**QUEEN'S
UNIVERSITY
BELFAST**

Tightness of Jensen's Bounds and Applications to MIMO Communications

Yuan, J., Matthaiou, M., Jin, S., & Gao, F. (2017). Tightness of Jensen's Bounds and Applications to MIMO Communications. *IEEE Transactions on Communications*, 65(2), 579-593.
<https://doi.org/10.1109/TCOMM.2016.2623945>

Published in:
IEEE Transactions on Communications

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

© 2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Tightness of Jensen's Bounds and Applications to MIMO Communications

Jide Yuan, *Student Member, IEEE*, Michail Matthaiou, *Senior Member, IEEE*,
Shi Jin, *Member, IEEE*, and Feifei Gao, *Member, IEEE*

Abstract—Due to the difficulty in manipulating the distribution of Wishart random matrices, the performance analysis of multiple-input-multiple-output (MIMO) channels has mainly focused on deriving capacity bounds via Jensen's inequality. However, to the best of our knowledge, the tightness of Jensen's bounds has not yet been rigorously quantified in the general MIMO context. This paper proposes a new methodology for measuring the tightness of Jensen's bounds via the sandwich theorem. In particular, we first compare the tightness of two different pairs of upper/lower bounds for a general class of MIMO channels based on the unordered eigenvalue of the instantaneous correlation matrix and for arbitrary numbers of antennas. The tightness of Jensen's bounds in different channel scenarios is investigated including multiuser MIMO with maximal ratio combining. Our analysis is facilitated by deriving some new results for finite-dimensional Wishart matrices, i.e., for the arbitrary moments of the unordered eigenvalue of central and non-central Wishart matrices. Our results provide very interesting insights into the implications of the system parameters, such as the number of antennas, and signal-to-noise ratio, on the tightness of Jensen's bounds, and showcase the suitability and limitations of Jensen's bounds.

Index Terms—Ergodic capacity, Jensen's inequality, MIMO, random matrix theory, upper and lower bounds.

I. INTRODUCTION

The use of multiple antennas at both ends of a wireless link, known as multiple-input-multiple-output (MIMO) technology, can deliver substantial improvement in spectral efficiency and robustness, with no penalty in either power or bandwidth. For this reason, a vast amount of literature has been devoted to the analysis of the ergodic capacity of both single user (SU)-MIMO systems [1,2] and multiuser (MU)-MIMO systems [3, 4].

Assuming a MIMO system with N_t transmitters and N_r receivers, the propagation channel is expressed by an $N_r \times N_t$ channel matrix $\mathbf{H}$, whose (i, j) th entry describes

the path fading between the j th transmit and the i th receive antenna. Due to the intimate relationship between the performance of a MIMO system and the eigenvalue profile of the instantaneous correlation matrix¹ $\mathbf{H}\mathbf{H}^\dagger$ (or $\mathbf{H}^\dagger\mathbf{H}$), the probabilistic characterization of these eigenvalues is necessary in order to derive analytical expressions for the ergodic capacity. For the prevalent Rayleigh and Ricean fading models, the instantaneous correlation matrix follows a complex Wishart (CW) distribution [5]. Fortunately, the eigenvalue statistics of CW matrices have been widely studied for many decades, for both the central [6,7] and non-central [7, 8] cases. These results have been effectively applied to analyze the performance of SU-MIMO systems in different particular cases. More specifically, the joint probability density function (pdf) of ordered eigenvalues [9,10] was used to evaluate the information theoretical limits of SU-MIMO systems, while the pdfs of the smallest eigenvalue [11] and largest eigenvalue [12,13] have been utilized for designing antenna selection and multichannel beamforming schemes.

Although the performance of MIMO systems has been thoroughly investigated using random matrix theory tools, starting from independent and identically distributed (i.i.d.) Rayleigh fading [1,2,14,15] to the case where correlation is present at one of the two sides [8,11,16,17] or at both sides [18] or even Ricean fading channels [6,19]–[21], exact, closed-form capacity expressions are inherently hard to derive due to the complex distribution of the eigenvalues of $\mathbf{H}\mathbf{H}^\dagger$. In order to obtain more informative analytical insights into MIMO systems, a large number of works has focused on formulating upper and lower capacity bounds for both SU-MIMO (see [2,3,13,22]–[26]) and MU-MIMO systems [4,27] via Jensen's inequality². Note that the Jensen's inequality is a fundamental tool with a plethora of applications in areas such as probability, statistics [28,29] and communications.

However, the tightness of Jensen's bounds (e.g. [23,25,30, 31]) varies according to the eigenvalue profile of $\mathbf{H}\mathbf{H}^\dagger$, signal-to-noise ratio (SNR) and number of antennas. To the best of our knowledge, there appears to be a dearth of references investigating the tightness of Jensen's bounds. A recent work [32], which leveraged the concentration of spectral measure of random matrices [33,34], applied a non-asymptotic analysis to characterize the power offset of Jensen's bounds. The authors of [35] proposed an asymptotic analysis by expanding capacity

Manuscript received Mar 30, 2016; revised September 17, 2016; accepted October 24, 2016. This work was supported in part by the National Science Foundation (NSFC) for Distinguished Young Scholars of China with Grant 61625106 and the NSFC with Grant 61531011. The associate editor coordinating the review of this paper and approving it for publication was S. Durrani.

Jide Yuan and Shi Jin are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing, 210096, P. R. China (e-mail: yuanjide@seu.edu.cn; jinshi@seu.edu.cn).

Michail Matthaiou is with the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast, BT3 9DT, U.K. (e-mail: m.matthaiou@qub.ac.uk).

Feifei Gao is with Tsinghua National Laboratory for Information Science and Technology (TNList), Beijing, 100084, P. R. China (e-mail: feifeigao@ieee.org).

¹The $(\cdot)^\dagger$ superscript denotes the Hermitian transpose.

²Hereafter, we name the bounds on ergodic capacity, which are derived by applying Jensen's inequality, as "Jensen's bounds" for convenience.

as an affine function adding a zero-order term, which is well suited for the high-SNR regime, while [36] gave a similar analysis in the low-SNR regime.

In this paper, we propose an alternative method to measure the tightness of Jensen's bounds based on a basic mathematic idea—sandwich theorem [37]. By capitalizing on the sandwich theorem, we evaluate the offset between two pairs of upper and lower bounds by working exclusively with the eigenvalues of CW matrices. The offsets are analytically determined for several fading scenarios, i.e., i.i.d. Rayleigh, semi-correlated Rayleigh, Ricean channels for SU-MIMO systems and MU-MIMO with maximum ratio combining (MRC) receivers. Our results are based on the theory of finite-dimensional Wishart random matrices and, in particular, on the first positive and negative moment of the unordered eigenvalue of the instantaneous correlation matrix. To facilitate our mathematical derivations, we present a set of new results for CW matrices, namely closed-form expressions for the arbitrary moments of the unordered eigenvalue of central and non-central Wishart matrices. All our results are given in tractable closed-form and can be very easily evaluated. Moreover, we draw several important engineering insights into the tightness of Jensen's bounds and how this is affected by the system parameters, namely number of antennas, spatial correlation, Ricean K -factor and SNR. The results explicitly illustrate the suitability of Jensen's bounds and their limitations as well. As an indicative example, our results indicate that the Jensen's bounds become exact for massive MIMO scenarios since the upper bound converges to the lower bound when the number of antennas on one side of the radio link grows large.

The paper is organized as follows: In Section II, we present the channel and system model under consideration. Section III presents the tightness of Jensen's bounds against the SNR for our general framework. The tightness of Jensen's bounds for different fading scenarios is provided with our new random matrix theory contributions in Section IV. In Section V, we discuss the suitability and limitations of Jensen's bounds. Finally, Section VI concludes the paper. The main mathematical proofs are relegated to the Appendices.

Throughout this paper, vectors and matrices are denoted in bold lowercase $\mathbf{a}$ and bold uppercase $\mathbf{A}$, respectively. The $\otimes$ symbol denotes the Kronecker product. We use the $\text{etr}(\mathbf{A})$ to indicate $e^{\text{tr}(\mathbf{A})}$, where $\text{tr}(\mathbf{A})$ is the trace of the matrix $\mathbf{A}$. The $\det(\mathbf{A})$ and $\mathbf{A}^{-1}$ stand for the determinant and inverse of $\mathbf{A}$, respectively. In addition, $\mathbf{I}_r$ denotes an $r \times r$ identity matrix and the (i, j) th entry of $\mathbf{A}$ is denoted as $\{\mathbf{A}\}_{i,j}$. In some cases, we will write the determinant of $\mathbf{A}$ in terms of its (i, j) th elements (e.g., as $\det\{a_{i,j}\}$). The complex number field is represented by $\mathbb{C}$, and $E[\cdot]$ evaluates the expectation of the input random entity. Also, we use the notation $\mathbf{X} \sim \mathcal{CN}_{p,q}(\mathbf{M}, \mathbf{\Sigma} \otimes \mathbf{\Psi})$ to denote that $\mathbf{X}$ is Gaussian distributed with mean matrix $\mathbf{M} \in \mathbb{C}^{p \times q}$ and covariance matrix $\mathbf{\Sigma} \otimes \mathbf{\Psi}$ where $\mathbf{\Sigma} \in \mathbb{C}^{p \times p}$ and $\mathbf{\Psi} \in \mathbb{C}^{q \times q}$ are Hermitian matrices with $p \leq q$. In this case, $\mathbf{W} = \mathbf{X}\mathbf{X}^\dagger$ has a complex noncentral Wishart distribution $W_p(q, \mathbf{\Sigma}, \mathbf{\Theta})$, where $\mathbf{\Theta} = \mathbf{\Sigma}^{-1}\mathbf{M}\mathbf{M}^\dagger$. It has to be mentioned that if $\mathbf{M} = \mathbf{0}$, $\mathbf{W}$ obeys the complex central Wishart distribution $W_p(q, \mathbf{\Sigma})$. Additionally, $\Gamma(\cdot)$ is the gamma function [38,

Eq. (6.1.5)] and ${}_lF_k(a_1, \dots, a_l; b_1, \dots, b_k; x)$ denotes the generalized hypergeometric function with l, k non-negative integers [39, Eq. (16.2.1)].

