i) *Bounds on mutual information [4].* When $M > N$,

$$\sum_{k=M-N+1}^{M}\log[1+(\rho/M)\cdot\chi^2_{2k}] < \mathcal{I} < \sum_{i=1}^{M}\log[1+(\rho/M)\cdot\chi^2_{2N,i}], \tag{7}$$

where $\{\chi^2_{2k} : k = M-N+1, \dots, M\}$ are independent chi-square variates with the given degrees of freedom as subscripts, and $\{\chi^2_{2N,i} : i = 1, \dots, M\}$ are independent chi-square variates each with $2N$ degrees of freedom. Equation (7) should be interpreted in a probabilistic sense: Two random variables A and B satisfy inequality $A < B$ if $\Pr(B < \tau) < \Pr(A < \tau)$, for any τ. When $M = N = n$, it can be shown that as $n \to \infty$, the lower bound, i.e., the left hand side of (7), divided by n, converges to a constant:

$$n^{-1}\sum_{k=1}^{n}\log[1+(\rho/n)\cdot\chi^2_{2k}] \to (1+\rho^{-1})\log(1+\rho) - \log(e), \quad \text{as } n \to \infty. \tag{8}$$

It has been deduced from this lower bound analysis and numerical tests (Monte-Carlo simulations) that for large SNR, there is about n bits per cycle increase in the rate R for each 3dB increase in SNR, for a fixed outage probability [4].

ii) *Asymptotic distribution of $\mathcal{I}$ for large M and/or large N* [5]. For all the four following cases, the distribution of $\mathcal{I}$ can be well approximated by a Gaussian distribution: i) large M and fixed N; ii) large N and fixed M; iii) large M, large N, and low SNR; and iv) large M, large N, and high SNR.

3 Main Results

Instead of looking only at asymptotic cases, we will derive in this section an expression for the outage probability that can be numerically evaluated accurately and quickly for practical values of M and N, beyond which the asymptotic results of [5] becomes accurate. We will present our result in theorems, whose proofs are collected in the appendix.

3.1 PDF of outage mutual information

Theorem 1 [Outage probability] *Let $k = \min(M, N)$, and $d = \max(M, N) - k$. The moment generating function (MGF) of $\mathcal{I}$ is*

$$\Phi_I(s) := E[e^{-\mathcal{I}s}] = B^{-1} \det[\mathbf{G}(s)], \tag{9}$$

where $B = \prod_{i=1}^{k} \Gamma(d+i)$, and $\mathbf{G}(s)$ is a $k \times k$ Hankel matrix whose (i,j)th entry is

$$g_{ij}(s) = \int_0^\infty \left(1 + \frac{\rho}{M}\lambda\right)^{-s} \lambda^{i+j+d} e^{-\lambda}\, d\lambda, \quad i, j = 0, \ldots, k-1. \tag{10}$$

The probability density function (PDF) of $\mathcal{I}$ is the inverse Laplace transform of $\Phi_I(s)$, and the outage probability is given by the inverse Laplace transform of $s^{-1}\Phi_I(s)$:

$$P_{out}(R; M, N, \rho) = \frac{1}{2\pi j} \int_{\sigma - j\infty}^{\sigma + j\infty} s^{-1} \Phi_I(s) e^{sR}\, ds, \tag{11}$$

where σ is a fixed positive number.

Notice that Theorem 1 neither assumes M or N large, nor it requires the SNR to be small or large. It offers an effective way of evaluating the outage probability for practical values of M and N. The integral in (10) can be evaluated either numerically or by using its hypergeometric series expansion of it obtained (using Mathematica^{TM}) as:

$$\begin{aligned} g_{ij}(s) =& \pi\Gamma^{-1}(s)\, \csc[\pi\ (d+i+j-s)] \cdot \\ & \left[(\rho/M)^{-1-d-i-j}\, \Gamma(1+d+i+j)\, {}_1\bar{F}_1(1+d+i+j, 2+d+i+j-s, M/\rho) - \right. \\ & \left. (\rho/M)^{-s}\, \Gamma(s)\, {}_1\bar{F}_1(s, -d-i-j+s, M/\rho)\right], \end{aligned}$$

where $\csc(z) = 1/\sin(z)$; ${}_1\bar{F}_1$ is the *regularized confluent hypergeometric function*: ${}_1\bar{F}_1(a,b,z) = {}_1F_1(a,b,z)/\Gamma(b)$, and ${}_1F_1$ stands for the *confluent hypergeometric function* [3]. The inverse Laplace transform in (11) can also be evaluated numerically using efficient algorithms such as [2], a MATLAB implementation of which [6], was used in the ensuing examples.

