ASYMPTOTIC OUTAGE CAPACITY OF MULTIANTENNA CHANNELS

Antonia M. Tulino

Universitá di Napoli
80125 Napoli, Italy
atulino@ee.princeton.edu

Sergio Verdú

Princeton University
Princeton, NJ 08540,USA
verdu@ee.princeton.edu

ABSTRACT

This paper characterizes the asymptotic distribution of the input-output mutual information of multiantenna channels. Using recent results on random matrix theory, we prove asymptotic normality of the unnormalized mutual information for arbitrary signal-to-noise ratios and fading distributions, allowing for correlation between the antennas at either transmitter or receiver.

1. INTRODUCTION

Most information-theoretic asymptotic studies of multiantenna channels have focused on the (ergodic) capacity, the fundamental operational limit in the regime in which a codeword spans many realizations of the fading coefficients. The quantity of interest is the mutual information averaged over such coefficients.

In the non-ergodic regime, where the fading is not such that the statistics of the channel are revealed to the receiver during the span of a codeword, the outage capacity (cumulative distribution of the mutual information) is of interest. Results on large random matrices (e.g. [1]) show that the mutual information per receive antenna converges a.s. (almost surely) to its expectation as the number of antennas goes to infinity (with a given ratio of transmit to receive antennas). Thus, as the number of antennas grows, a self-averaging mechanism hardens the mutual information to its expected value. However, the non-normalized mutual information still suffers random fluctuations that, although small with respect to the mean, are of vital interest in the study of the outage capacity. In this paper, we show the asymptotic normality of those random fluctuations and we characterize their variance for arbitrary signal-to-noise ratios.

For fixed numbers of antennas, the computation of mean and variance of the mutual information of the canonical channel (i.i.d. zero-mean Gaussian entries) was carried out in [2, 3], which laid out numerical evidence of an excellent Gaussian fit. These results were extended to Ricean fading in [4] and to correlated antennas (at the array with the most antennas) in [5]. Arguments supporting the normality of the c.d.f (cumulative distribution function) of the mutual information for large numbers of antennas were given in [6, 7, 8].[1] Ref. [6] used the replica method from statistical physics (which has yet to find a rigorous justification) and [7] showed the asymptotic normality only in the asymptotic regimes of low and high signal-to-noise ratios. Finally, in [8], the normality of the outage capacity is proved for the canonical channel using [9]. Our results differ in that:

- They allow for arbitrary fading distributions (not only Rayleigh) and for correlation between either transmit or receive antennas.
- They hold for arbitrary signal-to-noise ratios.
- They are rigorously proved using asymptotic random matrix tools.
- Succinct expressions for the asymptotic variance of the mutual information are given.

2. MATHEMATICAL BACKGROUND

Given an $N \times N$ Hermitian matrix $\mathbf{A}$, the empirical c.d.f of the eigenvalues (also referred to as the empirical spectral distribution (ESD)) of $\mathbf{A}$ is defined as

$$\mathsf{F}^N_{\mathbf{A}}(x) = \frac{1}{N}\sum_{i=1}^{N} 1\{\lambda_i(\mathbf{A}) \leq x\} \tag{1}$$

where $\lambda_1(\mathbf{A}), \ldots, \lambda_N(\mathbf{A})$ are the eigenvalues of $\mathbf{A}$ and $1\{\cdot\}$ is the indicator function. If $\mathsf{F}^N_{\mathbf{A}}(\cdot)$ converges a.s as $N \to \infty$, then the corresponding limit (asymptotic ESD) is denoted by $\mathsf{F}_{\mathbf{A}}(\cdot)$.

For our purposes, it is advantageous to make use of the η-transform and the Shannon transform, which were motivated by the application of random matrix theory to various problems in the information theory of noisy communication channels [1]. These transforms, intimately related with each other and with the Stieltjes transform traditionally used in random matrix theory [1], characterize the spectrum of a random matrix while carrying certain engineering intuition.