II. CHANNEL AND SYSTEM MODEL

A. SU-MIMO Systems

We consider a SU-MIMO system equipped with N_r receive antennas and N_t transmit antennas. We assume that the receiver has perfect channel state information (CSI) while no CSI is available at the transmitter. The transmitter applies a uniform power allocation across transmit antennas. The received complex vector $\mathbf{y}_{\text{su}} \in \mathbb{C}^{N_r \times 1}$ depends on the transmitted vector $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$ according to

$$\mathbf{y}_{\text{su}} = \sqrt{\frac{P}{N_t}} \mathbf{H}_{\text{su}} \mathbf{x} + \mathbf{n}, \quad (1)$$

where $\mathbf{n}$ is the complex noise term, and the entries of $\mathbf{x}$ and $\mathbf{n}$ are independent and identically distributed (i.i.d.) zero-mean Gaussian distributed random variables of unit variance, i.e., $E[\mathbf{x}\mathbf{x}^\dagger] = \mathbf{I}_{N_t}$, $E[\mathbf{n}\mathbf{n}^\dagger] = \sigma^2 \mathbf{I}_{N_r}$, where σ^2 is the noise power. The transmit power P is uniformly allocated to all data streams.

The channel $\mathbf{H}_{\text{su}}$ can be effectively characterized by the matrix

$$\mathbf{H}_{\text{su}} = \sqrt{\frac{K}{K+1}} \mathbf{M} + \sqrt{\frac{1}{K+1}} \mathbf{R}_r^{1/2} \bar{\mathbf{H}} \mathbf{R}_t^{1/2}, \quad (2)$$

where $\mathbf{M}$ is the deterministic channel component satisfying $\text{tr}(\mathbf{M}\mathbf{M}^\dagger) = N_r N_t$, and $\bar{\mathbf{H}}$ is the random channel component containing i.i.d. $\mathcal{CN}(0, 1)$ entries. The Ricean K -factor is defined as the ratio of the power in $\mathbf{M}$ and the average power in $\bar{\mathbf{H}}$. Moreover, $\mathbf{R}_r \in \mathbb{C}^{N_r \times N_r} > \mathbf{0}$ and $\mathbf{R}_t \in \mathbb{C}^{N_t \times N_t} > \mathbf{0}$ are the transmit and receive correlation matrices. In what follows, we refer to

$$m = \min(N_r, N_t), \quad n = \max(N_r, N_t) \quad (3)$$

and define the random matrix $\Phi \in \mathbb{C}^{m \times m} > \mathbf{0}$ as

$$\Phi = \begin{cases} \mathbf{H}_{\text{su}} \mathbf{H}_{\text{su}}^\dagger, & N_t \geq N_r, \\ \mathbf{H}_{\text{su}}^\dagger \mathbf{H}_{\text{su}}, & N_t < N_r. \end{cases} \quad (4)$$

Also, let us denote, for convenience,

$$(\mathbf{R}_m, \mathbf{R}_n, \bar{\mathbf{M}}) = \begin{cases} (\mathbf{R}_r, \mathbf{R}_t, \mathbf{M}), & N_t \geq N_r, \\ (\mathbf{R}_t, \mathbf{R}_r, \mathbf{M}^\dagger), & N_t < N_r. \end{cases} \quad (5)$$

Denoting by Λ_m and Λ_n the respective diagonal eigenvalue matrices of $\mathbf{R}_m$ and $\mathbf{R}_n$, the ergodic capacity can be written in the following two forms³ [41, Eq. (2)]

$$C^{\text{su}} = E_{\tilde{\mathbf{H}}} \left[\log_2 \det \left(\mathbf{I}_m + \frac{\rho}{N_t} \Lambda_m \tilde{\mathbf{H}} \Lambda_n \tilde{\mathbf{H}}^\dagger \right) \right], \quad (6)$$

and

$$C^{\text{su}} = m E_\lambda \left[\log_2 \left(1 + \frac{\rho}{N_t} \lambda \right) \right], \quad (7)$$

³When no CSI is available at the transmitter, the equal power allocation strategy, that appears in the expressions (6) and (7), is a robust transmission scheme since it maximizes the worst-case capacity in many cases [1,40] (i.e., the maxmin property). In the cases of general correlation structures and non-zero mean channels, this expression yields an achievable rate (lower bound on capacity) yet, we denote this expression as "capacity" for the sake of consistency with the vast body of MIMO literature [3,18,19,22,23,41,42].

where $\tilde{\mathbf{H}} \sim CN_{m,n} \left(\sqrt{\frac{K}{K+1}} \mathbf{R}_m^{-\frac{1}{2}} \mathbf{M} \mathbf{R}_n^{-\frac{1}{2}}, \frac{1}{K+1} \mathbf{I}_m \otimes \mathbf{I}_n \right)$, $\rho = P/\sigma^2$ is the average SNR, and λ is the single unordered eigenvalue of Φ . Assuming $f_\lambda(\Phi)$ is the marginal probability distribution function (pdf) of the unordered eigenvalue λ , we reformulate (7) as

$$C^{\text{su}} = m \int_0^\infty \log_2 \left(1 + \frac{\rho}{N_t} \lambda \right) f_\lambda(\Phi) d\lambda. \quad (8)$$

This is the general capacity formula of SU-MIMO systems which encompasses many MIMO channels. However, in most cases of interest, it is really complicated, if not impossible, to evaluate analytically the theoretic expression in (8). Thus, for convenience, many prior works (see [22]–[24,41]) choose to characterize the ergodic capacity via its upper/lower bounds using Jensen's inequality. In what follows, we elaborate on the tightness of Jensen's bounds and provide important insights into their ability to approximate the exact MIMO capacity.

B. MU-MIMO Systems

We now consider a MU-MIMO MRC system with N single-antenna users and an M -antenna BS, where each user transmits its signal to the BS in the same time-frequency channel. Assuming the system is single-cell with no interference from neighboring cells, the received complex vector can be written as [4]

$$\mathbf{y}_{\text{mu}} = \sqrt{P_{\text{mu}}} \mathbf{G} \mathbf{x} + \mathbf{n}, \quad (9)$$

where $\mathbf{n}$ is a vector of Gaussian noise defined as in (1), $\sqrt{P_{\text{mu}}} \mathbf{x}$ represents the $N \times 1$ vector containing the transmitted signals from all users, and $\mathbf{x}$ is an i.i.d. zero-mean Gaussian distributed random vector of unit variance, P_{mu} denotes the average transmitted power of each user, and $\mathbf{G}$ is the $M \times N$ MIMO channel matrix between the BS and the N users, which embraces independent fast fading, geometric attenuation and log-normal shadow fading [3], given as

$$\mathbf{G} = \mathbf{H} \mathbf{B}^{1/2}, \quad (10)$$

where $\mathbf{H}$ is the the fast fading matrix, whose element $\{\mathbf{H}\}_{i,j}$ represents the channel from j th user to the i th antenna of the BS with $\{\mathbf{H}\}_{i,j} \sim CN(0, 1)$, while $\mathbf{B}$ is the $N \times N$ diagonal matrix with $\{\mathbf{B}\}_{j,j} = \beta_j$ representing the large-scale fading (LSF) coefficient, which is assumed to be constant across the antenna array.

We consider the case that the BS has perfect CSI and deploys MRC receiver. Thus, the received signal vector at the BS is given by

$$\mathbf{r} = \sqrt{P_{\text{mu}}} \mathbf{G}^H \mathbf{G} \mathbf{x} + \mathbf{G}^H \mathbf{n}. \quad (11)$$

By the law of matrix multiplication, the spectral efficiency (in bit/s/Hz) of the j th user is given by [4]

$$R_j^{\text{mu}} = E \left[\log_2 \left(1 + \frac{\rho_{\text{mu}} \|\mathbf{g}_j\|^2}{\sum_{k=1, k \neq j}^N |\tilde{g}_k|^2 + 1} \right) \right] \quad (12)$$

where $\tilde{g}_k = \mathbf{g}_j^H \mathbf{g}_k / \|\mathbf{g}_j\|$, and $\rho_{\text{mu}} = P_{\text{mu}}/\sigma^2$ represents the

SNR. Conditioned on $\mathbf{g}_j$, $\tilde{g}_k$ is a Gaussian RV with zero mean and variance β_k which does not depend on $\mathbf{g}_j$. Therefore, $\tilde{g}_k$ is Gaussian distributed and independent of $\mathbf{g}_j$ with $\tilde{g}_k \sim \mathcal{CN}(0, \beta_k)$.

III. GENERAL ANALYSIS FRAMEWORK FOR THE TIGHTNESS OF JENSEN'S BOUNDS

In this section, we consider two pairs of Jensen's bounds and quantify rigorously their offsets against the SNR. Note that these two pairs of bounds have been extensively applied in characterizing the capacity of MIMO systems (see [23,35]). Noting the fact that the evaluation of ergodic capacity can be reduced from the matrix problem in (6) into a real-valued scalar problem in (7), we first define two pairs of Jensen's bounds for our general framework as

$$J_1 \triangleq \begin{cases} C^{\text{su}} \leq m \log_2 \left(1 + \frac{\rho}{N_t} E[\lambda] \right), \\ C^{\text{su}} \geq m \log_2 \left(1 + \frac{\rho}{N_t E[\lambda^{-1}]} \right), \end{cases} \quad (13)$$

and

$$J_2 \triangleq \begin{cases} C^{\text{su}} \leq E_\Phi \left[\log_2 \left(\frac{\rho}{N_t} \det(\Phi) \right) \right] \\ \quad + m \log_2 \left(1 + \frac{N_t}{\rho} E[\lambda^{-1}] \right) \\ C^{\text{su}} \geq E_\Phi \left[\log_2 \frac{\rho}{N_t} \det(\Phi) \right] + m \log_2 \left(1 + \frac{N_t}{\rho E[\lambda]} \right). \end{cases} \quad (14)$$

It is important to note that J_1 and J_2 can be easily derived via the convexity or concavity of the corresponding functions, respectively. To evaluate the tightness of the pairs of bounds, we measure the offsets between the corresponding upper and lower bounds in J_1 and J_2 .