As an application example, we depict in Fig. 1 the PDF of $\mathcal{I}$ for $M = N = 3$, and different SNR values from 0 to 20 dB. We have the following observations:

i) the PDF takes a Gaussian-like shape for all the SNR values ρ considered;

ii) naturally, the mean increases with SNR, and the increase for high SNR is linearly proportional to the SNR increase in dB;

iii) the variance also increases with SNR, and at high SNR, the variance converges to a fixed number;

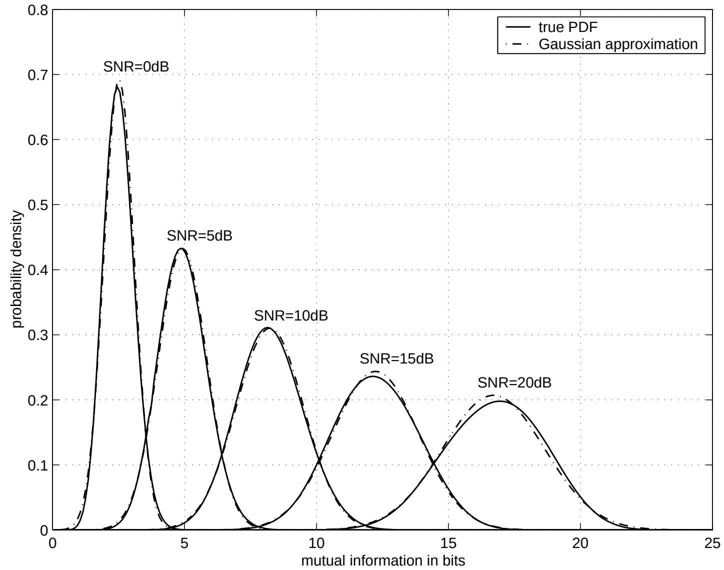

Figure 1: PDF of $\mathcal{I}$ for $M = N = 3$ and different SNR's

iv) the mess of the PDF is mostly above a certain level (e.g., for SNR=20dB, the PDF is nearly zero for $\mathcal{I} < 10$ bits)

Observation iv) suggests that although the capacity in the Shannon sense does not exist per channel realization, the outage probability can be quite small for reasonably high rates.

3.2 Outage Rate

For a fixed level of outage probability ϵ, we define the *outage rate* as the rate R for which the outage probability is at the given level ϵ; i.e.,

$$R_{\mathrm{out}}(\epsilon; M, N, \rho) := \arg_R \left[P_{\mathrm{out}}(R; M, N, \rho) = \epsilon\right]. \tag{12}$$

In order to find the outage rate for a given ϵ using Theorem 1, we can perform a bisection search over a starting interval, on whose left (right) boundary the outage probability takes values less (greater) than ϵ. As an example, we plot in Fig. 2 the outage rate for $\epsilon = 10\%$, for different SNR values and different number of antennas.

The following remarks are due on this example:

i) the slope of the curves converges to a constant at high SNR (the curves tend to straight lines);

ii) the asymptotic slope increases, when we increase both M and N. The numerical values corresponding to the graph are reported in Table 1.

We have also supported these observations analytically, as we summarize next (see the appendix for proof):

Theorem 2 *At high SNR, the slope of the outage rate curve (number of bits versus SNR in dB) is* $\log(10^{0.1})\min(M, N)$ *bits per dB, or, about* $0.9966 \cdot \min(M, N)$ *bits per 3 dB.*

We remark that the results we obtained in Fig. 2 match well with a similar result in [1, Fig. 2] that was obtained by Monte-Carlo simulation. This corroborates the validity of Theorem 1.