[1] For additional references, see [1]

Definition 1 *Given a nonnegative definite random matrix* $\mathbf{A}$*, its* η*-transform is*

$$\eta_{\mathbf{A}}(\gamma) = \mathbb{E}\left[\frac{1}{1+\gamma X}\right] \tag{2}$$

where X *is a nonnegative random variable whose distribution is the asymptotic ESD of* $\mathbf{A}$ *while* γ *is a nonnegative real number. Thus,* $0 < \eta_X(\gamma) \leq 1$.

Then, $\mathbb{E}[X^k]$ can be regarded as the kth asymptotic moment of $\mathbf{A}$, i.e., $\lim_{N\to\infty} \frac{1}{N}\mathrm{tr}\{\mathbf{A}^k\}$. As a consequence, $\eta_{\mathbf{A}}(\gamma)$ can be regarded as a generating function for the asymptotic moments of $\mathbf{A}$.

Definition 2 *Given a nonnegative definite random matrix* $\mathbf{A}$*, its Shannon transform is defined as*

$$\mathcal{V}_{\mathbf{A}}(\gamma) = \mathbb{E}[\log(1+\gamma X)] \tag{3}$$

where X *is a nonnegative random variable whose distribution is the asymptotic ESD of* $\mathbf{A}$ *while* γ *is a nonnegative real number.*

The rationale for introducing the Shannon transform is that it gives the Shannon capacity of various noisy coherent channels.

Theorem 1 *[1] If the entries of* $\mathbf{H}$ *are zero-mean i.i.d. with variance* $\frac{1}{N}$*, as* $K, N \to \infty$ *with* $\frac{K}{N} \to \beta$*, the ESD of* $\mathbf{H}\mathbf{H}^\dagger$ *converges a.s. to the Marčenko-Pastur law whose density function is*

$$\tilde{\mathsf{f}}_\beta(x) = (1-\beta)^+\,\delta(x) + \frac{\sqrt{(x-a)^+(b-x)^+}}{2\pi x} \tag{4}$$

while the η*- and Shannon transforms are*

$$\eta_{\mathbf{H}\mathbf{H}^\dagger}(\gamma) = 1 - \frac{\mathcal{F}(\gamma,\beta)}{4\,\gamma} \tag{5}$$

and

$$\begin{aligned}\mathcal{V}_{\mathbf{H}\mathbf{H}^\dagger}(\gamma) = {} & \beta\log\left(1+\gamma-\frac{\mathcal{F}(\gamma,\beta)}{4}\right) \\ & + N\log\left(1+\gamma\beta-\frac{\mathcal{F}(\gamma,\beta)}{4}\right) \\ & -\frac{N\log e}{4\,\gamma}\mathcal{F}(\gamma,\beta)\end{aligned} \tag{6}$$

with

$$\mathcal{F}(x,z) = \left(\sqrt{x(1+\sqrt{z})^2+1} - \sqrt{x(1-\sqrt{z})^2+1}\right)^2$$

Using the η-transform, we reformulate the following result from [10] in terms of the η-transform.

Theorem 2 *[1] Let* $\mathbf{H}$ *be an* $N \times K$ *complex random matrix whose entries are i.i.d. with variance* $\frac{1}{N}$*. Let* $\mathbf{T}$ *be a* $K \times K$ *nonnegative definite nonrandom matrix, whose ESD converges to a proper c.d.f. The ESD of* $\mathbf{H}\mathbf{T}\mathbf{H}^\dagger$ *converges a.s., as* $K, N \to \infty$ *with* $\frac{K}{N} \to \beta$*, to a distribution whose* η*-transform satisfies*

$$\beta = \frac{1-\eta}{1-\eta_{\mathbf{T}}(\gamma\eta)} \tag{7}$$

where we have compactly abbreviated $\eta_{\mathbf{H}\mathbf{T}\mathbf{H}^\dagger}(\gamma) = \eta$*. The corresponding Shannon transform is*

$$\mathcal{V}_{\mathbf{H}\mathbf{T}\mathbf{H}^\dagger}(\gamma) = \beta\mathcal{V}_{\mathbf{T}}(\eta\gamma) + \log\frac{1}{\eta} + (\eta-1)\log e \tag{8}$$

