Publication information

Title	Characterizing the mutual information distribution of MIMO systems: Beyond the Gaussian approximation
Author(s)	Li, Shang; McKay, Matthew R.; Chen, Yang
Source	Conference record - Asilomar Conference on Signals, Systems, & Computers , 2012, p. 1758-1762
Version	Published version
DOI	http://dx.doi.org/10.1109/ACSSC.2012.6489335
Publisher	IEEE

Copyright information

Copyright © 2012 IEEE. Reprinted from (S. Li, M. R. McKay and Y. Chen, "Characterizing the mutual information distribution of MIMO systems: Beyond the Gaussian approximation," 2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computer (ASILOMAR), Pacific Grove, CA, 2012, pp. 1758-1762).

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of The Hong Kong University of Science and Technology's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org. By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

Notice

This version is available at HKUST Institutional Repository via

<http://hdl.handle.net/1783.1/56637>

If it is the author's pre-published version, changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published version.

Characterizing the Mutual Information Distribution of MIMO Systems: Beyond the Gaussian Approximation

Shang Li*, Matthew R. McKay*, Yang Chen†

*Dept. of Elec. and Comp. Engineering, Hong Kong University of Science and Tech., Hong Kong

†Department of Mathematics, University of Macau, Macau

Abstract—This paper investigates the mutual information distribution of MIMO channels. We further develop our recent work [1], where an exact expression was derived for the moment generating function in terms of a Painlevé V differential equation, along with a Gaussian approximation. Based on the exact representation, we systematically compute closed-form expressions for the high-order cumulants of the mutual information distribution to leading order in the number of antennas, as well as first-order correction terms which provide improved accuracy for finite-length arrays. These results yield considerable new insight, for example, providing a technical explanation as to why the Gaussian approximation is quite robust to the system parameters for the case of unequal transmit and receive antenna arrays, whilst for equal antenna arrays it deviates strongly as the SNR increases. In addition, by employing our new expressions for the higher order cumulants, we draw upon the Edgeworth expansion technique to propose a *refined* Gaussian approximation. This approximation is shown to give a very accurate closed-form characterization of the mutual information distribution, both around the mean and also in the tail region of interest for the outage probability.

I. INTRODUCTION

Recently, there has been a flurry of research activity aimed at designing advanced multiple-input multiple-output (MIMO) technologies. However, despite the plethora of existing work in this field, some very fundamental problems still remain unsolved, even for the simplest point-to-point MIMO systems. One of the most important is the characterization of the *outage capacity*, which gives an achievable rate for transmission over quasi-static channels.

Characterizing the outage capacity is much more difficult than for the ergodic capacity, since it requires solving for the entire distribution of the channel mutual information, rather than simply the mean, and thus far this distribution is only partially understood. To address this problem, an exact expression was computed for the moment generating function (MGF) of the mutual information in [2]. This result however, was given in terms of a Hankel determinant which yields very little insight. Alternatively, many other papers have focused on giving a large-antenna asymptotic analysis which provides more intuitive results; see e.g., [1, 3–5] and also [6, 7]. Quoting [1], for a MIMO system subjected to independent and identically distributed (IID) Rayleigh fading with n_t transmit antennas, n_r receive antennas, and signal to noise ratio (SNR) P , and defining $n := \min\{n_t, n_r\}$, $m := \max\{n_t, n_r\}$, $\beta := m/n$, the mutual information distribution is approximated by

a Gaussian with the n -asymptotic mean μ_0 given in (1) (top of the next page) and variance

$$\sigma_0^2 = 2 \log \left[\frac{1}{2} \left(\frac{\beta + aP}{\beta + bP} \right)^{1/4} + \frac{1}{2} \left(\frac{\beta + bP}{\beta + aP} \right)^{1/4} \right] \quad (2)$$