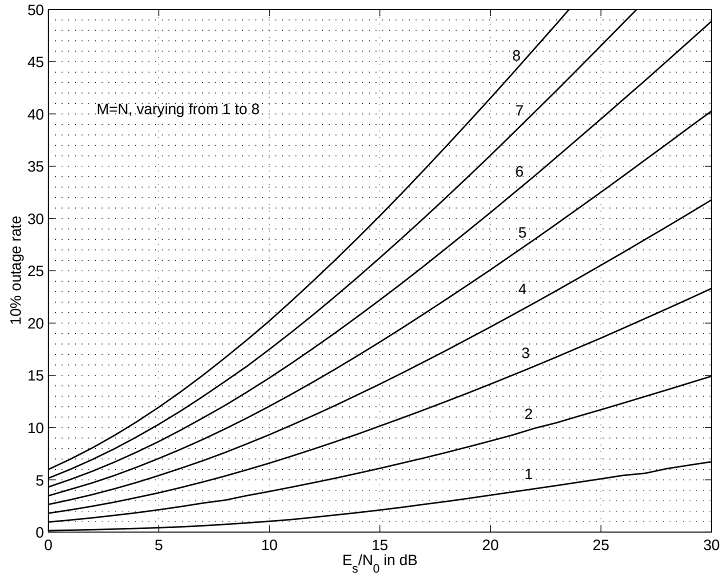

Figure 2: 10% outage rate for different SNR and $M = N = 1, 2, \ldots, 8$

3.3 Gaussian Approximation

Based on the observation that the PDF of $\mathcal{I}$ resembles that of a Gaussian random variable, we approximate it by a Gaussian distribution with equal mean and variance, as summarized in the following theorem.

Theorem 3 *Define the functions* $\tilde{\phi}_i(\lambda) := [i!/(i+d)!]^{\frac{1}{2}} L_i^d(\lambda)\lambda^{\frac{d}{2}} e^{-\frac{\lambda}{2}}$*, where* $i = 0, 1, \ldots,$ *and* $L_i^d(\lambda) := \frac{1}{i!} e^{\lambda}\lambda^{-d} \frac{d^i}{d\lambda^i}(e^{-\lambda}\lambda^{d+i})$ *is the associated Laguerre polynomial of order* i*. Define also* $K(x, y) := \sum_{i=0}^{k-1} \tilde{\phi}_i(x)\tilde{\phi}_i(y)$*.. The Gaussian approximation to the PDF of* $\mathcal{I}$ *in* (11) *is given by* $p(\mathcal{I}; M, N, \rho) \approx (\sqrt{2\pi}\sigma_I)^{-1} \exp[-(\mathcal{I} - \mu_I)^2/(2\sigma_I^2)]$*, where* $\mu_I = \int_0^\infty \log(1 + \lambda\rho/M)K(\lambda, \lambda)\, d\lambda$ *and*

$$\begin{aligned}\sigma_I^2 = &\int_0^\infty \log^2(1 + \lambda\rho/M)K(\lambda, \lambda)\, d\lambda \\ &- \int_0^\infty \int_0^\infty \log(1 + \lambda_1\rho/M)\log(1 + \lambda_2\rho/M)K^2(\lambda_1, \lambda_2)\, d\lambda_1\, d\lambda_2.\end{aligned} \tag{13}$$

To illustrate the usefulness of Theorem 3, we show in Fig. 1 the Gaussian approximations to the PDF. We can see that for virtually all SNR values, the Gaussian fit is surprisingly good. The Gaussianity can be intuitively explained by a central limit argument: the mutual information $\mathcal{I}$ can be written as a sum of $\log(\cdot)$ random variables that are correlated (see (4)). Should they be uncorrelated, we would be able to apply the classical central limit theorem directly. But here the eigenvalues are statistically dependent. The proof for normality for large M or large N, or, when both M and N are large and the SNR is either very small or very large, can be found in [5]. The missing part in the whole picture is when both M and N are large, but the SNR is neither too small nor too large.