Theorem 3 *[9] Let* $\mathbf{H}$ *be an* $N \times K$ *complex matrix defined as in Theorem 2. Let* $\mathbf{T}$ *be a* $K \times K$ *matrix defined as in Theorem 2 whose the spectral norm is bounded. Let* $g(\cdot)$ *be a continuous function on the real line with bounded and continuous derivatives, analytic on a open set containing the interval*[2]

$$[\liminf_K \phi_K \max{}^2\{0, 1-\sqrt{\beta}\}, \limsup_K \phi_1(1+\sqrt{\beta})^2]$$

where $\phi_1 \geq \ldots \geq \phi_K$ *are the eigenvalues of* $\mathbf{T}$*. Denoting by* λ_i *and* $\mathsf{F}_{\mathbf{H}\mathbf{T}\mathbf{H}^\dagger}(\cdot)$*, respectively, the* i*th eigenvalue and the asymptotic ESD of* $\mathbf{H}\mathbf{T}\mathbf{H}^\dagger$*, the random variable*

$$\mathbf{\Delta}_N = \sum_{i=1}^{N} g(\lambda_i) - N\int g(x)\,d\mathsf{F}_{\mathbf{H}\mathbf{T}\mathbf{H}^\dagger} \tag{9}$$

converges, as $K, N \to \infty$ *with* $\frac{K}{N} \to \beta$*, to a zero-mean Gaussian random variable with variance*

$$\mathbb{E}[\mathbf{\Delta}^2] = -\frac{1}{2\pi^2}\oint\oint \frac{\dot{g}(\mathcal{Z}(\sigma_1))g(\mathcal{Z}(\sigma_2))}{\sigma_2-\sigma_1}d\sigma_1 d\sigma_2 \tag{10}$$

where $\dot{g}(x) = \frac{d}{dx}g(x)$ *while*

$$\mathcal{Z}(\sigma) = -\frac{1}{\sigma}\left(1-\beta(1-\eta_{\mathbf{T}}(\sigma))\right) \tag{11}$$

In (10) the integration variables σ_1 *and* σ_2 *follow closed contours, which we may take to be non-overlapping and counterclockwise, such that the corresponding contours mapped through* $\mathcal{Z}(\sigma)$ *enclose the support of* $\mathsf{F}_{\mathbf{H}\mathbf{T}\mathbf{H}^\dagger}(\cdot)$.

3. ASYMPTOTIC NORMALITY

Using the foregoing results, in this section we characterize the asymptotic mean and variance of the difference

$$N(\mathcal{V}_{\mathsf{F}^N_{\mathbf{B}\mathbf{B}^\dagger}}(\gamma) - \mathcal{V}_{\mathsf{F}_{\mathbf{B}\mathbf{B}^\dagger}}(\gamma))$$

[2] In [11] this interval contains the spectral support of $\mathbf{H}^\dagger\mathbf{H}\mathbf{T}$.

for $N \times K$ random matrices $\mathbf{B}$ that represent multiantenna channels with one-sided correlation at either transmitter or receiver.[3]

Theorem 4 *[9] Let $\mathbf{H}$ be an $N \times K$ complex matrix defined as in Theorem 1. Let $\mathbf{T}$ be an Hermitian random matrix independent of $\mathbf{H}$ with bounded spectral norm and whose asymptotic ESD converges a.s. to a nonrandom limit. Denote by $\mathcal{V}_{\mathbf{HTH}^\dagger}(\gamma)$ the Shannon transform of $\mathbf{HTH}^\dagger$. As $K, N \to \infty$ with $\frac{K}{N} \to \beta$, the random variable*

$$\Delta_N = \log\det(\mathbf{I} + \gamma\mathbf{HTH}^\dagger) - N\mathcal{V}_{\mathbf{HTH}^\dagger}(\gamma) \tag{12}$$

