Distribution of MIMO Mutual Information: A Large Deviations Approach

Pavlos Kazakopoulos, Panayotis Mertikopoulos, Aris L. Moustakas and Giuseppe Caire

***Abstract*—Using a large deviations approach we calculate the probability distribution of the mutual information of MIMO channels in the limit of large antenna numbers. In contrast to previous methods that only focused to the distribution close to its most probable value, thus obtaining an asymptotically Gaussian distribution, we calculate the full distribution including its tails, which behave quite differently from the bulk of the distribution. Our resulting probability distribution seamlessly interpolates between the Gaussian approximation for rates R close to the ergodic value of the mutual information and the approach of Zheng and Tse [1], valid for large signal to noise ratios ρ. This provides us with a tool to analytically calculate outage probabilities at any point in the (R, ρ, N) parameter space, as long as the number of antennas N is not too small. In addition, this method also yields the probability distribution of eigenvalues constrained in the subspace where the mutual information per antenna is fixed to R for a given ρ. Quite remarkably, this eigenvalue density is of the form of the Marcenko-Pastur distribution with square-root singularities.**

I. Introduction

Considerable interest has followed from the prediction [2], [3] that the use of multiple antennas in transmitting and receiving signals can result to substantial gains in information throughput. To analyze the theoretical limits of such a MIMO (Multiple Input Multiple Output) system, it has been convenient to focus on the i.i.d. Gaussian noise and input case of the mutual information for a channel matrix $\mathbf{H}$, which takes the familiar form

$$I_N = \ln\det\left(\mathbf{I} + \rho\mathbf{H}^\dagger\mathbf{H}\right). \tag{1}$$

Here ln signifies the natural logarithm, ρ is the signal to noise ratio and $\mathbf{H}$ is the $M \times N$ channel matrix, assumed to have independent $\mathcal{CN}(0, 1/N)$ elements.

One first attempt to analytically quantify the gains of the use of MIMO was to assume that the number of antennas N is large and $\beta = M/N$ fixed and finite, in which case $\mathbf{H}$ can be viewed as a large random matrix [4]. By applying ideas and methods from the theory of random matrices, it was shown that the value of the mutual information per antenna $I(\rho, \mathbf{H})/N$ "freezes" to a deterministic value in the large N limit, corresponding to its most likely value, the so-called *ergodic average* $r_{erg}(\rho)$. Underlying this result is the fact that the distribution of eigenvalues of $\mathbf{H}^\dagger\mathbf{H}$ itself freezes to the celebrated Marcenko-Pastur (MP) distribution, given by

$$p(x) = \frac{\sqrt{(b-x)(x-a)}}{2\pi x} \tag{2}$$

where $a, b = (1 \pm \sqrt{\beta})^2$ are the end-points of its support.

For finite N, one is naturally interested in the distribution of the mutual information since this would give the probability of outage when transmitting at an arbitrary rate in a fading channel. Various approaches [5]–[7] have shown that the distribution of the mutual information R becomes asymptotically Gaussian, with mean equal to the ergodic capacity $R_{erg} = Nr_{erg}(\rho)$ and a variance of order unity in N. This Gaussian variability of the mutual information is due to the fluctuations of the eigenvalues of the matrix away from the most probable distribution described by the MP law. Since this Gaussian approximation is essentially a variation of the central limit theorem, it is only valid within a finite number of standard deviations from R_{erg}. As a result, it should fail to capture the tails of the distribution, for example the probability of the rate R being half the most probable one $R_{erg}/2$, since this would correspond to $O(N)$ standard deviations away from the mean.