4 Discussion

It has been shown that at high SNR and for large M and large N, the mutual information is approximately Gaussian distributed [5]. The proof in [5] resorts to a Lyapunov condition for a triangular array type of central limit theorem. It is possible to achieve

$M=N$	0dB	5dB	10dB	15dB	20dB	25dB	30dB
1	0.15	0.41	1.04	2.11	3.53	5.10	6.73
2	0.97	2.13	3.89	6.10	8.72	11.71	14.91
3	1.80	3.76	6.60	10.15	14.16	18.57	23.30
4	2.64	5.41	9.32	14.16	19.63	25.53	31.78
5	3.48	7.04	12.04	18.19	25.10	32.52	40.28
6	4.31	8.67	14.76	22.21	30.57	39.52	48.87
7	5.15	10.31	17.48	26.24	36.04	46.53	57.45
8	5.99	11.94	20.21	30.25	41.52	53.52	66.05

Table 1: 10% outage rate for different SNR's from 0 to 30 dB and $M = N = 1, 2, \ldots, 8$

the same goal of proving normality by noticing that for large ρ the MGF of the mutual information in Theorem 1 can be written as a product of gamma functions:

$$\Phi_I(s) \approx B^{-1}(\rho/M)^{-sk} \prod_{i=0}^{k-1} \Gamma(i + d - s + 1). \tag{14}$$

This can be obtained by invoking the approximation $(1 + \lambda_i \rho/M) \approx \lambda_i \rho/M$, factoring out all the elements in the first row of matrix $\mathbf{G}(s)$, and performing row reduction. Let us now consider the *polygamma functions* [3]

$$\psi_0(z) = -\gamma + \sum_{j=1}^{z-1} \frac{1}{j}, \quad \text{for integer } z \tag{15}$$

$$\psi_n(z) = \frac{d^{n+1}}{dz^{n+1}} \ln[\Gamma(z)] = (-1)^{n+1} n! \sum_{j=0}^{\infty} \frac{1}{(z+j)^{n+1}} = (-1)^{n+1} n! \zeta(n+1, z), \quad n \geq 1 \tag{16}$$

for $z = i + d + 1$, where $\zeta(a, z)$ is the *Hurwitz zeta function* and $\gamma = 0.5772\ldots$ is the Euler-Mascheroni constant [3]. Taking the logarithm of (14), we can write $\ln[\Phi_I(s)]$ as

$$\ln[\Phi_I(s)] \approx -sk \ln(\rho/M) + \sum_{n=1}^{\infty} c(n) s^n, \tag{17}$$

where the coefficient $c(n)$ for s^n can be expressed using the polygamma functions as:

$$c(1) = k\gamma - \sum_{i=0}^{k-1} \sum_{j=1}^{i} \frac{1}{j+d} \tag{18}$$

$$c(n) = \frac{1}{n} \sum_{i=0}^{k-1} \zeta(n, i+d+1) = \frac{1}{n} \sum_{i=0}^{k-1} \sum_{j=0}^{\infty} \frac{1}{(i+j+d+1)^n}, \quad n > 1. \tag{19}$$

By approximating the series using integrals, it can be shown that $\{c(n)\}$ converge to some constant as $k \to \infty$, for any $n > 2$. For $n = 2$, however, using the following asymptotics [5]

$$\zeta(2, n) = \frac{1}{n} + R_n, \quad \sum_{i=1}^{k} \frac{1}{i} = \ln(k) + \gamma + o(1), \tag{20}$$

where R_n is negligible compared with $1/n$, we infer that $c(2)$ grows logarithmically with k: $c(2) = \ln(k+d) - \sum_{i=1}^{d} \frac{1}{i} + \gamma + o(1)$. For $n = 1$, it can be seen by definition that $-c(1)$ grows in the same order as $\int_d^{k+d} \ln(x)\,\mathrm{d}x$, which is even larger in magnitude than $c(2) \approx \ln(k+d) - \ln(d)$, for large k. Therefore, the dominant terms in $\ln[\Phi_I(s)]$ are those of s and s^2. Or, for large k, $\Phi_I(s) \approx \exp[(-k\ln(\rho/M) + c(1))s + c(2)s^2]$, which is the MGF of a Gaussian PDF with mean $-c(1) + k\ln(\rho/M)$, and variance $2c(2)$. This proves the normality of the PDF of interest.