Theorem 1: Let Φ be a $m \times m$ matrix as in (4) with unordered eigenvalue λ . The offsets of the two pairs of bounds defined in (13) and (14) are given by

$$\Delta_1 = m \log_2 \left(1 + \frac{\omega - 1}{\frac{N_t}{\rho} E[\frac{1}{\lambda}] + 1} \right) \quad (15)$$

and

$$\Delta_2 = m \log_2 \left(1 + \frac{\omega - 1}{\frac{\rho}{N_t} E[\lambda] + 1} \right) \quad (16)$$

respectively, where $\omega = E[\frac{1}{\lambda}] E[\lambda]$.

Proof: The result can be derived directly by subtracting the lower bound from the upper bound in (13) and (14), respectively. ■

It is important to note that this result holds for arbitrary number of antennas and SNR. Since the offsets vary according to the statistical characteristics of the unordered eigenvalue, the tightness of both pairs of bounds is different for each MIMO channel. We now investigate the tightness of the proposed bounds as a function of the SNR. We compare the two offsets and determine which admits the smallest value depending on a SNR threshold.

Corollary 1: Define ρ_t as the threshold such that

$$\begin{cases} \Delta_1 \leq \Delta_2, & \rho \leq \rho_t, \\ \Delta_1 > \Delta_2, & \rho > \rho_t, \end{cases} \quad (17)$$

then, the relationship between ρ_t and the moments of the unordered eigenvalue is given by

$$\rho_t = N_t \sqrt{\frac{E\left[\frac{1}{\lambda}\right]}{E[\lambda]}}, \quad (18)$$

and the offset at ρ_t can be expressed as

$$\Delta_{\rho_t} = \frac{m}{2} \log_2(\omega). \quad (19)$$

Proof: The proof follows by comparing (15) and (16), and then simplifying the result. ■

It is important to note that ρ_t is only related to the number of transmit antennas and eigenvalue moments, namely, the first positive and negative moment of the unordered eigenvalue. The threshold indicates the behavior of J_1 and J_2 : particularly, J_1 is more suitable for low-SNR scenarios while J_2 performs better at high-SNRs. In other words, Δ_{ρ_t} can be regarded as the maximum offset in the whole SNR regime since the offset in other SNR regimes can be reduced by selecting different pairs of Jensen's bounds. To gain further insights, we further investigate the offsets for these two extreme conditions as follows:

- As the SNR grows large, i.e., $\rho \rightarrow \infty$, the two offsets converge to

$$\Delta_1 = m \log_2(\omega), \quad \Delta_2 = 0. \quad (20)$$

We see that the parameter N_t disappears in both offsets as ρ grows large, with Δ_2 converging to zero, while Δ_1 grows linearly with the minimum dimension of the channel matrix. In this case, the result shows that the capacity can be precisely predicted by J_2 at high-SNRs via the sandwich theorem.

- As the SNR decreases, i.e., $\rho \rightarrow 0$, the two offsets converge to

$$\Delta_2 = m \log_2(\omega), \quad \Delta_1 = 0. \quad (21)$$

Interestingly, we see that the two offsets have an inverse behavior in the low-SNR regime, which means that the properties of Δ_2 at low-SNR can be derived by analyzing Δ_1 in the high-SNR regime, and vice versa. In the latter case, the upper and lower bound in J_1 describe accurately the ergodic capacity.

- Two insights may be useful:

$$\Delta_{\rho_t} = \Delta_1/2, \quad \rho \rightarrow \infty. \quad (22)$$

$$\Delta_{\rho_t} = \Delta_2/2, \quad \rho \rightarrow 0. \quad (23)$$

Since the maximum offset happens to be half the offset in the extreme SNR conditions, this symmetry of Δ_1 and Δ_2 is again illuminated.

To recap, *Theorem 1* shows the symmetry of the two offsets around the SNR, which is very convenient for our following analysis since we can investigate the behavior of J_1 and J_2 in the high-SNR regime while omitting the low-SNR analysis.

IV. TIGHTNESS OF JENSEN'S BOUNDS IN FADING CHANNELS

Apart from the SNR, the tightness of Jensen's bounds depends on the random channel matrix via its unordered eigenvalue. In this section, we also present some new results on Wishart random matrix theory, which are used to showcase the tightness of Jensen's bounds for various fading models.

A. SU i.i.d. Rayleigh fading channels

For i.i.d. Rayleigh fading channels, the channel is statistically described by $\mathbf{H} \sim CN_{m,n}(\mathbf{0}, \mathbf{I}_m \otimes \mathbf{I}_n)$. As shown in *Theorem 1*, to assess the tightness of Δ_1 and Δ_2 , it is necessary to derive the first positive and first negative moment of the unordered eigenvalue of the channel matrix. Noting that the matrix $\mathbf{\Phi} \sim W_m(n, \mathbf{I}_m)$, the desired results are given as [36, Lemma 4, Lemma 6]

$$E[\lambda] = n, \quad (24)$$

and

$$E\left[\frac{1}{\lambda}\right] = \frac{1}{n-m}, \quad n \neq m. \quad (25)$$

Proposition 1: For i.i.d. Rayleigh fading channels, the two offsets of J_1 and J_2 with $n \neq m$, are given in closed-form by

$$\Delta_1 = m \log_2 \left(1 + \frac{m\rho}{N_t + \rho(n-m)} \right) \quad (26)$$

and

$$\Delta_2 = m \log_2 \left(1 + \frac{mN_t}{(\rho n + N_t)(n-m)} \right) \quad (27)$$

respectively. Moreover, ρ_t is expressed as

$$\rho_t = \frac{N_t}{\sqrt{n(n-m)}}. \quad (28)$$

Proof: Substituting (24) and (25) into *Theorem 1* we can obtain the desired result. ■

Our results in *Proposition 1* give mathematical conclusions for the tightness of the two pairs of Jensen's bounds, which are valid for all SNRs and arbitrary antenna configurations as long as $n \neq m$. The results are insightful since they show explicitly that the two offsets reduce with increasing n . A set of numerical results is presented in Fig. 1, where the theoretical results are compared against the empirical offsets, obtained via Monte-Carlo simulations. We see that the two offsets have indeed a mirror behavior with the symmetric center ρ_t varying with the dimensions of the channel matrix. Note that the Jensen's bounds become inherently loose when the minimum number of antennas grows large (e.g., Δ_1 grows up to 12.84 bit/s/Hz with $m = 5$), since the variances of the involved random quantities become much higher. However, it is important to note that the two pairs of bounds are rather tight in the cases with small m and large n , which correspond to a massive MIMO setup. To further illuminate the influence of matrix dimensions on the tightness of the bounds, we present the following corollaries.

Corollary 2: For i.i.d. Rayleigh fading channels, as the number of antennas on one side of the MIMO link grows large, i.e., $n \rightarrow \infty$, Δ_1 and Δ_2 converge to 0.

Fig. 1. Offsets Δ_1 and Δ_2 against the SNR ρ for i.i.d. Rayleigh fading channels. The different antenna configurations are denoted as (m, n) .

The result holds for arbitrary m in the arbitrary SNR regime, and implies that both J_1 and J_2 are intimately tight for MIMO systems with large array antennas. In order to get more insights, we further investigate the tightness of J_1 for finite dimensions in the high-SNR regime.

Corollary 3: As the SNR grows large, i.e., $\rho \rightarrow \infty$, the offset of J_1 converges to

$$\Delta_1 = m \log_2 \left(\frac{n}{n-m} \right). \quad (29)$$

We see that, for high-SNR, the tightness of J_1 is only dependent on the numbers of antennas. The offset grows with the minimum number of antennas, while varying inversely with the maximum number of antennas. The result is confirmed in Fig. 2 as a function of n , for a system with SNR = 20dB. The minimum dimension of the channel

Fig. 2. Δ_1 against the maximum number of antennas for i.i.d. Rayleigh fading channels with $\rho = 20$ dB.

matrices, which typically corresponds to the numbers of user antennas in practice, varies from 2 to 5. We find that all curves converge to zero with an increasing number of n , indicating that the Jensen's bounds are particularly tight for massive MIMO configurations (e.g., Δ_1 less than 0.5 bit/s/Hz with an antenna configuration (4, 120)).

Here, we further study J_1 by examining the following cases.

- For i.i.d. Rayleigh fading channels, adding k antennas at both ends of link, Δ_1 looses according to

$$\Delta_1 = (m+k) \log_2 \left(\frac{n+k}{n-m} \right). \quad (30)$$

- For i.i.d. Rayleigh fading channels, as the number of transmit and receive antennas grows with ratio $\beta = N_T/N_R \geq 1$, Δ_1 simplifies to

$$\Delta_1 = m \log_2 \left(\frac{\beta}{\beta-1} \right), \quad (31)$$

showing that the tightness of the bounds is inversely proportional to m with fixed β . We point out that (30) and (31) showcase the impact of the minimum number of antennas, m , on the tightness of J_1 in the high-SNR regime and for arbitrary number of antennas.

B. SU semi-correlated Rayleigh fading channels

For semi-correlated Rayleigh channels, we assume that the correlation is at the either end of the link, which makes the expressions for the marginal distribution of the unordered eigenvalue of Φ different. For convenience, we give a separate treatment of two cases, namely

$$(s, t, \Lambda_m, \Lambda_n) = \begin{cases} (m, n, \Theta_m, \mathbf{I}_n), & \Theta_m \neq \mathbf{I}_m, \\ (n, m, \mathbf{I}_m, \Theta_n), & \Theta_n \neq \mathbf{I}_n, \end{cases}$$

to represent the min semi-correlated and max semi-correlated Rayleigh fading channels, respectively. We now establish some new results on the unordered eigenvalue λ .

Lemma 1: For semi-correlated Rayleigh fading channels, with the unordered eigenvalue of Φ given as λ , the expectation of the arbitrary p th moment of λ is given as

$$E[\lambda^p] = \frac{1}{m \prod_{i < j}^s (\theta_j - \theta_i)} \times \sum_{j=(v-\text{sgn}(p)p)/2+1}^{(v+\text{sgn}(p)p)/2} \frac{\Gamma(t-s+j+p)}{\Gamma(t-s+j)} \det(\tilde{\mathbf{D}}_j^p), \quad (32)$$

where $v = s - p + \text{sgn}(p)s$, $\text{sgn}(\cdot)$ is the sign function, and $\tilde{\mathbf{D}}_j^p$ is an $s \times s$ matrix whose (l, k) th entry is

$$\{\tilde{\mathbf{D}}_j^p\}_{l,k} = \begin{cases} \theta_l^{k-1}, & l = 1, \dots, s, \quad k \neq j, \\ \theta_l^{k+p-1}, & l = 1, \dots, s, \quad k = j, \end{cases} \quad (33)$$

where θ_l is the l th eigenvalue of the correlation matrix Θ_m (or Θ_n).