Nevertheless, the tails of the distributions of the mutual information are important, since they correspond to regimes with low outage probability, where one would want to operate a MIMO system. This interplay between low outage and multiplexing gain was exemplified in [1]. In this seminal paper, the authors analyzed the asymptotics of the distribution of the mutual information in the limit of large ρ, with $R/\ln\rho$ fixed. They found that the asymptotic form of the logarithm of the distribution of the mutual information is a piecewise linear function of $R/\ln\rho$, linearly interpolating between the values

$$\ln P(R_n) \approx -\frac{(R_n - M\ln\rho)(R_n - N\ln\rho)}{\ln\rho} \tag{3}$$

where the intermediate rates are given by $R_n = n\ln\rho$ for integer $n \leq N$.[1] When, in addition to ρ, N is large, $\ln P(R)$ in (3) becomes to leading order a continuous function of R/N. It should be pointed out that this result is complementary to the large N asymptotics discussed above, since it provides insight in the distribution of the mutual information quite far from its peak, which for large ρ (and large N) is at $I_N \approx N\ln\rho$. However, it does not give any quantitative estimates of the

P. Kazakopoulos (pkazakop@phys.uoa.gr), P. Mertikopoulos (pmertik@phys.uoa.gr) and A. L. Moustakas (arislm@phys.uoa.gr) are with the Physics Dept., Athens Univ., 157 84 Athens, Greece. G. Caire (caire@usc.edu) is with the EE - Systems Dept., Univ. Southern California Los Angeles, CA 90007, USA

[1] It should be noted that [1] analyzed the *outage* probability, namely the $Prob(I_N < R)$, rather than the probability density $P(R) = dProb(I_N < R)/dR$. However for large ρ and $R_n < N\ln\rho$ the expression of both is identical up to corrections of O(1).

outage for finite ρ and therefore cannot be used to realistic outage predictions.

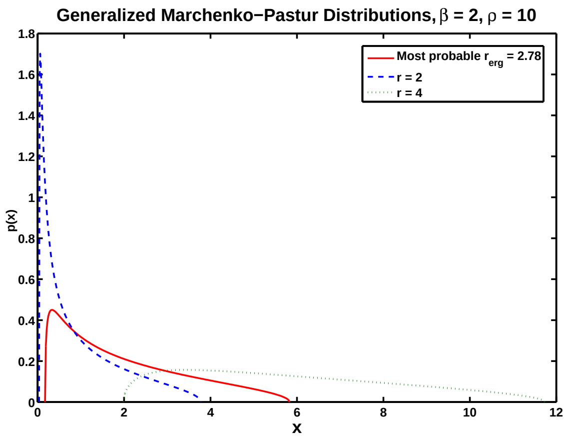

Fig. 1. Generalized MP distributions for SNR=10dB and different values of r

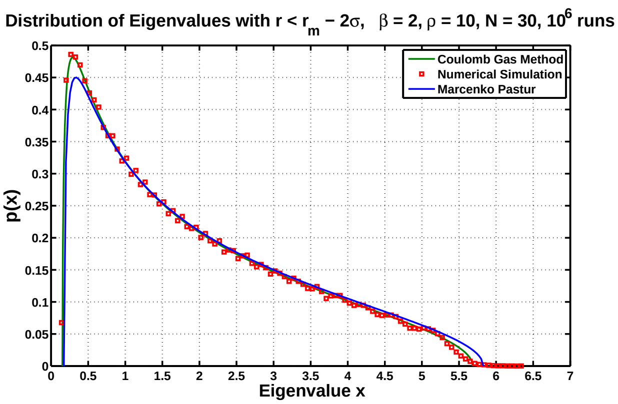

Fig. 2. Comparison of MP distribution and distribution resulting from shifting the rate by two standard deviations from the most probable one r_{erg}. The agreement is quite remarkable with the numerical distribution.