is asymptotically zero-mean Gaussian with variance

$$\mathbb{E}[\Delta^2] = -\log\left(1 - \beta\,\mathbb{E}\left[\left(\frac{\mathsf{T}\gamma\eta_{\mathbf{HTH}^\dagger}(\gamma)}{1+\mathsf{T}\gamma\eta_{\mathbf{HTH}^\dagger}(\gamma)}\right)^2\right]\right)$$

where the expectation is over the nonnegative random variable T whose distribution equals the asymptotic ESD of $\mathbf{T}$.
Proof: *See Appendix.*

Example 1 *Fig. 1 compares the limiting Gaussian distribution of Δ with a histogram of Δ_N with $K = 5$ transmit and $N = 10$ receive antennas, $\gamma = 10$ (the signal-to-noise ratio per receive antenna is $\gamma\frac{K}{N}$) and the (i, j)th entry of the transmit correlation matrix is*

$$(\mathbf{T})_{i,j} = e^{-0.2(i-j)^2} \tag{13}$$

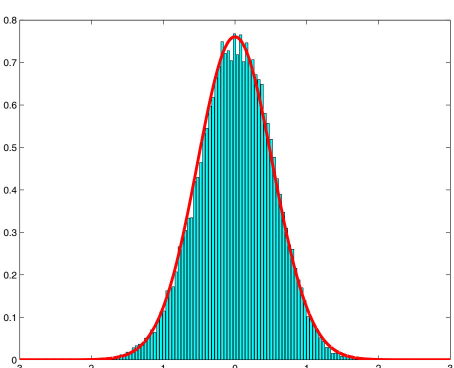

Fig. 1. Histogram of Δ_N for a Rayleigh-faded channel with $K = 5$ and $N = 10$. Transmit antennas are correlated as per (13). Receive antennas are uncorrelated. The solid line indicates the corresponding limiting Gaussian distribution.

For $K = N = 2$ and $K = N = 4$ (i.e., with $\beta = 1$) with the transmit correlation matrix in (13) and no receive correlations, we compute 10% outage capacities via Monte-Carlo and contrast them with the asymptotic formulas. The agreement is remarkable for even such small number of antennas. For $K = N = 2$ specifically, we find that with γ at 10 dB (resp. 0 dB) the asymptotic formula yields 2.2846 (resp. 0.5230) while the Monte Carlo simulation gives 2.2683 (resp. 0.4986). For $K = N = 4$, in turn, we find that at 10 dB (resp. 0 dB) the asymptotic formula yields 5.0129 (resp. 1.4472) while the Monte Carlo simulation gives 5.0109 (resp. 1.4468). The agreement is even closer with different numbers of transmit and receive antennas: for γ at 10 dB (resp. 0 dB), $K = 2N = 4$ and the same correlations as before, we find that the asymptotic formula dictates 3.7795 (resp. 1.1103) while the Monte Carlo simulation gives 3.7759 (resp. 1.1062).

From Jensen's inequality and (7), a tight lower bound for the variance of Δ in Theorem 4 is [1, Eq. 2.239]

$$\mathbb{E}[\Delta^2] \geq -\log\left(1 - \frac{(1-\eta_{\mathbf{HTH}^\dagger}(\gamma))^2}{\beta}\right) \tag{14}$$

with strict equality if $\mathbf{T} = \mathbf{I}$. In fact, Theorem 4 can be particularized to the case $\mathbf{T} = \mathbf{I}$ to obtain:

Corollary 1 *Let $\mathbf{H}$ be an $N \times K$ complex matrix whose entries are i.i.d. zero-mean random variables with variance $\frac{1}{N}$ such that $\mathbb{E}[|\mathsf{H}_{i,j}|^4] = \frac{2}{N^2}$. As $K, N \to \infty$ with $\frac{K}{N} \to \beta$, the random variable*

$$\Delta_N = \log\det(\mathbf{I} + \gamma\mathbf{HH}^\dagger) - N\mathcal{V}_{\mathbf{HH}^\dagger}(\gamma) \tag{15}$$

is asymptotically zero-mean Gaussian with variance

$$\mathbb{E}\left[\Delta^2\right] = -\log\left(1 - \frac{(1-\eta_{\mathbf{HH}^\dagger}(\gamma))^2}{\beta}\right) \tag{16}$$

where $\eta_{\mathbf{HH}^\dagger}(\gamma)$ and $\mathcal{V}_{\mathbf{HH}^\dagger}(\gamma)$ are given in (5) and (6).

Appendix

Using (10) with $g(x) = \log(1 + \gamma x)$,

$$\mathbb{E}[\Delta^2] = \frac{-1}{2\pi^2}\oint\oint\frac{\log(1+\gamma\mathcal{Z}(\sigma_2))\gamma\dot{\mathcal{Z}}(\sigma_1)}{1+\gamma\mathcal{Z}(\sigma_1)(\sigma_2-\sigma_1)}d\sigma_1 d\sigma_2$$

with $\mathcal{Z}(\sigma)$ in (11). With no loss of generality we assume that, besides satisfying the condition of Theorem 3, the σ_1, σ_2 contours do not overlap and the σ_2 contour encloses the σ_1 contour. Defining $f(\sigma) = 1 + \gamma\mathcal{Z}(\sigma)$, from (11)

$$f(\sigma) = 1 + \frac{\gamma}{\sigma}\left(1 - \beta(1 - \eta_{\mathbf{T}}(\sigma))\right) \tag{17}$$

Consequently, $\mathbb{E}[\Delta^2]$ can be re-written as

$$\mathbb{E}[\Delta^2] = -\frac{1}{2\pi^2}\oint\log(f(\sigma_2))\oint\frac{\dot{f}(\sigma_1)}{f(\sigma_1)}\frac{d\sigma_1 d\sigma_2}{\sigma_2-\sigma_1} \tag{18}$$

Let us reproduce the following theorem, which follows from Cauchy's residue Theorem.

[3] Note that $\mathcal{V}_{\mathsf{F}^N_{\mathbf{BB}^\dagger}}(\gamma)$ and $\mathcal{V}_{\mathsf{F}_{\mathbf{BB}^\dagger}}(\gamma)$ give, respectively, the normalized mutual information and the corresponding asymptotic mean for various coherent communication channels that can be described through $\mathbf{B}$.

Theorem 5 *[12, p. 116] Let $\phi(\cdot)$ be a function analytic on and within a closed contour $\mathcal{C}$, except for a finite number of poles $b_1, \ldots, b_n$. Further let $\phi(\cdot)$ be nonzero on $\mathcal{C}$. If $g(\cdot)$ is another function, analytic on and within $\mathcal{C}$, then*

$$\frac{1}{2\pi i}\oint_{\mathcal{C}} \frac{\phi'(z)}{\phi(z)} g(z)dz = \sum_{\mu=1}^{m} \ell_\mu g(a_\mu) - \sum_{\nu=1}^{n} j_\nu \, g(b_\nu)$$

where $a_1, \ldots, a_m$ are the zeros of $\phi(\cdot)$ enclosed in $\mathcal{C}$.

Finding the zeros of $f(\sigma_1)$ is equivalent to solving

$$1 - \frac{\gamma}{\sigma}\left(1 - \beta(1 - \eta_{\mathbf{T}}(\sigma))\right) = 0 \tag{19}$$

whose solution, using (5), is $\hat{\sigma} = \gamma \eta_{\mathbf{HTH}^\dagger}(\gamma)$. Furthermore, from $\dot{f}(\sigma_1)$, using Jensen's inequality and (7), it is seen that $\hat{\sigma}$ is a zero of order one. Without lost of generality, the σ_1 contour can be chosen such that only the simple pole at $\sigma_1 = 0$ of $f(\sigma_1)$ is enclosed. Then, in the neighborhood of this point