Proof: See Appendix A. ■

This lemma presents a new result for the statistical density of an unordered eigenvalue of a complex semi-correlated central Wishart matrix. It is important to mention that this

result holds for any p , even for $p < 0$. From (32), we notice that when $p < 0$ with $t < s$, some items in the expression become ill-defined. This is due to the fact that for max semi-correlated Rayleigh channels, the negative moments do not exist.

Lemma 2: For semi-correlated Rayleigh fading channels, the expectation of the unordered eigenvalue λ is given as

$$E[\lambda] = \frac{t}{m} \sum_{i=1}^s \theta_i \quad (34)$$

and its first negative moment for the min semi-correlated case is given by⁴

$$E\left[\frac{1}{\lambda}\right] = \frac{1}{m(n-m)} \sum_{i=1}^m \frac{1}{\theta_i}, \quad (35)$$

while for the max semi-correlated case, $E\left[\frac{1}{\lambda}\right]$ does not exist.

Proof: See Appendix B. ■

We note that (34) and (35) can be reduced to (24) and (25) when the correlation matrix is an identity matrix. Now, we are ready to derive the desired offsets for min semi-correlated Rayleigh fading channels.

Proposition 2: For min semi-correlated Rayleigh fading channels, the two offsets of J_1 and J_2 are given in closed-form as follows

$$\Delta_1 = m \log_2 \left(1 + \frac{n\rho A - H\rho(n-m)}{N_t + H\rho(n-m)} \right) \quad (36)$$

and

$$\Delta_2 = m \log_2 \left(1 + \frac{nN_t A - HN_t(n-m)}{(n\rho A + HN_t)(n-m)} \right), \quad (37)$$

where $A = \frac{1}{m} \sum_{i=1}^m \theta_i$ represents the arithmetic mean of $\theta_{i=1,\dots,m}$ and $H = \frac{m}{\sum_{i=1}^m \theta_i^{-1}}$ is the harmonic mean of $\theta_{i=1,\dots,m}$. Moreover, ρ_t is given by

$$\rho_t = N_t \sqrt{\frac{A}{(n-m)H}}. \quad (38)$$

Proof: Substituting (34) and (35) into Theorem 1 and (18) yields the desired result. ■

It is important to note that the results give exact closed-form expressions for the offsets, which apply for all SNRs and arbitrary antenna configurations. Fig. 3 shows the offsets of two Jensen's bounds for min semi-correlated Rayleigh channels. The correlation matrix $\mathbf{R}_m$, whose (i, j) th entry is $e^{-0.05d^2(i-j)^2}$, corresponds to a d -wavelength antenna separation and a broadside Gaussian power azimuth spectrum with 2° root-mean-square spread [41]. The analytical expressions based on (36) and (37) match precisely with the simulation results, and perform quite similar as Δ_1 and Δ_2 in Fig. 1. By noting that the antenna configurations are identical to those considered in Fig. 1, it is straightforward to see that J_1 and J_2 are much looser in the semi-correlated channel scenario, since, in general, correlation increases the spread of

⁴Note that the proof of $E[\lambda^{-1}]$ generalizes the result of [24] by introducing the concept of generalized Vandermonde determinants and holds for any arbitrary moment.

Fig. 3. Offsets Δ_1 and Δ_2 for a min semi-correlated channel with correlation matrix modeled as $\{\mathbf{R}_m\}_{i,j} = e^{-0.05(2i-2j)^2}$. The different antenna configurations are denoted as (m, n) .

the unordered eigenvalue.

Corollary 4: For min semi-correlated Rayleigh fading channels, as the SNR grows large, i.e., $\rho \rightarrow \infty$, the offsets of J_1 and J_2 converge respectively to

$$\Delta_1 = m \log_2 \left(\frac{nA}{(n-m)H} \right) \text{ and } \Delta_2 = 0. \quad (39)$$

Proof: This result can be directly derived by substituting (34) and (35) into (20). ■

As expected, the results in (39) reduce to (29) when the correlation matrix is an identity matrix. Therefore, the offsets in (39) show similar trend against the maximum number of antennas, as in the case of i.i.d. Rayleigh fading channels. It is important to note that both A and H depend on the m eigenvalues of the correlation matrix so that the impact of m cannot be separately analyzed.

Corollary 5: For min semi-correlated Rayleigh fading channels, as $n \rightarrow \infty$, Δ_1 reduces to

$$\Delta_1 = m \log_2 \left(\frac{A}{H} \right) \geq 0 \text{ and } \Delta_2 = 0. \quad (40)$$

Note that the equality of Δ_1 holds if and only if all elements in Θ_m are the same.

From (40), we see that in a massive MIMO scenario, the tightness of J_1 is only related with the eigenvalues of the correlation matrix. We also note that the tightness of the bounds is influenced by the ratio of the arithmetic mean and harmonic mean of the eigenvalues of Θ_m . This ratio, as we known, is equal or greater than 1, and decreases with decreasing variance of θ_i . This implies that the Jensen's bounds are less tight for semi-correlated Rayleigh channels compared with i.i.d. Rayleigh channels.

The results of Corollary 4 and Corollary 5 are confirmed in Fig. 4, where we compare the exact closed-form of offset Δ_1 based on (36), and its asymptotic limit based on (40) for an SNR = 30dB. The correlation matrix is again modeled as

Fig. 4. Δ_1 for min semi-correlated Rayleigh channels against the maximum number of antennas with $m = 2, 3, 4$, and $\rho = 30\text{dB}$.

$\{\mathbf{R}_m\}_{i,j} = e^{-0.2(i-j)^2}$. In all cases, given as $m = 2, 3, 4$, there is an exact agreement between the analytical results and the Monte-Carlo simulations. However, we see that the asymptotic limit of Δ_1 is larger than 6.5 bits/s/Hz when $m = 4$; this means that J_1 is not suitable to measure the ergodic capacity in the high SNR regime when the number of user antennas grows large, even for a very large n . As shown in Fig. 3, J_2 would be more appropriate for semi-correlated Rayleigh channels when the SNR grows large.

C. SU i.i.d. Ricean fading channels

For i.i.d. Ricean fading channels, we model $\tilde{\mathbf{H}} \sim CN_{m,n}(\sqrt{\frac{K}{K+1}}\tilde{\mathbf{M}}, \frac{1}{K+1}\mathbf{I}_m \otimes \mathbf{I}_n)$ by simplifying $\mathbf{\Lambda}_m = \mathbf{I}_m$ and $\mathbf{\Lambda}_n = \mathbf{I}_n$. With this model, Φ follows a complex noncentral Wishart distribution, denoted as $\Phi \sim W_m(n, \frac{1}{K+1}\mathbf{I}_m, \Xi)$, where $\Xi \triangleq K\tilde{\mathbf{M}}\tilde{\mathbf{M}}^\dagger$.

Lemma 3: Let $\epsilon_1, \dots, \epsilon_m$ be the m eigenvalues of Ξ . Then, the first positive and first negative moment of the singular unordered eigenvalue λ of Φ are

$$E[\lambda] = \frac{\kappa}{2(K+1)} \det\{\mathbf{U}\} \quad (41)$$

and

$$E\left[\frac{1}{\lambda}\right] = 2\kappa(K+1) \det\{\tilde{\mathbf{U}}\}, \quad (42)$$

respectively, where $\kappa \triangleq \frac{\sqrt{2}^{m-mn} (\prod_{i=1}^m \epsilon_i)^{(m-n)/2}}{m \prod_{k=1}^m (\epsilon_k - \epsilon_{k-1})}$. Note that $\mathbf{U}$ and $\tilde{\mathbf{U}}$ are $m \times m$ matrices whose (l, k) th entry is given by

$$\{\mathbf{U}\}_{l,k} = \begin{cases} Q_{n-m+2k-1, n-m}(\sqrt{2\epsilon_l}, 0), & k = 1, \dots, m-1, \\ Q_{n+m+1, n-m}(\sqrt{2\epsilon_l}, 0), & k = m. \end{cases} \quad (43)$$

and

$$\{\tilde{\mathbf{U}}\}_{l,k} = \begin{cases} Q_{n-m-1, n-m}(\sqrt{2\epsilon_l}, 0), & k = 1, \\ Q_{n-m+2k-1, n-m}(\sqrt{2\epsilon_l}, 0), & k = 2, \dots, m, \end{cases} \quad (44)$$

respectively. In the above, $Q_{p,q}(a, b)$ is the Nuttall Q -function, defined in [21] as

$$Q_{p,q}(a, b) = \int_b^\infty x^p e^{-\frac{x^2+a^2}{2}} I_q(ax) dx, \quad (45)$$

and $I_q(\cdot)$ is the q th order modified Bessel function of the first kind.

Proof: See Appendix C. ■

We see that the results in (41) and (42) can be easily programmed and efficiently calculated. Additionally, the Nuttall Q -functions in (43) and (44) have closed-form representations shown in (93) since the lower limit of the integral in (45) is equal to 0.

Proposition 3: For i.i.d. Ricean fading channels, the two offsets of J_1 and J_2 are given in closed-form by

$$\Delta_1 = m \log_2 \left(1 + \frac{\rho \kappa^2 \det\{\tilde{\mathbf{U}}\} \det\{\mathbf{U}\} - \rho}{\gamma \kappa \det\{\tilde{\mathbf{U}}\} + \rho} \right), \quad (46)$$

and

$$\Delta_2 = m \log_2 \left(1 + \frac{\gamma \kappa^2 \det\{\tilde{\mathbf{U}}\} \det\{\mathbf{U}\} - \gamma}{\rho \kappa \det\{\mathbf{U}\} + \gamma} \right), \quad (47)$$

respectively, where $\gamma \triangleq 2N_t(K+1)$. Moreover, ρ_t is given by

$$\rho_t = 2(K+1)N_t \sqrt{\frac{\det(\tilde{\mathbf{U}})}{\det(\mathbf{U})}}. \quad (48)$$

Proof: Substituting (41) and (42) into *Theorem 1* and (18) yields the desired result. ■

It is important to note that the results in *Proposition 3* apply for arbitrary dimensions of the channel matrix and all SNRs. However, it is hard to unveil the impact of each parameter on the offsets due to the presence of the determinant inside the logarithmic expression. In order to derive more informative insights, we analyze the single-antenna user case, and obtain a simpler result given in the next corollary.