with $a := (\sqrt{\beta+1} - 1)^2$ and $b := (\sqrt{\beta+1} + 1)^2$. Though extensively employed due to its simplicity, the Gaussian approximation becomes highly inaccurate in some cases. We provide two examples to demonstrate this. As a first example, Fig. 1(a) and Fig. 1(b) show the probability density function (PDF) of the Gaussian approximation and the true (simulated) PDF of the mutual information distribution for finite antenna arrays and for increasing SNR. Comparing these two sub-figures, we see that the Gaussian deviates from the true PDF in all cases as the SNR increases, whilst interestingly the deviation is significantly stronger when $n_r = n_t$ compared with the alternative case, despite the fact that the number of antennas are larger. Thus, when $n_t \neq n_r$ it appears that the Gaussian approximation is more robust to increasing SNR. This key observation is unexpected and hitherto lacks any rigorous explanation. As a second example, for finite length antenna arrays, the Gaussian approximation has been shown to be inaccurate in the left-hand tail region of the distribution, which is of prime interest for the outage probability [8, 9]. This is observed in Fig. 2(b), where the solid curve (i.e., Gaussian) fails markedly in capturing the simulations for small but practical outage probabilities.

Only very recently have tools been introduced which aim to capture the Gaussian deviations; most notably [9] and [1]. In [9], a new large deviations approach was adopted based on Coulomb-fluid heuristics from statistical physics. On a more rigorous footing, our recent work [1] provided tools for systematically capturing the finite- n corrections to the asymptotic Gaussian results; although only deviations for the specific case $n_r = n_t$ were investigated. Related work for multi-user MIMO systems was presented in [10].

In this paper, we further develop our existing studies in [1], giving a broader understanding of the mutual information distribution for arbitrary numbers of antennas. First, we will systematically compute new closed-form expressions for the higher order cumulants of the mutual information distribution. Based on these results, we then provide a rigorous technical explanation for the large SNR deviation behavior observed in Figs. 1(a) and 1(b). Finally, we draw upon the Edgeworth expansion technique to give a refined closed-form distribution approximation, which is shown to be very accurate both around the mean and also in the tail region.

$$\mu_0 = n \left[\frac{a+b}{2} \log \left(\frac{\sqrt{\beta+aP} + \sqrt{\beta+bP}}{2\sqrt{\beta}} \right) - \sqrt{ab} \log \left(\frac{\sqrt{a(\beta+bP)} + \sqrt{b(\beta+aP)}}{\sqrt{a\beta} + \sqrt{b\beta}} \right) - \frac{(\sqrt{\beta+aP} - \sqrt{\beta+bP})^2}{4P} \right] \quad (1)$$

(a) $n_t = n_r = 3$ (b) $n_t = 3, n_r = 2$

Fig. 1. PDF of mutual information for different antenna settings and SNRs.

II. SYSTEM MODEL

Under flat-fading, the linear MIMO channel model takes the form:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \quad (3)$$

where $\mathbf{y} \in \mathbb{C}^{n_r}$ and $\mathbf{x} \in \mathbb{C}^{n_t}$ denote the received and transmitted vector respectively, whilst $\mathbf{H} \in \mathbb{C}^{n_r \times n_t}$ represents the channel matrix, and $\mathbf{n}_{n_r \times 1} \in \mathbb{C}^{n_r}$ represents noise. Assuming rich scattering, $\mathbf{H}$ is modeled as Rayleigh fading, having IID entries $h_{i,j} \sim \mathcal{CN}(0,1)$, known to the receiver only. The noise is assumed $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{n_r})$. The input is selected to be the ergodic-capacity-achieving input distribution $\mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \frac{P}{n_t} \mathbf{I}_{n_t}\right)$, where P is the transmitted power constraint and also represents the SNR due to the normalized noise. For this model, the mutual information between the channel input and output is [11]:

$$\mathcal{I}(\mathbf{x}; \mathbf{y}) = \begin{cases} \ln \det \left(\mathbf{I}_n + \frac{P}{n} \mathbf{H} \mathbf{H}^\dagger \right), & n_t \geq n_r \\ \ln \det \left(\mathbf{I}_n + \frac{P}{n} \mathbf{H}^\dagger \mathbf{H} \right), & n_t < n_r \end{cases} \quad (4)$$

According to (4), we can assume $n_t \geq n_r$ without loss of generality; otherwise, if $n_t < n_r$, we only need replace P with βP .