Meantime, all variants of the large N Gaussian approximation of the mutual information [5]–[7] fail spectacularly for large ρ, even close to the peak of the distribution. Specifically, they all predict that the probability distribution is given asymptotically by

$$\ln P(R) \approx -\frac{(R - N\ln\rho)^2}{2\ln\left(1-\beta^{-1}\right)} \tag{4}$$

where $\beta = M/N > 1$, which is in striking disagreement compared to (3). It should be noted that for $\beta = 1$ the asymptotic form of (3) is recovered within the Gaussian approximation [5], [7]. This discrepancy for $\beta \neq 1$ indicates that the limits $N \to \infty$ and $\rho \to \infty$ cannot naïvely be interchanged. When one takes the large N limit first, one focuses on the most probable eigenvalue distribution, which tends to the Marcenko-Pastur distribution (2), and small variations around it. However, as can be seen in (2), this distribution (almost surely) produces no eigenvalues close to zero when $\beta > 1$. Nevertheless, the analysis for large ρ focuses at the regime where the eigenvalues are of order $z \equiv \rho^{-1}$. As a result, it is not surprising that the large-N Gaussian approximation of the mutual information distribution misses the correct behavior.

In summary, we have two methods, the large-N Gaussian approximation and the large ρ analysis, both of which have their own regions of validity but fail to produce quantitative results for the outage probability outside that region. Thus, one still needs an approach that correctly describes the outage behavior of the mutual information distribution for arbitrary ρ and R.

In this paper, we introduce a large deviations approach: this is formally valid only for large N, but it actually works over the whole range of values of R and ρ. This method bridges the two regions of small/intermediate and large signal to noise ratios within a single formalism. In effect it amounts to calculating the rate function of the logarithm of the average moment generating function of the mutual information. Our method was first introduced in the context of random matrix theory by Dyson [8] and is quite intuitive because it looks at the eigenvalues as point charges repelling each other logarithmically. As a byproduct of this approach, we obtain the most probable eigenvalue distribution constrained on the manifold of fixed total rate R and signal to noise ratio ρ. This is a generalized Marcenko-Pastur equation that gives the eigenvalue distribution at the tails of the distribution of R.

This generalized Marcenko-Pastur distribution can also be seen as the inverse of the so-called Shannon transform introduced in [9] in the following sense: While the Shannon transform yields the value of normalized mutual information I_N/N as a functional of the asymptotic eigenvalue distribution of $\mathbf{H}^\dagger\mathbf{H}$ (the Marcenko-Pastur distribution), in contrast the generalized Marchenko-Pastur distribution introduced here corresponds to the asymptotic eigenvalue distribution of $\mathbf{H}^\dagger\mathbf{H}$ for a given value of the mutual information $R = Nr$, i.e., when $\mathbf{H}^\dagger\mathbf{H}$ is constrained on the manifold defined by $r = I_N/N$.

II. Methodology

In this section we will describe the basic steps of the mathematical methodology, following the elegant approach introduced in [10], [11]. We wish to calculate the probability distribution of the mutual information (1), which can be written in terms of λ_k, the eigenvalues of the Wishart matrix $\mathbf{H}^\dagger\mathbf{H}$ as

$$I_N = \sum_{k=1}^{N} \ln\left(1+\rho\lambda_k\right) \tag{5}$$

We will assume that $\beta = M/N > 1$. In the opposite case $\beta < 1$, we just need to redefine $\rho_{new} = \rho\beta$ and interchange M with N. Clearly, the probability distribution of the mutual information will depend on the joint probability distribution of the eigenvalues $\{\lambda_k\}$. This is well known to be

$$\begin{aligned} P(\{\lambda_k\}) &= A\Delta(\{\lambda_i\})^2 \prod_{k=1}^{N} \lambda_k^{M-N} e^{-N\lambda_k} \\ &= A e^{-N\sum_k(\lambda_k - (\beta-1)\ln\lambda_k) + 2\sum_{j>k}\ln|\lambda_j-\lambda_k|} \end{aligned} \tag{6}$$