The difference in the proof given here relative to the one in [5] is that the complete MGF of $\mathcal{I}$ is given by (17) which not only allows for a more intuitive understanding of the nature of the problem of interest, but also enables estimating the speed of convergence to normality as k grows. The formulas for the mean and the variance in this analysis are less important because the exact values for them (instead of the high SNR approximants) have been already given in Theorem 3.

The fact that for large k, the mean is much larger than the square root of the variance is a sign of "channel hardening" [5]. It shows that when both M and N get large, the capacity behaves more and more like a deterministic quantity. In this sense, the Shannon capacity exists, and is infinite in the limit of infinitely many antennas at both sides.

5 Conclusions

We have derived analytic expressions for the probability density function of the random mutual information between the transmitted and the received vector signals of a random space-time i.i.d. MIMO channel, assuming that the transmitted signals from the multiple antennas are independent and Gaussian. The outage probability is also given in a closed integral form. The expressions are easy to evaluate numerically for any practical setup. As an application example, we presented the PDF of a 3×3 MIMO channel for SNR ranging from 0 to 20 dB. It was observed that the PDF can be well approximated by Gaussian densities. We have given formulas for the mean and the variance of the mutual information, which allow for accurate and simple Gaussian approximations to the PDF.

Outage rate can also be evaluated by numerically solving a single variable equation. We showed as an example the outage rate for $n \times n$ MIMO channels with n varying from 1 to 8, for SNR ranging from 0 to 30 dB. We established that at high SNR, the outage rate increases about $\min(M, N)$ bits for every 3 dB increase in SNR, with M transmit and N receive antennas.

A Proof of Theorem 1

To prove Theorem 1, we first provide several lemmas. Define the functions $\tilde{\phi}_i(\lambda) := [i!/(i+d)!]^{\frac{1}{2}} \mathrm{L}_i^d(\lambda) \lambda^{\frac{d}{2}} e^{-\frac{\lambda}{2}}$, where $i = 0, 1, ...$, and $\mathrm{L}_i^d(\lambda) := \frac{1}{i!} e^{\lambda} \lambda^{-d} \frac{d^i}{d\lambda^i}(e^{-\lambda}\lambda^{d+i})$ is the associated Laguerre polynomial of order i. Let $\int^k := \underbrace{\int_0^\infty \int_0^\infty \cdots \int_0^\infty}_{k}$ and $\mathrm{d}^k x := \mathrm{d}x_0 \ldots \mathrm{d}x_{k-1}$.

Lemma 1 [8] [Joint PDF of eigenvalues of $\mathbf{H}\mathbf{H}^{\mathcal{H}}$] *Let* $\mathbf{H}$ *be a* $k \times n$ *matrix, where* $n \geq k$. *The joint PDF of the unordered eigenvalues of* $\mathbf{H}\mathbf{H}^{\mathcal{H}}$ *is*

$$p(\lambda_0, \ldots, \lambda_{k-1}) = C_{n,k} \prod_i e^{-\lambda_i} \lambda_i^{n-k} \prod_{i<j} (\lambda_i - \lambda_j)^2 = \frac{1}{k!} \det{}^2 [\tilde{\phi}_i(\lambda_j)]_{ij=0}^{k-1} \qquad (21)$$

where $C_{n,k}$ *is a normalizing factor.* □

Lemma 2 *Let $\{f_i(x) : i = 0, 1, \ldots, N-1\}$ be N functions for which the following integrals exist (i.e.,they decay fast enough at infinity). It holds that*

$$\int^k \det{}^2[f_i(x_j)]_{ij=0}^{k-1}\, d^k x = k! \det[F_{ij}]_{ij=0}^{k}, \tag{22}$$

where $F_{ij} = \int_0^\infty f_i(x) f_j(x)\, dx$.