To characterize the distribution of (4), a natural approach is to employ the MGF:

$$\mathcal{M}(\lambda) := E[\exp(\lambda \mathcal{I}(\mathbf{x}; \mathbf{y}))]. \quad (5)$$

With this, the cumulant generating function (CGF) can be

expressed as a power series about $\lambda = 0$:

$$\mathcal{K}(\lambda) := \ln \mathcal{M}(\lambda) = \sum_{\ell=1}^{\infty} \kappa_{\ell} \frac{\lambda^{\ell}}{\ell!} \quad (6)$$

where the coefficient κ_{ℓ} is the ℓ th cumulant of $\mathcal{I}(\mathbf{x}; \mathbf{y})$.

III. CUMULANTS ANALYSIS

A. Exact Painlevé Representation of the MGF

We first quote the following exact representation for the MGF (5) in [1]; a key result derived by the authors by drawing upon methods from random matrix theory (see e.g., [12–14]).

Proposition 1: The MGF (5) admits the following compact representation:

$$\mathcal{M}(\lambda) = \exp \left(\int_{\infty}^{\beta/P} \frac{G_n(x)}{x} dx \right) \quad (7)$$

where $G_n(x)$ satisfies a version of the Painlevé V continuous Jimbo-Miwa-Okamoto σ -form:

$$(xG_n'')^2 = n^2 (xG_n' + G_n' (1 + \beta + \lambda/n) - G_n)^2 - 4 (xG_n' - G_n + nm) (G_n'^2 + \lambda n G_n'), \quad (8)$$

with $'$ denoting the derivative with respect to x .

Whilst an explicit solution to the Painlevé V differential equation does not exist in general, we will show in the next subsection that it can be used to great effect to extract deep insight into the behavior of the mutual information distribution. In particular, we will use it to provide a systematic method for computing closed-form expressions for the leading-order and correction terms to the mean, variance, and higher order cumulants, as the numbers of antennas grow large. This will allow us to obtain more accurate characterizations than the asymptotic Gaussian approximation, which is based on only the leading order terms of the mean and variance.

B. Systematic Computation of the Cumulants

Applying a similar expansion procedure to that in [1, Sect. IV], we can observe the following large- n series structure:

$$\begin{aligned} \kappa_1 &= n C_{1,0} + \frac{1}{n} C_{1,1} + \frac{1}{n^3} C_{1,2} + \frac{1}{n^5} C_{1,3} \cdots \\ \kappa_2 &= C_{2,0} + \frac{1}{n^2} C_{2,1} + \frac{1}{n^4} C_{2,2} + \cdots \\ \kappa_3 &= \frac{1}{n} C_{3,0} + \frac{1}{n^3} C_{3,1} + \frac{1}{n^5} C_{3,2} \cdots \\ &\vdots \quad \ddots \quad \ddots \end{aligned} \quad (9)$$

The $C_{i,j}$'s are expressions involving P and β , but are independent of n . This gives an important structural understanding of the mutual information distribution in terms of cumulants, and lays the foundation of our computing procedure.

Recalling the definition of the cumulants in (6) and their n -series structure in (9), we start by making the replacement $\lambda \rightarrow n\lambda$ and taking n large, such that $G_n(x)$ admits

$$G_n(x) = n^2 Y(x) + Z(x) + O(n^{-2}) \quad (10)$$

where $Y(x) = \lambda y_1(x) + \lambda^2 y_2(x) + \cdots$, $Z(x) = \lambda z_1(x) + \lambda^2 z_2(x) + \cdots$, both independent of n . Here, y_{ℓ}, z_{ℓ} correspond

to the leading term (in n) and first-order correction terms for the ℓ -th cumulant respectively by the following integrals:

$$C_{\ell,0} = \ell! \int_{-\infty}^{\beta/P} \frac{y_{\ell}}{x} dx, \quad C_{\ell,1} = \ell! \int_{-\infty}^{\beta/P} \frac{z_{\ell}}{x} dx, \quad \ell = 1, 2, \dots \quad (11)$$

The remaining challenge is to compute closed-form formulas for y_{ℓ} and z_{ℓ} , $\ell = 1, 2, \dots$, in order to obtain the corresponding $C_{\ell,0}$ and $C_{\ell,1}$. These computations will be considered in turn.