where A is a constant and $\Delta(\{\lambda_k\}) = \prod_{i>j}(\lambda_i - \lambda_j)$ is the so-called Vandermonde determinant of the eigenvalues λ_k. An intuitive way to interpret the eigenvalues λ_k in the exponent of (6) as the positions of N positive unit charges located on a line repelling each other and interacting with an external electric field. Thus the exponent in (6) can be viewed as the configuration energy of these charges. This picture was first proposed by Dyson [8]. Within this interpretation, the last term in the exponent corresponds to the logarithmic repulsion energy, while the first term is the potential due to a constant field, and the second term is the repulsion of a point charge located at the origin. Indeed, these are simply the potentials that one obtains in classical two-dimensional electrostatics. Dyson also pointed out that, just as in electrostatics, when the number of charges is large it is natural to analyze the individual charges collectively as a charge density $p(x)$, which in this case has to be normalized properly. It is therefore convenient to express the sums in the exponent above as integrals over a density, e.g. $\sum_k \lambda_k = N \int dx p(x) x$, where $Np(x) = \sum_k \delta(x - \lambda_k)$. In this framework, an average of any function of λ_k over their joint probability distribution is interpreted as an integration over all possible normalized charge densities $p(x)$.

Our task is to calculate the probability density of the mutual information $P(R) = E_p\left[\delta(I_N - Nr)\right]$, where r is the value of the (normalized) mutual information per antenna, $r = R/N$. Representing the δ-function through its Laplace transform and expressing I_N as $N \int dx p(x) \ln(1 + \rho x)$ yields

$$\delta(I_N - Nr) = \int Dk\, e^{N^2 k[\int dx p(x) \ln(1+\rho x) - r]} \quad (7)$$

where k is integrated along the path $(t - i\infty, t + i\infty)$ for some appropriate real number t with the integration measure $Dk = N dk/(2\pi i)$. Similarly, one can express the normalization condition as a δ-function $\delta(N \int dx p(x) - N)$ inside the p-average, which can then also be expressed by its Laplace transform. Following an approach introduced by Dyson [8] and more recently discussed in [11] the probability density of the mutual information $P(R)$ can be written as

$$\begin{aligned}
P(R) &= \frac{Z_1}{Z_0} && (8)\\
Z_0 &= \int \mathcal{D}p \int Dc\, e^{-N^2 E_0[p]} && (9)\\
E_0[p] &= \int_0^\infty dx\, x p(x) - (\beta - 1) \int_0^\infty dx\, p(x) \ln x \\
&- \int_0^\infty dx \int_0^\infty dx' p(x) p(x') \ln|x - x'| \\
&+ c\left[\int_0^\infty dx\, p(x) - 1\right] && (10)\\
Z_1 &= \int \mathcal{D}p \int Dc \int Dk e^{-N^2 E_1[p]} && (11)\\
E_1[p] &= E_0[p] + k\left[\int_0^\infty dx\, p(x) \ln(1 + \rho x) - r\right] && (12)
\end{aligned}$$

The notation $\int \mathcal{D}p$ signifies a functional integral over all possible charge density functions $p(x) \geq 0$.[2] Similar to (7), the integral over c ensures the normalization of $p(x)$. Z_1 and Z_0 represent volumes in configuration space of $p(x)$, the latter constrained to normalized $p(x)$, while the former also constrained in the subspace where the mutual information is fixed to $I_N = R(= Nr)$ (7). In the same spirit, $E_1[p]$ and $E_0[p]$ represent the energies of the configurations $p(x)$ with and without the mutual information constraint.[3]

Following Dyson [8], it is instructive to interpret the terms in $E_0[p]$, $E_1[p]$ above in the spirit of a Coulomb gas. Thus the first term in (10) corresponds to the energy due to a constant electric field attracting the charge density towards the origin. The second term is a logarithmic repulsion due to a δ - function charge density at the origin with strength $\beta - 1$, while the third one is the logarithmic repulsion between charges. The fourth term adds a constant c shift in the potential of all charges, which, just as in electrostatics, determines the total charge, i.e. the normalization condition. Finally, the second term in (12) is a *virtual* δ - function charge density at the (forbidden) location $x = -z$. This acts as a shift of the total charge density to the left or the right, depending on the sign of k, to ensure that the mutual information per channel is r. It should be noted that this intuition holds only for real c and k, however, as we shall see these will be the only relevant values for large N.