Corollary 6: For i.i.d. Ricean fading systems with a single-antenna user, i.e., $m = 1$, the closed-form expressions of the two offsets are given as

$$\Delta_1 = \log_2 \left(1 + \frac{2\rho\xi(n+\epsilon) - 2\rho}{\gamma\xi + 2\rho} \right), \quad (49)$$

and

$$\Delta_2 = \log_2 \left(1 + \frac{\gamma\xi(n+\epsilon) - \gamma}{2\rho(n+\epsilon) + \gamma} \right), \quad (50)$$

respectively, where $\xi \triangleq \frac{\Gamma(n-1) \left(e^\epsilon \sum_{k=0}^{n-2} \frac{(-\epsilon)^k}{k!} - 1 \right)}{(-\epsilon)^{n-1} e^\epsilon}$ and ϵ is the eigenvalue of $\Xi \in \mathbb{C}^{1 \times 1}$.

Proof: See Appendix D. ■

The above results place no restrictions on the maximum dimension of the matrix. Most importantly, the two offsets for single input multiple output systems can be easily evaluated since they only involve standard functions.

Corollary 7: For i.i.d. Ricean fading systems with single-

antenna users, as $n \rightarrow \infty$,

$$\lim_{n \rightarrow \infty} \Delta_1 = 0. \quad (51)$$

Proof: By applying a Taylor series expansion, e^ϵ can be divided into two parts, given as

$$e^{-\epsilon} = \sum_{k=0}^{n-2} \frac{(-\epsilon)^k}{k!} + \sum_{k=n-1}^{\infty} \frac{(-\epsilon)^k}{k!}. \quad (52)$$

By noting that ξ involves the first part on the right side of (52), we substitute (52) in ξ and after some basic manipulations, we obtain

$$\xi = O\left(\frac{1}{n-1}\right), \quad (53)$$

where $O(\cdot)$ is expressed as $f(n) = O(g(n))$ as $n \rightarrow \infty$, if $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 1$. Then, substituting (53) into (49) and (50) yields the result. ■

Fig. 5. Offsets Δ_1 and Δ_2 against the SNR for i.i.d. Ricean fading channels. The Ricean factor $K = 5$ dB and the different antenna configurations are denoted as (m, n) .

From *Corollary 7*, the tightness of bounds for i.i.d. Ricean fading channels shows the same trend against the maximum dimension of the channel matrix. Therefore, Jensen's bounds remain very tight for single-input-multiple-output (SIMO) systems in i.i.d. Ricean fading channels with large receive arrays.

Fig. 5 gives the analytical Δ_1 based on (46), analytical Δ_2 based on (47), and Monte-Carlo results for different antenna configurations. The entries of matrix $\mathbf{M}$ are generated randomly while satisfying $\text{tr}(\mathbf{M}\mathbf{M}^H) = mn$. The figure shows perfect agreement between the analytical expressions and the simulations. We also see that J_1 and J_2 are tighter for the Ricean scenario compared with the semi-correlated Rayleigh channel (e.g., Δ_1 is 1.6 bits/s/Hz in Fig. 5 compared with 2.9 bits/s/Hz in Fig. 3, for $(m, n) = (2, 6)$ with SNR = 25 dB). This is due to the fact that the MIMO channel is now more deterministic and, hence, the spread of the unordered eigenvalue is smaller.

D. MU-MIMO MRC Systems

For a MU-MIMO MRC system modeled as in (11), whose spectral efficiency is given in (12), we note that the desired signal power and interference power are independent of each other, which makes it possible to derive them separately. We first notice that $\{\mathbf{g}_j^H \mathbf{g}_j\} \sim W_1(M, \mathbf{I}_M)$ in (12), thus, we simply get [36, Lemma 4]

$$E[\rho_{\text{mu}} \mathbf{g}_j^H \mathbf{g}_j] = \rho_{\text{mu}} M \beta_j \quad (54)$$

and

$$E[(\rho_{\text{mu}} \mathbf{g}_j^H \mathbf{g}_j)^{-1}] = \frac{1}{\rho_{\text{mu}} (M-1) \beta_j}. \quad (55)$$

Then, an exact expression of the expectation of the noise-plus-interference (NPI) power is derived, as well as an approximate formula of the first negative moment of the NPI power.

Lemma 4: The first positive and first negative moment of the noise-plus-interference power $I = 1 + \rho_{\text{mu}} \sum_{k=1, k \neq j}^N |\tilde{g}_k|^2$ are given as

$$E[I] = 1 + \rho_{\text{mu}} \varpi \quad (56)$$

and

$$E[I^{-1}] = \sum_{k=1, k \neq j}^N \frac{-L(\rho_{\text{mu}} \beta_k)}{\rho_{\text{mu}} T(j, k, \mathbf{B})}, \quad (57)$$

where $L(x) = e^{\frac{1}{x}} \text{Ei}(-1/x)$, $T(k, l, \mathbf{B}) = \beta_k \prod_{l=1, l \neq k, l \neq j}^N (1 - \beta_l / \beta_k)$, $\varpi = \sum_{k=1, k \neq j}^N \beta_k$ and $\text{Ei}(\cdot)$ is the exponential integral function [38, Eq. (5.1.2)].

Proof: See Appendix E. ■

Note that the results in (56) can be very easily evaluated. Both $E[I]$ and $E[I^{-1}]$ are only related to the users' positions while independent of the number of antennas at the BS. This is confirmed in Fig. 6, in which we compare the analytical expression of $E[I^{-1}]$ based on (57) and Monte-Carlo results for a MU-MIMO MRC system. We set the number of antennas at BS as $M = 128$, and assume that the number of users $N = 2, 5, 15$. All users are dropped randomly in a cell with radius of 1000 meters, and we assume that no user is closer to the BS than $r_h = 100$ meters. Without loss of generality, the large-scale fading is modeled via $\beta_j = z_j / (r_j / r_h)^\nu$, where z_j obeys a log-normal distribution with standard deviation σ_{shadow} , and r_j is the distance from the j th user to BS. In all cases, the curves remain constant at 1 at low SNRs, and undergo a sharp decrease until decreasing to 0 at high SNR. Moreover, there appears to be a significant property that the number of users has an profound impact on $E[I^{-1}]$; in particular, $E[I^{-1}]$ decreases with increasing N since interference I increases linearly with the number of users.

Lemma 5: For the function as $L(x)$, the following properties are satisfied:

$$\lim_{x \rightarrow \infty} \frac{L(x)}{\ln(x)} = -1, \quad \text{and} \quad \lim_{x \rightarrow 0} \frac{L(x)}{x} = -1. \quad (58)$$

Proof: The results can be obtained by using *L'Hôpital's Rule* directly. ■

Proposition 4: For a MU-MIMO system with an MRC receiver, the expressions of the two offsets of J_1 and J_2 are

given by

$$\Delta_1 = \log_2 \left(1 + M\beta_j \sum_{k=1, k \neq j}^N \frac{-L(\rho_{\text{mu}}\beta_k)}{T(j, k, \mathbf{B})} \right) - \log_2 \left(1 + \frac{\rho_{\text{mu}}(M-1)\beta_j}{1 + \rho_{\text{mu}}\varpi} \right) \quad (59)$$

and

$$\Delta_2 = \log_2 \left(1 + \frac{\rho_{\text{mu}}\varpi + 1}{\rho_{\text{mu}}(M-1)\beta_j} \right) - \log_2 \left(1 + \frac{1}{M\beta_j \sum_{k=1, k \neq j}^N \frac{-L(\rho_{\text{mu}}\beta_k)}{T(j, k, \mathbf{B})}} \right). \quad (60)$$

Proof: Substituting (56) and (57) into *Theorem 1* yields the desired result. ■

This result shows the relationship between the large-scale fading coefficient and tightness of bounds, which varies with the different users' positions. In order to gain more insights into the expressions in *Proposition 4*, we particularize our results for three cases.

- As $\rho_{\text{mu}} \rightarrow 0$, the two offsets reduce to

$$\Delta_2 = \log_2 \left(\frac{M}{M-1} \sum_{k=1, k \neq j}^N \frac{\beta_k}{T(j, k, \mathbf{B})} \right), \text{ and } \Delta_1 = 0 \quad (61)$$

The results are obtained by utilizing *Lemma 5* in (59) and (60). We see that J_1 can perfectly characterize the capacity in the low-SNR regime. Moreover, the tightness of J_2 depends on the LSF from each user, and when M grows large, Δ_2 converges to

$$\Delta_2 = \log_2 \left(\sum_{k=1, k \neq j}^N \frac{\beta_k}{T(j, k, \mathbf{B})} \right). \quad (62)$$

Fig. 6. Comparison of analytical expression and Monte-Carlo simulation of $E[I^{-1}]$ for a MU-MIMO MRC system with $M = 128$, $\sigma_{\text{shadow}} = 8\text{dB}$ and $\nu = 2.1$.

Fig. 7. Δ_1 and Δ_2 for a MU-MIMO MRC system with 10 users in the cell. The propagation channel parameters are $\sigma_{\text{shadow}} = 8\text{dB}$ and $\nu = 2.1$.

- As $\rho_{\text{mu}} \rightarrow \infty$, the two offsets approximately reduce to

$$\Delta_1 = \log_2 \left(\frac{\varpi + \varpi M\beta_j \sum_{k=1, k \neq j}^N \frac{\ln \beta_k}{T(j, k, \mathbf{B})}}{(M-1)\beta_j} \right) \quad (63)$$

and

$$\Delta_2 = \log_2 \left(1 + \frac{\varpi}{(M-1)\beta_j} \right) - \log_2 \left(1 + \frac{1}{M\beta_j \sum_{k=1, k \neq j}^N \frac{\ln \beta_k}{T(j, k, \mathbf{B})}} \right). \quad (64)$$

This result is obtained by using *Lemma 5* and [43, Lemma 2] in *Proposition 4*. The result shows the significant impact of LSF on the tightness of bounds in the high-SNR regime with a finite number of antennas..