C. Cumulants to Leading Order in n

We start by substituting (10) into (8) and keeping only the leading order terms in n (i.e., $O(n^6)$), giving the following equation involving $Y(x)$:

$$[(x + \beta + \lambda + 1)Y' - Y]^2 = 4(xY' - Y + \beta)(Y' + \lambda)Y'. \quad (12)$$

Note that this equation completely characterizes all of the cumulants to leading order in n (i.e., based on this, we may compute $C_{\ell,0}$, $\ell = 1, 2, \dots$ in the structure (9)). This can be done upon substituting $Y = \lambda y_1 + \lambda^2 y_2 + \dots$ into (12) and matching the coefficients of λ^k on the left and right-hand side. Using this approach, we systematically obtain equations involving y_k , $k = 1, 2, \dots$ which are elaborated upon below.

1) *Mean and variance*: The case $k = 1$ characterizes the mean of the mutual information. In this case, we obtain the non-linear differential equation involving y_1 :

$$\begin{aligned} & \left[x^2 + 2(\beta + 1)x + (\beta - 1)^2 \right] (y_1')^2 \\ & - 2(x + \beta + 1)y_1' y_1 - 4\beta y_1'^2 + y_1^2 = 0, \end{aligned} \quad (13)$$

which can be explicitly solved to give:

$$y_1 = \frac{\sqrt{x^2 + 2(\beta + 1)x + (\beta - 1)^2} - x - (\beta + 1)}{2}. \quad (14)$$

Integrating y_1 via (11), we derive a closed-form expression for $C_{1,0}$. It is found that $nC_{1,0} = \mu_0$, where μ_0 is the n -asymptotic approximation for the mean mutual information in (1).

The case $k = 2$ characterizes the variance of the mutual information. In this case, we obtain the non-linear differential equation involving y_1 and y_2 :

$$\begin{aligned} & \left\{ \left[x^2 + 2(\beta + 1)x + (\beta - 1)^2 \right] y_1' - (x + \beta + 1)y_1 - 2\beta \right\} y_2' \\ & - [(x + \beta + 1)y_1' - y_1] y_2 = 2x(y_1')^3 - (x - \beta - 1)(y_1')^2 \\ & + 2(y_1')^2 y_1 + y_1' y_1. \end{aligned} \quad (15)$$

Whilst seemingly complicated, quite remarkably, once we substitute y_1 with (14), the coefficient of y_2' becomes zero, and the differential equation (15) collapses to a simple *algebraic equation* in y_2 . Solving this equation for y_2 and integrating via (11), we obtain a closed-form expression for $C_{2,0}$. It is found that $C_{2,0} = \sigma_0^2$, where σ_0^2 is the n -asymptotic approximation for the variance of the mutual information in (2).

2) *Higher cumulants*: The higher order cumulants, beyond the mean and variance, are characterized by the case $k > 2$. Evaluating and investigating these cumulants is important, since it allows one to study deviations from Gaussian for finite dimensions. As we will see in Section V, these higher order cumulants may also be used to “refine” the Gaussian approximation to provide increased accuracy.

For $k > 2$, y_k are found to satisfy the recursion given in (16) at the top of the next page, with initial conditions y_1 and y_2 computed previously. In theory, the recursive equation

(16) enables us to systematically compute the leading order expressions (in n) in closed-form for any desired number of higher order cumulants, in sequence. In its current form, however, (16) appears very complicated. Fortunately, this expression can be simplified considerably by observing that the only term in (16) which involves y_k (or its derivative) is:

$$\begin{aligned} & \left\{ \left[x^2 + 2(\beta + 1)x + (\beta - 1)^2 \right] y_1' - (x + \beta + 1)y_1 - 2\beta \right\} y_k' \\ & - [(x + \beta + 1)y_1' - y_1] y_k, \end{aligned} \quad (17)$$

with all other terms involving the previously computed y_i , $i < k$. As for the variance, quite remarkably, the term (17) simplifies considerably upon noting that the coefficient of y_k' is identically equal to zero. This interesting observation indicates that the computation of the higher cumulants, for $k > 2$, only involves solving simple *algebraic* equations in y_k , rather than non-linear differential equations involving y_k and y_k' . More specifically, the recursive equation (16) collapses to (18) for $k = 3, 4, \dots$, where R_k represents the right-hand-side of (16), which depends only on the previously computed y_i , $i < k$. With this result, y_k can be easily computed, systematically in closed-form, for any value of k . By integrating y_k via (11), we obtain $C_{k,0}$ in (9) in closed form. For the sake of space, we simply write down $C_{3,0}$ in (19) by way of example.