For large N both Z_0 and Z_1 can be approximated by the value of their integrands at the saddle points of the corresponding exponents. The saddle points are determined by taking the functional derivatives of $E_0[p]$ and $E_1[p]$ respectively, with respect to the function $p(x)$ and setting these to zero. As a result we need to find the density $p_1(x)$ that satisfies the following relation

$$\left.\frac{\delta E_1[q]}{\delta q}\right|_{q=p_1} = 0 \quad (13)$$

with a similar relation for $E_0[p]$, which is minimized with a different density $q = p_0$. In addition, $E_1[p]$ needs to be at a saddle-point with respect to the variables c and q. Is easy to see that, in order for the final result to be positive, the saddle-point values of c, k have to be real. As a result we also need

$$\begin{aligned}
\int dx p(x) &= 1 && (14)\\
\int dx p(x) \ln(1 + \rho x) &= r && (15)
\end{aligned}$$

Furthermore, $p(x)$ has to be non-negative in the region $x > 0$, since it represents a density of eigenvalues. We will henceforth denote $E_1 = E_1[p_1]$ and $E_0 = E_0[p_0]$. From the extremal values of E_1 and E_0 the asymptotic probability distribution of the mutual information $P(R)$ (8) can be evaluated, in the sense that

$$\lim_{N\to\infty} \frac{\ln P(R)}{N^2} = -\left(E_1 - E_0\right) \quad (16)$$

[2] We refer the reader to any standard advanced physics book, most notably [12] for details on functional derivatives and functional integrals, or *path integrals*.

[3] It should be pointed out that the above equations are correct to order $o(N)$ in the exponent; see e.g. [8], [11].

i.e. for finite N, $P(R) \approx e^{-N^2(E_1-E_0)}$. The above analysis yields identical results with a more standard large deviations approach. For large N, Varadhan's lemma [13] would imply that $\ln Z_0/N^2$ would be asymptotically equal to $-\max_p \min_c E_0[p]$, while $\ln Z_1/N^2 \to -\max_p \min_{c,k} E_1[p]$.

Our task now is to find a solution of (13), subject to the above conditions. The solution for E_0 will result as a special case when we relax the condition (15), which will correspond to setting $k = 0$ in the final result, with $p_0(x)$ the MP distribution. (13) results in

$$2\int_0^\infty dx' p(x') \ln|x-x'| = x-(\beta-1)\ln x + c + k\ln(1+\rho x) \tag{17}$$

With a fair amount of hindsight we know that the solution of the above integral equation will be positive between two as of yet undetermined endpoints $0 \le a < b < \infty$. To eliminate c for the moment, we differentiate with respect to x of (17):

$$2\mathcal{P}\int_a^b \frac{p(x')}{x-x'}dx' = 1 - \frac{\beta-1}{x} + \frac{k}{x+z} \equiv f(x) \tag{18}$$

where $\mathcal{P}$ denotes the principal value of the expression. The above equation has a straightforward physical meaning. It represents a balance of forces at every location $b > x \ge a$: The repulsion from all other charges of the distribution located at x' (expression at the LHS) is equal to the external forces (RHS). When $\beta > 1$, $p(x)$ cannot have any support at $x = 0$, because in this case the force from the finite charge density located at $x = 0$ (second term above) would be infinite. As a result we expect that for $\beta > 1$, $a > 0$. The solution of this integral equation for general $f(x)$ can be obtained using standard methods [14].