- In the massive MIMO regime, i.e., as $M \rightarrow \infty$, the two offsets approximately converge to

$$\Delta_1 = \log_2 \left(\frac{1 + \rho_{\text{mu}}\varpi}{\rho_{\text{mu}}} \sum_{k=1, k \neq j}^N \frac{-L(\rho_{\text{mu}}\beta_k)}{T(j, k, \mathbf{B})} \right),$$

and

$$\Delta_2 = 0. \quad (65)$$

Note that J_2 is considerably tight for massive MIMO for any the arbitrary SNRs. Moreover, we see that as the SNR grows large, Δ_1 further converges to

$$\Delta_1 = \log_2 \left(\varpi \sum_{k=1, k \neq j}^N \frac{\ln \beta_k}{T(j, k, \mathbf{B})} \right). \quad (66)$$

We note that the LSF of each user has significant influence on the tightness of Jensen's bounds. Therefore, in contract of SU-MIMO scenarios, the two offsets do not converge to the same value in the high and low SNR regime for MU-MIMO MRC systems. However, from (65), we see that J_2 is rather tight for arbitrary SNRs in massive MIMO systems.

TABLE I
TIGHTNESS OF JENSEN'S BOUNDS FOR DIFFERENT SCENARIOS.

ρ [dB]	$m \times n$	Single user						Multiuser	
		i.i.d. Rayleigh		semi-correlated Rayleigh		i.i.d. Ricean		i.i.d. Rayleigh	
		Δ_1	Δ_2	Δ_1	Δ_2	Δ_1	Δ_2	Δ_1	Δ_2
-10	1×2	0.067	0.933	0.067	0.933	0.004	0.040		
	2×4	0.134	1.866	0.191	3.315	0.142	2.246	0.028	0.396
	4×8	0.269	3.732	0.460	10.21	0.401	7.615	0.007	0.185
	2×128	0.005	0.053	0.111	1.452	0.124	1.785	0.006	0.009
	4×128	0.021	0.214	0.378	6.531	0.354	5.844	0.004	0.004
15	1×2	0.956	0.044	0.956	0.044	0.045	0.006		
	2×4	1.913	0.087	3.303	0.202	2.268	0.119	1.024	0.704
	4×8	3.826	0.174	9.976	0.877	7.497	0.519	0.314	0.277
	2×128	0.057	0.002	1.501	0.063	1.823	0.086	0.812	0.024
	4×128	0.228	0.007	6.514	0.395	5.865	0.333	0.292	0.013

In Fig. 7, the offsets of J_1 based on (59) and J_2 based on (60) are depicted with 10 users in the cell. Here, we consider the same parameters as before. We first note that the absolute values of offsets of both bounds are rather small, which implies that the tightness of J_1 and J_2 is confirmed to be very satisfactory with single-antenna users. Moreover, we see that Δ_1 does not have a mirror behavior against Δ_2 in contrast with SU-MIMO systems, i.e., both bounds are tighter in the low-SNR regime. However, it still holds that J_1 performs better at low-SNRs while J_2 is tighter at high-SNRs. Similar to the SU cases, the tightness of bounds also improves with increasing the maximum number of antennas.

V. DISCUSSION

In this section, we further investigate the analytical offsets of Jensen's bounds for the SU-MIMO and MU-MIMO systems, and present a method of minimizing the offsets by selecting proper bounds in different SNR regimes. Note that all the results in this section are derived by prior expressions in Section IV.

In order to show the implications of our results, we analyze the most common scenarios of interest shown in Table I. We assume that the system works at -10dB and 15dB corresponding to low-SNR and high-SNR regimes respectively. For single user cases, we consider that the user is equipped with 1, 2 or 4 antennas while the base station (BS) has 2, 4, 8 or 128 antennas (i.e., $m = 1, 2, 4$ and $n = 2, 4, 8, 128$) to represent the finite-dimension MIMO systems and massive MIMO systems. Note that in i.i.d Ricean fading channels, the deterministic component $\mathbf{M}$ with singular values $\epsilon_{k=1,\dots,m}$ is randomly generated shown in Table II. For multiuser cases, we only consider there are 2 or 4 single-antenna users (i.e., $m = 2, 4$) in the system. The deterministic LSF β_k is derived based on the location of k th user who is randomly distributed in the area, as shown in Table II.

For SU-MIMO, from Table I, it can be confirmed that J_1 is more suitable for low-SNR scenarios while J_2 performs better at high-SNRs. The offset Δ_2 is rather small, numerically, less than 0.5bit/s/Hz in the most of cases when the system is working at 15dB regardless of the fading channels, especially for the cases of small number of terminal antennas. Note that

the offset can be further reduced as ρ grows based on our prior analysis in Section III. In other words, the numerical results indicate that the capacity can be approximately expressed by both the upper and lower bound in J_2 with the offset less than Δ_2 at 15dB (i.e., less than 0.087 bits/Hz for $m \times n = 2 \times 4$). On the contrary, J_1 behaves poorly in this regime while it becomes rather tight at -10dB. Therefore, we can infer that the capacity can be simply measured by selecting the right pairs of Jensen's bounds in different SNR regimes without calculating it, and show how tight the bounds are. We also note that Jensen's bounds perform even tighter with 128 antennas at the BS as expected, which validates that the Jensen's bounds become tighter for massive MIMO.

For MU-MIMO, recall that the tightness of Jensen's bounds is effected by the location of users. Without loss of generality, we randomly generate the users and model LSF via $\beta_k = z_k / (r_k / r_h)^\nu$ shown in Table II. We first note that the Jensen's bounds can be used to measure the capacity in the low-SNR regimes since both pairs of them behave really tight at -10dB. However, we are more interested in the behavior of Jensen's bounds in the high-SNR regime. The numerical results show that the Jensen's bound is not as tight as in the cases of low-SNRs, but still acceptable in representing capacity especially in the massive MIMO scenarios.

So far, we have confirmed that the capacity can be precisely predicted by proper Jensen's bound in both low-SNR and high-SNR regimes. We now investigate the worst performance of Jensen's bounds for SU-MIMO with respect to SNR. As shown in *Corollary 1*, the offset Δ_{ρ_t} at ρ_t can be regarded as the maximum offset since we can select different Jensen's bound

TABLE II
SYSTEM PARAMETERS.

$m \times n$	Single user i.i.d. Ricean	Multiuser i.i.d. Rayleigh
	$\{\epsilon_{k=1,\dots,m}\}$	$\{\beta_{k=1,\dots,m}\}$
1×2	{2}	
2×4	{0.334, 7.617}	
4×8	{1.463, 2.358, 3.519, 24.66}	{0.051, 0.150, 0.056, 0.139}
2×128	{30.44, 225.6}	{0.065, 0.218}
4×128	{24.12, 31.09, 40.99, 415.8}	{0.051, 0.150, 0.056, 0.139}

TABLE III
MAXIMUM OFFSET OF JENSEN'S BOUNDS FOR SU-MIMO WITH RESPECT TO SNR.

m	n	i.i.d. Rayleigh		semi-correlated Rayleigh		i.i.d. Ricean	
		ρ_t [dB]	Δ_{ρ_t}	ρ_t [dB]	Δ_{ρ_t}	ρ_t [dB]	Δ_{ρ_t}
1	2	0.71	0.500	1.51	0.500	0.06	0.022
	4	0.29	0.208	0.62	0.206	0.02	0.011
	8	0.13	0.096	0.29	0.096	0.01	0.005
	128	0.01	0.006	0.02	0.006		0
2	4	0.71	1.000	2.64	1.753	2.71	1.201
	8	0.29	0.415	1.76	1.168	1.44	1.017
	128	0.02	0.023	1.17	0.776	1.43	0.931
4	8	0.71	2.000	4.02	5.337	2.52	4.061
	128	0.03	0.092	2.58	3.428	2.24	2.021

to reduce the offset at other SNR values. Thus, we evaluate both Δ_{ρ_t} and ρ_t shown in Table III with the same simulation parameters shown in Table II.

For the cases of $m = 1$, we note two features: (a) ρ_t admits values which are quite smaller than practical values of interest; (b) the maximum offset Δ_{ρ_t} is rather small. The results indicate that Jensen's bounds are reliable in characterizing the capacity of SIMO with arbitrary fading channels. For the cases of $m = 2, 4$, we see that the offsets become larger, especially when the user is equipped with 4 antennas. Therefore, Jensen's bounds can hardly predict the capacity around ρ_t when multi-antennas are deployed at terminal side. However, it is important to note that ρ_t is still much less than the usual working SNR of the concurrent practical systems.

VI. CONCLUSION

This paper has investigated a problem in the context of MIMO systems that has been surprisingly overlooked in the relevant literature, namely the tightness of Jensen's capacity bounds. In order to perform this task, we have first derived some new results for finite-dimensional Wishart random matrix theory, by elaborating on central and non-central Wishart matrices with arbitrary dimensionality. We point out that our results for the unordered eigenvalue of different classes of Wishart matrices are closed-form expressions that involve standard functions. These results were subsequently used in the tightness analysis of Jensen's bounds. We have investigated the joint impact of minimum dimensionality, maximum dimensionality and SNR on the tightness of Jensen's bounds for two different cases: (a) SU-MIMO systems in different fading scenarios; (b) MU-MIMO systems with MRC receivers. Based on our analysis, we first note that it is improper to utilize Jensen's bounds to characterize the capacity around ρ_t for a multi-antenna user. Otherwise, choosing proper Jensen's bounds according to the operating SNR can precisely predict the performance of MIMO systems. We also have confirmed that the tightness of bounds improves by increasing the maximum number of antennas, while their tightness deteriorates by increasing the minimum number of antennas. This implies that the Jensen's bounds are perfectly suitable in characterizing the capacity of massive MIMO systems.

APPENDIX A PROOF OF LEMMA 1

We prove this lemma in two separate cases.