D. Cumulants to Second Order in n (A Brief Sketch)

In addition to computing the cumulants to leading order in n , it is of interest to compute the correction terms (in n), to capture deviations and achieve higher accuracy at finite n . Therefore, we consider $Z(x)$ in (10), which captures the first-order corrections (i.e., characterized by $C_{\ell,1}$, $\ell = 1, 2, \dots$ in (9)). Here we provide some brief details.

Substituting (10) into (8) and setting the second leading terms (in n) to be zero, we can obtain the exact equation involving $Z(x)$, which contains the exact first-order corrections to all leading order cumulants. We omit the equation here due to the space limitation. Through the same procedure as in the last subsection, substituting $Y = \lambda y_1 + \lambda^2 y_2 + \dots$ and $Z = \lambda z_1 + \lambda^2 z_2 + \dots$ and matching the coefficients of λ^k , we are able to compute the z_k 's systematically. Integrating the expressions for z_k , we obtain the first-order correction terms in closed-form. For example, $C_{1,1}$ and $C_{2,1}$ are given in (20) and (21) respectively. The correction terms to the higher cumulants $\kappa_3, \kappa_4, \dots$ are omitted here for the sake of space.

Note that by invoking the same procedure as what has been used for the first two order terms, we can compute the higher order correction terms (i.e. $C_{i,j}$, $i = 1, 2, \dots, j \geq 2$) to each cumulant, which only takes more algebraic efforts.

IV. LARGE SNR ANALYSIS

In this section, we focus on the scenario of fixed n and large SNR P . In this case, as shown in Figs. 1(a)–1(b), the mutual information distribution appears to deviate from Gaussian as the SNR increases, whilst the deviation appears to be much stronger for the case of equal numbers of transmit and receive antennas (i.e., $n_t = n_r$, thus $\beta = 1$) compared with the case $n_t \neq n_r$ ($\beta \neq 1$). Here we draw deep insight into this phenomenon by employing our new closed-form high-order cumulant expansions.

Interestingly, by taking P large in (1), (2), (19), (20) and (21), we obtain different limiting results depending on whether

$$\begin{aligned} & \sum_{i=1}^k y'_i \left\{ \left[x^2 + 2(\beta+1)x + (\beta-1)^2 \right] y'_{k-i+1} - 2(x+\beta+1)y_{k-i+1} \right\} + \sum_{i=1}^k y_i y_{k-i+1} - 4\beta y'_k \\ &= 2 \sum_{i=1}^{k-1} y'_i \left[\sum_{j=1}^{k-i} \left(2x y'_j y'_{k-i-j+1} - 2y'_j y_{k-i-j+1} \right) + (x+\beta-1) y'_{k-i} - y_{k-i} \right] - \sum_{i=1}^{k-2} y'_i y'_{k-i-1} \end{aligned} \quad (16)$$

$$y_k = \frac{1}{2[(x+\beta+1)y'_1 - y_1]} \left\{ R_k - \sum_{i=2}^{k-1} \left[(x^2 + 2(\beta+1)x + (\beta-1)^2) y'_i y'_{k-i+1} - 2(x+\beta+1) y'_i y_{k-i+1} + y_i y_{k-i+1} \right] \right\} \quad (18)$$

$$C_{3,0} = \frac{\beta+1}{2\beta} - \frac{3\beta P}{\beta^2 + 2(\beta+1)\beta P + (\beta-1)^2 P^2} - \frac{1}{2} \frac{(\beta-1)^2 [(\beta^2+1)P^3 + 3\beta(\beta+1)P^2] + 3\beta^2(\beta^2+1)P + \beta^3(\beta+1)}{\beta(\beta^2 + 2(\beta+1)\beta P + (\beta-1)^2 P^2)^{3/2}} \quad (19)$$

$$C_{1,1} = -\frac{\beta+1}{24\beta} + \frac{\beta^3(\beta+1) + 3\beta^2(\beta+1)^2 P + 3\beta(\beta+1)(\beta^2+1)P^2 + (\beta^2+1)(\beta-1)^2 P^3}{24\beta(\beta^2 + 2\beta(\beta+1)P + (\beta-1)^2 P^2)^{3/2}} \quad (20)$$