$$\begin{aligned} p(x) &= \frac{\mathcal{P}\int_a^b \frac{\sqrt{(y-a)(b-y)}f(y)}{y-x}dy + C}{2\pi^2\sqrt{(x-a)(b-x)}} \\ &= \frac{\sqrt{b-x}}{2\pi\sqrt{x-a}}\left[1 - \frac{k\sqrt{a+z}}{(x+z)\sqrt{b+z}} - \frac{(\beta-1)\sqrt{a}}{x\sqrt{b}}\right] \end{aligned} \tag{19}$$

In the second line above we have eliminated the constant C by the condition $p(b) = 0$.[4] The unknowns a, b can be determined as a function of k by the normalization condition (14) and the condition $p(a) = 0$.

III. Results

When $\beta > 1$, i.e. $M > N$, the finite density of zero eigenvalues repels other eigenvalues from the region therefore the continuous charge density has a lower limit $a > 0$, at which it goes to zero. If we impose the condition $p_1(a) = 0$ on (19) we get

$$\frac{k}{\sqrt{(a+z)(b+z)}} + \frac{\beta-1}{\sqrt{ab}} = 1 \tag{20}$$

Thus $p_1(x)$ takes the form

$$p_1(x) = \frac{1}{2\pi}\frac{\sqrt{(b-x)(x-a)}}{x(x+z)}\left(x + \frac{(\beta-1)z}{\sqrt{ab}}\right) \tag{21}$$

[4]It can be shown that when a, b are free to vary, this results to the extra conditions $p(a) = p(b) = 0$ [15].

Integrating $p_1(x)$ and imposing (14) implies

$$(a+b-2-2k-2\beta+2z)\sqrt{ab} = 2(\beta-1)z \tag{22}$$

We will show elsewhere that (20), (22) admit a unique solution a, b for given k. It is worth commenting on (21). This equation gives the asymptotic density of eigenvalues constrained on the subspace with fixed total rate $R = Nr$ in the limit of large N. Its form is reminiscent of the Marcenko-Pastur equation, to which it actually reduces when $k = 0$, or equivalently $r = r_{erg}$, i.e. when the r - constraint (15) is lifted. It is quite remarkable that even away from the (most probable) MP distribution, the constrained eigenvalue distribution "hardens" to a deterministic one. In Fig. 1 we plot examples of such eigenvalue distributions. Finally, we can readily integrate (15) with $p_1(x)$ given by (21) to obtain

$$\begin{aligned} r &= \ln\rho + \frac{\Delta}{4}\left[w_-^2\ln\left(\frac{\Delta}{e}\right) + 2w_+^2\ln\frac{w_+}{2}\right] \\ &- \frac{\Delta}{8}\left[(w_+^2-w_-^2)\ln\frac{(w_+^2-w_-^2)}{4}\right] \\ &+ \frac{(\beta-1)\Delta\ln\Delta}{8\sqrt{ab}}\left(\frac{w_+^2-w_-^2-v_+^2+v_-^2}{4}-1\right) \\ &+ \frac{(\beta-1)\Delta(v_+^2-v_-^2)}{8\sqrt{ab}}\ln v_+^2 \\ &+ \frac{(\beta-1)\Delta(w_+^2-w_-^2)}{8\sqrt{ab}}\ln\left(\frac{w_+^2-w_-^2}{4}\right) \\ &+ \frac{(\beta-1)\Delta}{2\sqrt{ab}}\left[(2v-w_+^2+1)\ln w_+ + z\ln 4\right] \\ &- \frac{(\beta-1)(v_+^2-v_-^2)}{4\sqrt{ab}}\ln\left(\sqrt{v(1+w)}+\sqrt{w(1+v)}\right) \end{aligned} \tag{23}$$

where $\Delta = b-a$, $w = \frac{a+z}{\Delta}$, $v = \frac{a}{\Delta}$, $w_\pm = \sqrt{1+w} \pm \sqrt{w}$, $v_\pm = \sqrt{1+v} \pm \sqrt{v}$. This equation can be solved to give k as a function of r. For $k = 0$ (20) and (22) result to the MP values of $a, b = (1 \pm \sqrt{\beta})^2$ and (21) reduces to (2). In Fig. 2 we compare the distribution with the corresponding numerical results. We see that the agreement is quite remarkable.