Proof: Gram's result states the following [7, Appendix A.13]: Let $\mathbf{v}_i$, $i = 1, 2, \ldots, k$, be m vectors and let $v_{i\alpha}$, $\alpha = 1, 2, \ldots, n$, be their components along some basis. Form the scalar products $b_{ij} = (\mathbf{v}_i, \mathbf{v}_j) = \sum_{\alpha=1}^n v_{i\alpha} v_{j\alpha}^*$, $i, j = 1, 2, \ldots, m$; where $(\cdot)^*$ stands for conjugate, then $\det[b_{ij}]_{ij=1}^m = \frac{1}{m!}\sum_{\alpha_1,\alpha_2,\ldots,\alpha_m} \det{}^2[v_{i\alpha_j}]_{ij=1}^m$, where on the right hand side the sum is over all possible ways of choosing $\alpha_1, \alpha_2, \ldots, \alpha_m$ among $1, 2, \ldots, m$. The desired result (22) can be proved by making the correspondences: $m \to k, v_{i\alpha} \to f_i(x)$, and $\sum_\alpha \to \int_0^\infty$ dx. The conjugate in the scalar product is immaterial and can be dropped even for complex functions $f_i(\cdot)$'s, because Gram's result is an algebraic one that does not depend on the scalar product properties. It can be noticed that the range of integration is not essential in the statement of the lemma: any interval would have worked equally well. □

We now prove the theorem itself. Consider the *moment generating function* of $\mathcal{I}$; that is, $\Phi_I(s) := \mathrm{E}[e^{-\mathcal{I}s}]$. Using the PDF of Lemma 1, we have

$$\Phi_I(s) = C_{M,N} \int^k \prod_i e^{-\lambda_i} \lambda_i^d (1 + \lambda_i \rho/M)^{-s} \prod_{i<j} (\lambda_i - \lambda_j)^2\, \mathrm{d}^k\lambda,$$

where $C_{M,N}$ is a normalizing constant. Now writing $\prod_{i<j}(\lambda_i - \lambda_j)$ as the squared Vandermonde determinant $\det[\lambda_i^j]_{ij=0}^{k-1}$, and defining $f_{i,s}(\lambda) = [e^{-\lambda}\lambda^d(1 + \lambda\rho/M)^{-s}]^{\frac{1}{2}}\lambda^i$, we can rewrite $\Phi_I(s)$ as $C_{M,N}\det^2[f_{i,s}(\lambda_j)]_{ij=0}^{k-1}$. Applying Lemma 2, we obtain $\Phi_I(s) = (k! C_{M,N}) \det[g_{ij}(s)]_{ij=0}^{k-1}$, where $g_{ij}(s) = \int_0^\infty \left(1 + \frac{\rho}{M}\lambda\right)^{-s} \lambda^{i+j+d} e^{-\lambda}\, \mathrm{d}\lambda, i, j = 0, \ldots, k-1$. In order to figure out the constant in $\Phi_I(s)$, we notice that $\Phi_I(0) = 1$, because the integral of any PDF is one. We have $g_{ij}(0) = \Gamma(i + j + d + 1)$, and $B := \det[\Gamma(i + j + d + 1)]_{ij=0}^{k-1} = \prod_{i=0}^{k-1}\Gamma(i + d + 1)$, which is the product of the first row elements in the matrix (it can be proved by factoring out the first row elements from each column followed by row reduction). Therefore, $k! C_{M,N} = B^{-1}$ and (9) is thus proved.

To prove (11), we only need to notice that the Laplace transform of $\int_\infty^x f(x) dx$ is the Laplace transform of $f(x)$, divided by s. □

B Proof of Theorem 2

An intuitively appealing proof is the following: For large ρ, the term ρ/M in (4) will dominate the sum $1 + \lambda_i \rho/M$ as long as λ_i is non-zero. Therefore, for large ρ, $\mathcal{I} \approx k \log(\rho) + \sum_{i=0}^{k-1} \log(\lambda_i/M)$. As a result, every dB increase in ρ will only result in an increase of $k \log(10^{0.1})$ bits in the mean of $\mathcal{I}$. Therefore, for any fixed outage probability ϵ, the outage rate at high SNR will increase $k \log(10^{0.1})$ bits for every dB increase in SNR. Equivalently, since $k := \min(M, N)$, this amounts to about $0.9966 \min(M, N)$ or just $\min(M, N)$ bits increase for every 3 dB increase in SNR.