$$\begin{aligned} C_{2,1} = & \frac{\beta^2+1}{24\beta^2} - \frac{4(\beta+1)\beta^3 P^3 + (9\beta^2 - 42\beta + 9)\beta^2 P^4 + 6\beta(\beta+1)(\beta-1)^2 P^5 + (\beta-1)^4 P^6}{12(\beta^2 + 2\beta(\beta+1)P + (\beta-1)^2 P^2)^3} - \frac{1}{24\beta^2(\beta^2 + 2\beta(\beta+1)P + (\beta-1)^2 P^2)^{7/2}} \left[(\beta^2+1)\beta^7 \right. \\ & + 7\beta^6(\beta+1)(\beta^2+1)P + 7\beta^5(\beta^2+1)(3\beta^2 + 4\beta + 3)P^2 + (\beta+1)(35\beta^4 + 62\beta^2 + 35)\beta^4 P^3 + (35\beta^6 + 9\beta^4 + 8\beta^3 + 9\beta^2 + 35)\beta^3 P^4 \\ & \left. + (\beta+1)(\beta-1)^2(21\beta^4 - 14\beta^3 - 2\beta^2 - 14\beta + 21)\beta^2 P^5 + 7(\beta+1)^2(\beta-1)^6\beta P^6 + (\beta+1)(\beta-1)^8 P^7 \right] \end{aligned} \quad (21)$$

$\beta = 1$ or not. For $\beta = 1$, we obtain¹:

$$\kappa_1 \sim n \log P + \frac{1}{16} \frac{\sqrt{P}}{n} + \dots \quad (22)$$

$$\kappa_2 \sim \frac{1}{2} \log \left(\frac{P}{2} \right) + \frac{1}{32} \left(\frac{\sqrt{P}}{n} \right)^2 + \dots \quad (23)$$

$$\kappa_3 \sim \frac{1}{4n} + \frac{1}{32} \left(\frac{\sqrt{P}}{n} \right)^3 + \dots \quad (24)$$

These large- n —large- P series expansions were also documented in [1]. From these expressions, we see that the correction terms grow faster with P than the leading terms and becomes significant for large P . Therefore the Gaussian approximation (i.e., considering only the leading order mean μ_0 and variance σ_0^2) performs inaccurately in this scenario.

Now, considering $\beta \neq 1$, we obtain:

$$\kappa_1 \sim n \left[\log P - (\beta-1) \log \left(\frac{\beta-1}{\beta} \right) - 1 \right] + \frac{1}{n} \left[\frac{1}{12(\beta-1)} - \frac{1}{12\beta} \right] + \dots \quad (25)$$

$$\kappa_2 \sim \log \left(\frac{\beta}{\beta-1} \right) + \frac{1}{n^2} \left[-\frac{1}{12(\beta-1)^2} + \frac{1}{12\beta^2} \right] + \dots \quad (26)$$

$$\kappa_3 \sim \frac{1}{n} \left[-\frac{1}{(\beta-1)} + \frac{1}{\beta} \right] + \frac{1}{n^3} \left[\frac{1}{6(\beta-1)^3} - \frac{1}{6\beta^3} \right] + \dots \quad (27)$$

where the leading terms of κ_1 and κ_2 match perfectly with results in [4]. Note that these large- n —large- P series expansions are invalid for $\beta = 1$. Based on (25)–(27) we can make some key observations:

Remark 1: In contrast to (22)–(24), as P increases, all terms other than the leading term of κ_1 are *strictly bounded*, converging to constants depending on β . Moreover, as n or β increase, the correction terms have less effect, eventually becoming negligible, even when P is very large.

Remark 2: By virtue of the comments above, for unequal antenna arrays, the n -asymptotic power series representations for the cumulants remain valid for *arbitrary* SNR P ; whilst

in contrast, for equal antenna arrays they break down for sufficiently large P .