We may now calculate the values of E_0, E_1. The simplest way is to start by evaluating the saddle point value of c by evaluating the expression in (17) for $x = a$ and then substituting the RHS of that equation with its LHS in the double integral of (10) and (12) [10]. The final result for E_1

$$\begin{aligned} E_1 &= \frac{\Delta^2}{8}\left[\frac{1}{2} + \frac{4kz}{w_+^2-w_-^2}w_-^2\right] \\ &- \frac{(\beta-1)\Delta}{4}\left[\frac{\ln\Delta}{2}w_-^2 - \frac{1}{2}v_-^2 + \ln\frac{v_+}{4}\right. \\ &+ w\ln\frac{v_+^2}{4} + \frac{w_+^2-w_-^2}{2} \\ &- \left.2\ln w_+\ln\left(\sqrt{v(1+w)}+\sqrt{w(1+v)}\right)\right] \\ &- \frac{\Delta}{4}\left[\ln\Delta w_-^2 + \frac{w_+^2-w_-^2}{2}\ln\frac{w_+^2}{w} - (1+2w)\ln 4\right] \\ &+ \frac{1}{2}\left[kr + a - (\beta-1)\ln a - k\ln\left(1+\frac{a}{z}\right)\right] + \frac{\Delta}{4} \end{aligned} \tag{24}$$

To obtain $p_0(x)$ and E_0 and one can simply set $k = 0$ in (20), (22). Due to the factor N^2 in the exponent of $P(R) \sim \exp\left[-N^2(E_1 - E_0)\right]$, $P(R)$ falls off rapidly away from its peak. Therefore, it can be shown that to leading order in N the logarithm of the outage probability $\ln Prob(I_N < R)$ is equal to $\ln P(R)$ when $R < R_{erg}$, while for $R > R_{erg}$, $\ln Prob(I_N < R) = 0$.

In Fig. 3 we plot the logarithm of the outage probability as a function of throughput r and we compare the result with the two other asymptotic forms, namely the Gaussian approximation of the mutual information [5] and the large-ρ asymptotic result given by (3) [1]. We see that our result performs much better at low outage, even at large $\rho = 100$. Fig. 4 shows more clearly the different trends of this curve compared to the other two approximations.

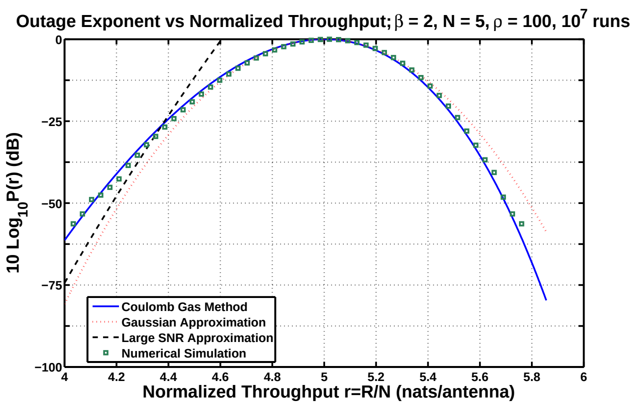

Fig. 3. Plot of the logarithm of the outage probability as a function of throughput r for $\beta = 2$ and comparison to the Gaussian approximation and the large-ρ asymptotic result obtained by Zheng and Tse given by (3) [1]. The numerical result for $N = 5$ follows closely our result, even at large $\rho = 100$.