This intuitive proof relies on the fact that the probability of λ_i being "small" is negligible. A more rigorous proof can be established by noticing that in (10) the constant 1 can be omitted for large ρ. Then for every dB increase in ρ, $g_{ij}(s)$ will be scaled by $10^{s/10}$, which means the determinant will be scaled by $10^{sk/10}$, or $e^{(k \ln 10^{0.1})s}$. It follows from (11) that the outage rate will increase by $k \ln 10^{0.1}$ nats, or, $k \log(10^{0.1})$ bits. □

C Proof of Theorem 3

Since the Gaussian approximant matches the PDF of $\mathcal{I}$ in the mean and variance, we only need to derive the mean μ_I and and variance σ_I^2. We need the following lemma.

Lemma 3 [Joint PDF of a few eigenvalues] *Let the joint distribution of $\{\lambda_i\}_{i=0}^{k-1}$ be $p(\lambda_0, \dots, \lambda_{k-1}) = \frac{1}{k!} \det{}^2[\phi_i(\lambda_j)]_{ij=0}^{k-1}$, where $\phi_i(\lambda)$ are orthonormal functions such that $\int \phi_i(\lambda)\phi_j(\lambda) = \delta_{ij}$. Define $K(x,y) := \sum_{i=0}^{k-1} \phi_i(x)\phi_i(y)$. Then the joint density of any m of the k eigenvalues ($m \leq k$), say, $\lambda_0, \dots, \lambda_{m-1}$, is $p(\lambda_0, \dots, \lambda_{m-1}) = \frac{(k-m)!}{k!} \det[K(\lambda_i, \lambda_j)]_{ij=0}^{m-1}$.*

Proof: The proof is essentially the same as the derivation in [7, Sec. 6.1.1], except for some normalizations. □

Specializing the result in Lemma 3 to the case of Lemma 1, we only need to replace $\phi_i(\cdot)$ by $\tilde{\phi}_i(\cdot)$, where the latter has been defined at the beginning of this appendix. Specifically, the densities for one and two eigenvalues are, respectively,

$$p(\lambda) = \frac{1}{k} K(\lambda, \lambda), \quad \text{and} \;\; p(\lambda_0, \lambda_1) = \frac{1}{k(k-1)} [K(\lambda_0, \lambda_0)K(\lambda_1, \lambda_1) - K^2(\lambda_0, \lambda_1)], \quad (23)$$

where in the last equation we have used that $K(x,y) = K(y,x)$.

Now, using (4), we have

$$\mu_I = \mathrm{E}\left[\sum_{i=0}^{k-1} \log(1 + \lambda_i \rho/M)\right] = k\mathrm{E}[\log(1 + \lambda_0\rho/M)] = \int_0^\infty \log(1 + \lambda\rho/M) K(\lambda, \lambda)\, \mathrm{d}\lambda,$$

and σ_I^2 can be obtained using $p(\lambda_0, \lambda_1)$ and Lemma 2.

Notice that the marginal density $p(\lambda)$ and μ_I have also been derived in [8]. □

References

[1] G. Bauch and J. Hagenauer, "Multiple antenna systems: capacities, transmit diversity and turbo processing," in *Proc. 4th Int'l ITG Conf. Source and Channel Coding*, Berlin, Germany, Jan. 2002, pp. 387–398.

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

[3] Eric Weisstein's World of Mathematics (MathWorld^{TM}), http://mathworld.wolfram.com.

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

[5] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multi-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Transactions on Information Theory*, May 2002 (submitted).

[6] K. J. Hollenbeck, "INVLAP.M: A matlab function for numerical inversion of Laplace transforms by the de Hoog algorithm," *unpublished work*, 1998; downloadable from `http://www.isva.dtu.dk/staff/karl/invlap.htm`

[7] M. L. Mehta, *Random matrices*, Academic Press, San Diego, 1st edition, 1967.

[8] I. Emre Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. on Telecomm.*, vol. 10, no. 6, pp. 585–596, Nov.–Dec. 1999.

[9] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. ISSSE-98*, Pisa, Italy, Sept. 29 1998.