These observations explain technically why the Gaussian approximation, based purely on the leading order terms of κ_1 and κ_2 , is relatively more robust to increasing SNRs in Fig. 1(b), compared with the results in Fig. 1(a). Nevertheless, even in the former case, the higher cumulants (e.g., of order $O(1/n)$) still contribute to some deviations from Gaussian. These deviations become particularly significant in the tail (see [9] and [1] for more discussion on this), yielding the Gaussian approximation unsuitable for capturing low outage probabilities (of practical interest). This behavior is shown in Fig. 2(b). In the following, we will make use of our closed-form cumulant expansions to “correct” for these deviations, both in the tail and around the mean, and thereby refine the Gaussian distribution.

V. CHARACTERIZATION WITH EDGEWORTH EXPANSION

Armed with closed-form expressions for $C_{i,j}$ in (9), we draw upon the Edgeworth expansion technique. This approach allows us to start with a Gaussian distribution and to systematically correct this by including higher cumulant effects (i.e., other than the mean and variance), giving an explicit expression for the corrected PDF. Moreover, the cumulative distribution function (CDF), which directly defines the outage probability, can be obtained explicitly through a straightforward integration. With the Edgeworth expansion method, the PDF of the mutual information takes the form [15, Eq. (43)]:

$$p_{I(\mathbf{x};\mathbf{y})}(x) \approx \frac{1}{\sqrt{2\pi}\sigma} e^{-z^2/2} [1 + \mathcal{D}(z)] \quad (28)$$

where $z := (x - \mu)/\sigma$ and

$$\mathcal{D}(z) = \sum_{s=1}^L \sum_{\{k\}} \frac{\text{He}_{s+2r}(z)}{\sigma^{s+2r}} \prod_{\ell=1}^s \frac{1}{k_\ell!} \left(\frac{\kappa_{\ell+2}}{(\ell+2)!} \right)^{k_\ell} \quad (29)$$

is the quantity which determines any deviation from Gaussian. Here, L is a positive integer characterizing how many cumulants are included in the corrected PDF. Note that the second

¹Note that the correction term for the third cumulant κ_3 (i.e., $C_{3,1}$ in (9)) was obtained in the derivation described in Section III-D, but was not presented in this paper due to space limitation.

summation enumerates all sets $k = \{k_1, k_2, \dots, k_s\}$ containing the non-negative integer solutions of the Diophantine equation $k_1 + 2k_2 + \dots + sk_s = s$. For each of these solutions, a corresponding constant r is defined as $r = k_1 + k_2 + \dots + k_s$. In [15], a practical algorithm for computing the $\{k\}$ solutions is proposed in general. The $\text{He}_\ell(z)$ is the ℓ -th Chebyshev-Hermite polynomial, with the explicit form [15, Eq. (13)]

$$\text{He}_\ell(z) = \ell! \sum_{k=0}^{\lfloor \ell/2 \rfloor} \frac{(-1)^k z^{\ell-2k}}{k!(\ell-2k)!2^k} \quad (30)$$

where $\lfloor \cdot \rfloor$ denotes the floor function. These are generated by differentiating the standard normal distribution:

$$\text{He}_\ell(z) = (-1)^\ell g_0^{(\ell)}(z)/g_0(z) \quad (31)$$

where $g_0(z) := \exp(-z^2/2)/\sqrt{2\pi}$ and the superscript (ℓ) denotes the ℓ -th derivative w.r.t. z .

In theory, we can approximate the mutual information distribution with arbitrary accuracy, by taking L sufficiently large in $\mathcal{D}(z)$. We compare the Edgeworth expansions with the Monte Carlo simulations and the Gaussian approximation in Fig. 2(a). Note that for the Edgeworth expansion curves we only use the leading terms of the cumulants (i.e., $C_{\ell,0}/n^{2-\ell}$ in (9)), since as we have shown, the leading terms of the first few cumulants remain valid approximations for arbitrary P if $\beta \neq 1$ (c.f. Remark 2). It is seen that adding only the leading order contribution of κ_3 (i.e., the $L = 1$ case) is sufficient to refine the bulk of the PDF at large SNR (e.g., $P = 30$ dB). To examine the outage probability, we need to derive the CDF of the mutual information. Recalling (31), we have

$$\int_{-\infty}^z \text{He}_\ell(x) g_0(x) dx = -\text{He}_{\ell-1}(z) g_0(z). \quad (32)$$