$$f(\sigma_1) = (\sigma_1 - \hat{\sigma})\phi(\sigma_1)\frac{1}{\sigma_1} \tag{20}$$

with $\phi(\sigma_1)$ analytic and nonzero on and within the σ_1 contour. (Identical considerations hold for the σ_2 contour). Applying Theorem 5 to (18),

$$\mathbb{E}[\Delta^2] = \oint \frac{\log(f(\sigma_2))}{2\pi i}\left(\frac{1}{\sigma_2 - \hat{\sigma}} - \frac{1}{\sigma_2}\right) d\sigma_2$$

Using (20) and the fact that

$$\oint \log \frac{\sigma_2 - \hat{\sigma}}{\sigma_1}\left(\frac{1}{\sigma_2 - \hat{\sigma}} - \frac{1}{\sigma_2}\right)\frac{d\sigma_2}{2\pi i} = 0$$

because the integrand has a primitive function which is single valued along the σ_2 contour, we obtain

$$\begin{aligned}\mathbb{E}[\Delta^2] &= \oint \frac{\log(f(\sigma_2))}{2\pi i}\left(\frac{1}{\sigma_2 - \hat{\sigma}} - \frac{1}{\sigma_2}\right) d\sigma_2 \\ &= -\log\left(\frac{\phi(\hat{\sigma})}{\phi(0)}\right) \end{aligned} \tag{21}$$

where (21) is obtained via Theorem 5. From (20) and (17),

$$\phi(0) = \frac{mf(\sigma)}{\hat{\sigma}}|_{\sigma=0} = \frac{\gamma}{\hat{\sigma}}, \qquad \hat{\sigma}\phi(\hat{\sigma}) = \frac{d}{d\sigma}\sigma^2 f(\sigma)|_{\hat{\sigma}}$$

and the result is proved.

4. REFERENCES

[1] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1(1), pp. 1–182, 2004.

[2] P. J. Smith and M. Shafi, "On a Gaussian approximation to the capacity of wireless MIMO systems," *Proc. IEEE Intern. Conf. in Commun. (ICC'02)*, pp. 406–410, Apr. 2002.

[3] Z. Wang and G. Giannakis, "Outage mutual information of space-time MIMO channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 4, pp. 657–663, 2004.

[4] M. Kang, L. Yang, and M. S. Alouini, "Capacity of MIMO Rician channels with multiple correlated Rayleigh co-channel interferers," in *Proc. 2003 IEEE GLOBECOM*, Dec. 2003.

[5] P. J. Smith, S. Roy, and M. Shafi, "Capacity of MIMO systems with semi-correlated flat-fading," *IEEE Trans. on Inform. Theory*, vol. 49, no. 10, pp. 2781–2788, Oct. 2003.

[6] A. Moustakas, S. Simon, and A. Sengupta, "MIMO capacity through correlated channels in the presence of correlated interferers and noise: a (not so) large N analysis," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2545–2561, 2003.

[7] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multi-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inform. Theory*, pp. 1893–1909, Oct 2004.

[8] M. Kamath, B. Hughes, and Y. Xinying, "Gaussian approximations for the capacity of MIMO Rayleigh fading channels," in *Asilomar Conf. on Signals, Systems and Computers*, Nov. 2002.

[9] Z. D. Bai and J. W. Silverstein, "CLT of linear spectral statistics of large dimensional sample covariance matrices," *The Annals of Probability*, vol. 32, no. 1A, pp. 553–605, 2004.

[10] J. W. Silverstein, "Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices," *J. of Multivariate Analysis*, vol. 55, pp. 331–339, 1995.

[11] Z. D. Bai and J. W. Silverstein, "Exact separation of eigenvalues of large dimensional sample covariance matrices," *The Annals of Probability*, vol. 27, no. 3, pp. 1536–1555, 1999.

[12] E. C. Titchmarsh, *The theory of functions*, Oxford University Press, 1991.