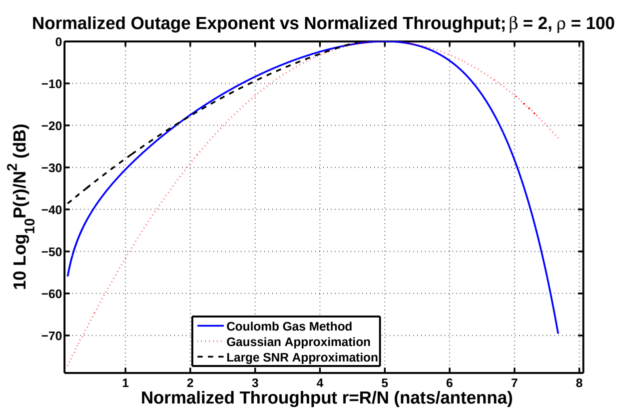

Fig. 4. Further comparison of outage probabilities of our result with that of the Gaussian approximation and the Zheng - Tse result

IV. Conclusion

In this paper we have used a large deviation approach, first introduced in the context of statistical mechanics [8], [10], to calculate the probability distribution of the mutual information of MIMO channels in the limit of large antenna numbers. In contrast to previous approaches that focused only close to the most probable eigenvalue distribution, [5]–[7], we also calculate the distribution for *rare events* in the tails of the distribution, corresponding to instances where the observed mutual information differs by $O(N)$ from the most probable value of the asymptotic distribution and hence the Gaussian approximation for the mutual information is invalid. We find that the distribution in those tails behaves markedly different from the center. Our resulting probability distribution seamlessly interpolates between the Gaussian approximation for rates close to the ergodic mutual information to the results of [1] for large signal to noise ratios, where the outage probability is given asymptotically by (3). Our method thus provides an analytic tool to calculate outage probabilities in any point in the (R, ρ, N) parameter space, as long as N is not too small. Additionally, this approach also provides the probability distribution of eigenvalues constrained in the subspace where the mutual information is fixed to R for a given signal to noise ratio ρ. Interestingly, this eigenvalue density is of the form of the Marcenko-Pastur distribution with square-root singularities. In a forthcoming work we will analyze the $\beta = 1$ case where a phase transition occurs, in which $a > 0$ and $p(a) = 0$ beyond a critical value of $r_c(\rho)$.

References

[1] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[2] 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, pp. 311–335, 1998.

[3] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications and Related Technologies*, vol. 10, no. 6, pp. 585–596, Nov. 1999.

[4] P. B. Rapajic and D. Popescu, "Information capacity of a random signature multiple-input multiple-output chanel," *IEEE Trans. Commun.*, vol. 48, no. 8, p. 1245, Aug. 2000.

[5] A. L. Moustakas, S. H. Simon, and A. M. 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, Oct. 2003.

[6] 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*, vol. 50, no. 9, pp. 1893–1909, Sept. 2004.

[7] W. Hachem *et al.*, "A new approach for capacity analysis of large dimensional multi-antenna channels," *IEEE Trans. Inform. Theory*, vol. 54, pp. 3987–4004, Sep. 2008.

[8] F. Dyson, "Statistical theory of the energy levels of complex systems. I," *J. Math. Phys.*, vol. 3, p. 140, 1962.

[9] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, 2004.

[10] P. Vivo, S. N. Majumdar, and O. Bohigas, "Large deviations of the maximum eigenvalue in Wishart random matrices," *J. Phys. A*, vol. 40, pp. 4317–4337, 2007.

[11] D. S. Dean and S. N. Majumdar, "Extreme value statistics of eigenvalues of Gaussian random matrices," *Phys. Rev E*, vol. 77, p. 041108, 2008.

[12] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*. New York: McGraw-Hill, 1965.

[13] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*. New York, USA: Springer-Verlag Inc., 1998.

[14] S. G. Mikhlin, *Integral Equations*. New York: Pergamon, 1964.

[15] C. W. J. Beenakker, "Random-matrix theory of quantum transport," *Rev. Mod. Phys.*, vol. 69, pp. 731–808, 1997.