The $(\Lambda_m, \Lambda_n) = (\Theta_m, \mathbf{I}_n)$ Case: For this case, the pdf of the unordered eigenvalue is given by [23, Eq. (14)] as

$$f_\lambda(\lambda) = \frac{1}{m \prod_{i < j}^m (\theta_j - \theta_i)} \times \sum_{i=1}^m \sum_{j=1}^m \frac{\lambda^{n-m+j-1} e^{-\lambda/\theta_i} \theta_i^{m-n-1}}{\Gamma(n-m+j)} D_{i,j}, \quad (67)$$

where $D_{i,j}$ the (i, j) -cofactor of an $m \times m$ matrix whose (l, k) -th entry is

$$\{\mathbf{D}\}_{l,k} = \theta_l^{k-1}. \quad (68)$$

It is straightforward to see that the p th moment of λ is given as

$$E[\lambda^p] = \frac{1}{m \prod_{i < j}^m (\theta_j - \theta_i)} \times \sum_{i=1}^m \sum_{j=1}^m \frac{\Gamma(n-m+j+p)}{\Gamma(n-m+j)} \theta_i^{j+p-1} D_{i,j} = \frac{1}{m \prod_{i < j}^m (\theta_j - \theta_i)} \sum_{j=1}^m \frac{\Gamma(n-m+j+p)}{\Gamma(n-m+j)} \det(\tilde{\mathbf{D}}_j^p). \quad (69)$$

The $(s, \Lambda_m, \Lambda_n) = (n, \mathbf{I}_m, \Theta_n)$ Case: For this case, we start by employing a result from [23, Eq. (14)] to express the pdf of the unordered eigenvalue λ , conditioned on Θ_n , as follows:

$$f_\lambda(\lambda) = \frac{1}{m \prod_{i < j}^n (\theta_j - \theta_i)} \times \sum_{i=1}^n \sum_{j=n-m+1}^n \frac{\lambda^{m-n+j-1} e^{-\lambda/\theta_i} \theta_i^{n-m-1}}{\Gamma(m-n+j)} D_{i,j}. \quad (70)$$

The arbitrary order moment of λ is calculated in the same way as in (69).

By setting s equal to the number of antennas on the side with correlation and t equal to the number of antennas on the

other side, we obtain

$$E[\lambda^p] = \frac{1}{m \prod_{i < j}^s (\theta_j - \theta_i)} \sum_{j=s-m+1}^s \frac{\Gamma(t-s+j+p)}{\Gamma(t-s+j)} \det(\tilde{\mathbf{D}}_j^p). \quad (71)$$

For the $p > 0$ case, $\det(\tilde{\mathbf{D}}_j^p) = 0$ when $j \leq s-p$, while for the $p < 0$ case, $\det(\tilde{\mathbf{D}}_j^p) = 0$ when $j > -p$. By noting this property, we conclude the proof.

APPENDIX B PROOF OF LEMMA 2

In order to prove this corollary, we first analyze the properties of the generalized Vandermonde determinants.

Lemma 6: Let us define a generalized Vandermonde matrix according to

$$\mathbf{V}_{n-k} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ a_1 & a_2 & a_3 & \cdots & a_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1^{n-k-1} & a_2^{n-k-1} & a_3^{n-k-1} & \cdots & a_n^{n-k-1} \\ a_1^{n-k+1} & a_2^{n-k+1} & a_3^{n-k+1} & \cdots & a_n^{n-k+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1^n & a_2^n & a_3^n & \cdots & a_n^n \end{pmatrix}, \quad (72)$$

where $k = 0, 1, \dots, n$. The determinant of $\mathbf{V}_{n-k}$ is equal to

$$\det(\mathbf{V}_{n-k}) = \sum_{q > \dots > j > i}^n \underbrace{a_i a_j \cdots a_q}_k \prod_{i < j}^n (a_j - a_i). \quad (73)$$

Proof: We first extend (72) by adding one column y^j and one row a_i^{n-k} with $i, j = 0, 1, \dots, n$, as follows

$$\mathbf{V} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ a_1 & a_2 & a_3 & \cdots & a_n & y \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_1^{n-2} & a_2^{n-2} & a_3^{n-2} & \cdots & a_n^{n-2} & y^{n-2} \\ a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \cdots & a_n^{n-1} & y^{n-1} \\ a_1^n & a_2^n & a_3^n & \cdots & a_n^n & y^n \end{pmatrix}. \quad (74)$$

Recalling the property of Vandermonde determinants, we have

$$\begin{aligned} \det(\mathbf{V}) &= \prod_{i=1}^n (y - a_i) \prod_{i < j}^n (a_j - a_i) \\ &= \left(y^n + (-1)y^{n-1} \sum_{i=1}^n a_i + \cdots \right. \\ &\quad \left. + (-1)^k y^{n-k} \sum_{q > \dots > j > i}^n \underbrace{a_i a_j \cdots a_q}_k + \cdots \right. \\ &\quad \left. + (-1)^n \prod_{i=1}^n a_i \right) \prod_{i < j}^n (a_j - a_i). \end{aligned} \quad (75)$$

We also notice that $\det(\mathbf{V})$ can be calculated by expanding it along its last row as

$$\det(\mathbf{V}) = \sum_{i=0}^n (-1)^i y^{n-i} \det(\mathbf{V}_{n-i}). \quad (76)$$

Comparing (75) and (76) in terms of the factor y^k , we obtain the result. ■

By establishing *Lemma 6*, the expectation of λ can be simplified as

$$E[\lambda] = \frac{t}{m} \sum_{i=1}^s \theta_i. \quad (77)$$

To derive $E[\frac{1}{\lambda}]$, we generate a generalized Vandermonde matrix as

$$\mathbf{U} = \begin{pmatrix} \frac{1}{a_1} & \frac{1}{a_2} & \frac{1}{a_3} & \cdots & \frac{1}{a_n} \\ 1 & 1 & 1 & \cdots & 1 \\ a_1 & a_2 & a_3 & \cdots & a_n \\ a_1^2 & a_2^2 & a_3^2 & \cdots & a_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1^{n-2} & a_2^{n-2} & a_3^{n-2} & \cdots & a_n^{n-2} \end{pmatrix}, \quad (78)$$

whose determinant can be easily obtained as

$$\det(\mathbf{U}) = \frac{\prod_{i < j}^n (a_j - a_i)}{\prod_{i=1}^n a_i}. \quad (79)$$

Then, we calculate $\det(\mathbf{U}_{n-k})$ in the same way as in (73). Thus, we get

$$\det(\mathbf{U}_{n-k}) = \sum_{q > \dots > j > i}^n \underbrace{a_i a_j \cdots a_q}_{k+1} \frac{\prod_{i < j}^n (a_j - a_i)}{\prod_{i=1}^n a_i}. \quad (80)$$

We notice the fact that the expectation of $1/\lambda$ can also be calculated as

$$E\left[\frac{1}{\lambda}\right] = \frac{\Gamma(t-s) \det(\hat{\mathbf{D}}_1^{-1})}{m \Gamma(t-s+1) \prod_{i < j}^s (\theta_j - \theta_i)}, \quad (81)$$

which has a zero mass, only if $t-s > 0$. Thus, the first negative moment of λ does not exist when correlation is on the side with the maximum number of antennas, while

$$E\left[\frac{1}{\lambda}\right] = \frac{1}{m(n-m)} \sum_{i=1}^s \frac{1}{\theta_i} \quad (82)$$

when correlation is on the side with the minimum number of antennas.

APPENDIX C PROOF OF LEMMA 3

We first give some preliminary results.

Lemma 7: Let $\epsilon_1, \dots, \epsilon_m$ be the m eigenvalues of Ξ . The marginal density distribution of a single eigenvalue λ of $\Phi \sim W_m\left(n, \frac{1}{K+1} \mathbf{I}_m, \Xi\right)$ is given by (83) shown at the top of the next page, where $U_{i,j}$ is the (i,j) -cofactor of an $m \times m$ matrix whose (l,k) -th entry is

$$\{\mathbf{U}\}_{l,k} = (n-m)! \epsilon_l^{\frac{m-n}{2}} 2^{\frac{m-n-2k+2}{2}} e^{\epsilon_l} Q_{n-m+2k-1, n-m}(\sqrt{2\epsilon_l}, 0). \quad (84)$$

Proof: The joint pdf of the unordered eigenvalues of $\Psi \sim W_m(n, \mathbf{I}_m, \Xi)$ has been given in [19, Theorem 1]. After some basic transformations, we derive the joint pdf of the eigenvalues $\lambda_1, \dots, \lambda_m$ of Φ , given by (85). To evaluate the

$$f_\lambda(\lambda) = \frac{\text{etr}(\Xi)}{m((n-m)!)^m \prod_{k < l}^m (\epsilon_l - \epsilon_k)} \times \sum_{i=1}^m \sum_{j=1}^m \frac{e^{-(K+1)\lambda}}{\lambda} ((K+1)\lambda)^{n-m+j} {}_0F_1(n-m+1, (K+1)\epsilon_i \lambda) U_{i,j} \quad (83)$$

$$f_\lambda(\lambda_1, \dots, \lambda_m) = \frac{\text{etr}(\Xi)}{((n-m)!)^m \prod_{k < l}^m (\epsilon_l - \epsilon_k)} \times \det \{ {}_0F_1(n-m+1, (K+1)\epsilon_i \lambda) \} \det \{ ((K+1)\lambda_i)^{j-1} \} \prod_{i=1}^m \frac{e^{-(K+1)\lambda_i} \lambda_i^{n-m}}{(K+1)^{m-n-1}} \quad (85)$$

$$\begin{aligned} \{\mathbf{U}\}_{l,k} &= (K+1)^{n-m+k} \int_0^\infty e^{-(K+1)\lambda} \lambda^{n-m+k-1} {}_0F_1(n-m+1, (K+1)\epsilon_l \lambda) d\lambda \\ &= (n-m)! \epsilon_l^{\frac{m-n}{2}} 2^{\frac{m-n-2k+2}{2}} e^{\epsilon_l} Q_{n-m+2k-1, n-m}(\sqrt{2\epsilon_l}, 0). \end{aligned} \quad (86)$$

pdf of the unordered single eigenvalue of Ψ , we apply [42, Lemma 2] in (85). It is then straightforward to get (83). The (l, k) -th entry of $\mathbf{U}$ can be calculated in the same way as in [21, Eq. (60)] and is given in (86). ■

With Lemma 7, using the linearity of the determinants and after some basic manipulations, the expectation of λ can be derived as (87) shown at the top of the next page, where $\mathbf{U}_j$ is $m \times m$ matrix whose (l, k) th entry is

$$\{\mathbf{U}_j\}_{l,k} = \begin{cases} Q_{n-m+2k-1, n-m}(\sqrt{2\epsilon_l}, 0), & k \neq j, \\ Q_{n-m+2k+1, n-m}(\sqrt{2\epsilon_l}, 0), & k = j. \end{cases} \quad (88)$$

Noting that $\det(\mathbf{U}_j) = 0$ when $j \leq m-1$, we obtain the result. Note that the expectation of $1/\lambda$ can be derived in the same way as in (87).