Therefore, the Edgeworth PDF formula (28) can be immediately integrated to give the CDF:

$$F_{I(\mathbf{x};\mathbf{y})}(x) \approx 1 - Q(z) - g_0(z) \sum_{s=1}^L \sum_{\{k\}} \frac{\text{He}_{s+2r-1}(z)}{\sigma^{s+2r}} \prod_{\ell=1}^s \frac{1}{k_\ell!} \left(\frac{\kappa_{\ell+2}}{(\ell+2)!} \right)^{k_\ell} \quad (33)$$

where $Q(z) := \frac{1}{\sqrt{2\pi}} \int_z^\infty e^{-x^2/2} dx$. Comparison of the CDF curves computed by the Edgeworth expansion (33) with the Gaussian approximation is made in Fig. 2(b) for finite numbers of antennas and large SNR. We see that in the tail region, the Edgeworth expansions with higher cumulants nicely approach the simulations (in this case, $L = 4$ was required to achieve good accuracy), whilst the Gaussian curve strongly deviates. This confirms that, by the virtue of our new cumulant expressions and the Edgeworth expansion, we can compute the outage probability of MIMO systems with high accuracy in closed-form, even at high SNR. This is an important contribution of the paper.

REFERENCES

- [1] Y. Chen and M. R. McKay, "Coulomb fluid, Painlevé transcendents and the information theory of MIMO systems," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4594–4634, Jul. 2012.
- [2] Z. Wang and G. B. Giannakis, "Outage mutual information of space-time MIMO channels," *IEEE Trans. Inf. Theory*, vol. 50, pp. 657–662, Apr. 2004.
- [3] 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. Inf. Theory*, vol. 49, pp. 2545–2561, Oct. 2003.

Fig. 2. PDF and CDF of the mutual information for $n_t = 4$, $n_r = 2$, SNR = 30 dB, computed by Gaussian approximation, Edgeworth expansion, and Monte Carlo simulations.

- [4] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multiantenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1893–1909, Sept. 2004.
- [5] W. Hachem, O. Khorunzhiy, P. Loubaton, J. Najim, and L. Pastur, "A new approach for capacity analysis of large dimensional multiantenna channels," *IEEE Trans. Inf. Theory*, vol. 54, pp. 3987–4004, Sept. 2008.
- [6] P. J. Smith and M. Shafi, "On a Gaussian approximation to the capacity of wireless MIMO systems," in *Proc. of the IEEE Int. Conf. Commun. (ICC)*, New York, Apr. 28–May 2, 2002, pp. 406–410.
- [7] M. R. McKay and I. B. Collings, "General capacity bounds for spatially correlated Rician MIMO channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3121–3145, Sept. 2005.
- [8] H. Shin, M. Z. Win, and J. H. Lee, "Saddlepoint approximation to the outage capacity of MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 2697–2684, Oct. 2006.
- [9] P. Kazakopoulos, P. Mertikopoulos, A. L. Moustakas, and G. Caire, "Living at the edge: A large deviations approach to the outage MIMO capacity," *IEEE Trans. Inf. Theory*, vol. 57, pp. 1984–2007, Apr. 2011.
- [10] S. Li, Y. Chen, and M. R. McKay, "Mutual information distribution of interference-limited MIMO: A joint Coulomb fluid and Painlevé based approach," in *Proc. of the 45th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 6–Nov. 9, 2011.
- [11] E. Teletar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Commun.*, pp. 585–595, Nov–Dec. 1999.
- [12] A. Magnus, "Painlevé-type differential equations for the recurrence coefficients of semi-classical orthogonal polynomials," *J. Comput. Appl. Math.*, vol. 57, pp. 215–237, 1995.
- [13] Y. Chen and A. R. Its, "Painlevé III and a singular linear statistics in Hermitian random matrix ensembles," *J. Approx. Theory*, vol. 162, pp. 270–297, 2010.
- [14] E. Basor and Y. Chen, "Perturbed Hankel determinants," *J. Phys. A: Math. Gen.*, vol. 38, 2005.
- [15] S. Blinnikov and R. Moessner, "Expansions for nearly Gaussian distributions," *Astron. Astrophys. Suppl. Ser.*, no. 130, pp. 193–205, May 1998.