APPENDIX D PROOF OF COROLLARY 6

Letting $m = 1$, it is straightforward to obtain

$$\kappa = 2^{\frac{1-n}{2}} \epsilon^{\frac{1-n}{2}}, \quad (89)$$

$$E[\lambda] = Q_{n-2, n-1}(\sqrt{2\epsilon}, 0), \quad (90)$$

and

$$E\left[\frac{1}{\lambda}\right] = Q_{n+2, n-1}(\sqrt{2\epsilon}, 0). \quad (91)$$

We notice that $I_q(\cdot)$ in (45) can be transformed into a first kind Bessel function [44, Eq. (8.406.3)] as

$$I_v(z) = i^{-v} J_v(iz), \quad (92)$$

where $i = \sqrt{-1}$. Substituting (92) in (45) and using [44, Eq. (6.631.1)], $Q_{p,q}(a, 0)$ can be derived as

$$\begin{aligned} Q_{p,q}(a, 0) &= i^{-q} \int_0^\infty x^p e^{-\frac{x^2+a^2}{2}} J_q(iax) dx \\ &= i^{-q} e^{-\frac{a^2}{2}} \int_0^\infty x^p e^{-\frac{x^2}{2}} J_q(iax) dx \\ &= \frac{a^q \Gamma\left(\frac{p+q+1}{2}\right) {}_1F_1\left(\frac{p+q+1}{2}; q+1; \frac{a^2}{2}\right)}{e^{\frac{a^2}{2}} 2^{\frac{q-p+1}{2}} \Gamma(q+1)}. \end{aligned} \quad (93)$$

We can now simplify the expectation of λ as

$$E[\lambda] = \frac{(\sqrt{2\epsilon})^{n-1}}{2(n-1)} {}_1F_1(n-1; n; \epsilon) e^{-\epsilon}. \quad (94)$$

Then, we use the relation [45, Eq. (07.20.03.0004.01)] and [45, Eq. (07.34.03.0456.01)] to get a closed-form representation of (94)

$$E[\lambda] = \frac{\left(-\sqrt{\frac{2}{\epsilon}}\right)^{n-1} \Gamma(n-1)}{2e^\epsilon} \left(1 - e^\epsilon \sum_{k=0}^{n-2} \frac{(-\epsilon)^k}{k!}\right). \quad (95)$$

The first negative moment of λ can also be evaluated through (93), given as

$$E\left[\frac{1}{\lambda}\right] = 2n \left(\sqrt{2\epsilon}\right)^{n-1} {}_1F_1(n+1; n; \epsilon) e^{-\epsilon}. \quad (96)$$

Using the relation [45, Eq. (07.20.03.0008.01)], the closed-form expression of (96) can be simplified as

$$E\left[\frac{1}{\lambda}\right] = 2 \left(\sqrt{2\epsilon}\right)^{n-1} (n + \epsilon). \quad (97)$$

Substituting (95) and (97) in Proposition 3, we obtain the desired result.

APPENDIX E PROOF OF LEMMA 4

We note that $E[I]$ can be evaluated straightforwardly since $\tilde{g}_k \sim CN(0, \beta_k)$. Let $\Omega = \sum_{k=1, k \neq j}^N |\tilde{g}_k|^2$ for convenience. To obtain $E[I^{-1}]$, it is necessary to establish the pdf of Ω . Note that $\tilde{g}_k \sim CN(0, \beta_k)$, and the sum of $|\tilde{g}_k|^2$ follows a generalized chi-squared distribution, given as [43, Theorem 4]

$$\begin{aligned} f_\Omega(x; N-1, \beta_1, \dots, \beta_{j-1}, \beta_{j+1}, \dots, \beta_N) \\ = \sum_{k=1, k \neq j}^N \frac{e^{-\frac{x}{\beta_k}}}{\beta_k \prod_{l=1, l \neq k, l \neq j}^N \left(1 - \frac{\beta_l}{\beta_k}\right)}. \end{aligned} \quad (99)$$

The first negative moment of NPI can be derived as (98), where (a) is derived by using [44, Eq. (3.352.4)].

$$\begin{aligned}
E[\lambda] &= \frac{e^{-\sum_i \epsilon_i}}{m \prod_{k < l}^m (\epsilon_l - \epsilon_k)} \sum_{j=1}^m \sum_{i=1}^m \frac{\epsilon_i^{\frac{m-n}{2}} 2^{\frac{2i-m-n-2}{2}}}{K+1} e^{\epsilon_i} Q_{n-m+2k+1, n-m}(\sqrt{2\epsilon_i}, 0) U_{i,j} \\
&= \frac{2^{\frac{m-n}{2}-1} (\prod_{i=1}^m \epsilon_i)^{(m-n)/2}}{m \prod_{k < l}^m (\epsilon_l - \epsilon_k) (K+1)} \sum_{j=1}^m \det \{\mathbf{U}_j\}
\end{aligned} \tag{87}$$

$$\begin{aligned}
E[\mathcal{I}^{-1}] &= \sum_{k=1, k \neq j}^N \frac{1}{\rho_{\text{mu}} \beta_k \prod_{l=1, l \neq k, l \neq j}^N \left(1 - \frac{\beta_l}{\beta_k}\right)} \int_0^\infty \frac{1}{\frac{1}{\rho_{\text{mu}}} + x} e^{-\frac{x}{\beta_k}} dx \\
&\stackrel{(a)}{=} \sum_{k=1, k \neq j}^N \frac{-e^{\frac{1}{\rho_{\text{mu}} \beta_k}} \text{Ei}\left(-\frac{1}{\rho_{\text{mu}} \beta_k}\right)}{\rho_{\text{mu}} \beta_k \prod_{l=1, l \neq k, l \neq j}^N \left(1 - \frac{\beta_l}{\beta_k}\right)}
\end{aligned} \tag{98}$$

REFERENCES

- [1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov. 1999.
- [2] T. Ratnarajah, "Topics in complex random matrices and information theory," Ph.D. dissertation, Department of Mathematics and Statistics, University of Ottawa, 2003.
- [3] M. Chiani, M. Z. Win, and H. Shin, "MIMO networks: The effects of interference," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 336–349, Jan. 2010.
- [4] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.
- [5] J. Wishart, "The generalised product moment distribution in samples from a normal multivariate population," *Biometrika*, vol. 20A, no. 1-2, pp. 32–52, Jul. 1928.
- [6] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *The Annals of Mathematical Statistics*, vol. 35, no. 2, pp. 475–501, Jun. 1964.
- [7] M. R. McKay, "Random matrix theory analysis of multiple antenna communication systems," Ph.D. dissertation, School of Electrical and Information Engineering, University of Sydney, 2007.
- [8] L. G. Ordóñez, D. P. Palomar, and J. R. Fonollosa, "Ordered eigenvalues of a general class of Hermitian random matrices with application to the performance analysis of MIMO systems," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 672–689, Feb. 2009.
- [9] T. Ratnarajah and R. Vaillancourt, "Quadratic forms on complex random matrices and multiple-antenna systems," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2976–2984, Aug. 2005.
- [10] M. Chiani, M. Z. Win, and A. Zanella, "On optimum combining of M-PSK signals with unequal-power interferers and noise," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 44–47, Jan. 2005.
- [11] A. Zanella, M. Chiani, and M. Z. Win, "On the marginal distribution of the eigenvalues of Wishart matrices," *IEEE Trans. Commun.*, vol. 57, no. 4, pp. 1050–1060, Apr. 2009.
- [12] C. G. Khatri, "Distribution of the largest or the smallest characteristic root under null hypothesis concerning complex multivariate normal populations," *The Annals of Mathematical Statistics*, vol. 35, no. 4, pp. 1807–1810, Dec. 1964.
- [13] M. Kang and M.-S. Alouini, "Largest eigenvalue of complex Wishart matrices and performance analysis of MIMO MRC systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 418–426, Apr. 2003.
- [14] C. G. Khatri and K. C. S. Pillai, "On the non-central distributions of two test criteria in multivariate analysis of variance," *The Annals of Mathematical Statistics*, vol. 39, no. 1, pp. 215–226, Mar. 1968.
- [15] N. R. Goodman, "Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction)," *The Annals of Mathematical Statistics*, vol. 34, no. 1, pp. 152–177, Mar. 1963.
- [16] A. Maaref and S. Aïssa, "Eigenvalue distributions of Wishart-type random matrices with application to the performance analysis of MIMO MRC systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2678–2689, Jul. 2007.
- [17] H. Gao and P. J. Smith, "A determinant representation for the distribution of quadratic forms in complex normal vectors," *Journal of Multivariate Analysis*, vol. 73, no. 2, pp. 155–165, May 2000.
- [18] M. R. McKay, A. J. Grant, and I. B. Collings, "Performance analysis of MIMO-MRC in double-correlated Rayleigh environments," *IEEE Trans. Commun.*, vol. 55, no. 3, pp. 497–507, Mar. 2007.
- [19] G. Alfano, A. Lozano, A. M. Tulino, and S. Verdú, "Mutual information and eigenvalue distribution of MIMO Ricean channels," in *Proc. IEEE ISITA*, Oct. 2004, pp. 1040–1045.
- [20] C. G. Khatri, "On certain distribution problems based on positive definite quadratic functions in normal vectors," *The Annals of Mathematical Statistics*, vol. 37, no. 2, pp. 468–479, Apr. 1966.
- [21] S. Jin, M. R. McKay, X. Gao, and I. B. Collings, "MIMO multichannel beamforming: SER and outage using new eigenvalue distributions of complex noncentral Wishart matrices," *IEEE Trans. Commun.*, vol. 56, no. 3, pp. 424–434, Mar. 2008.
- [22] L. Hanlen and A. Grant, "Capacity analysis of correlated MIMO channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 11, pp. 6773–6787, Nov. 2012.
- [23] S. Jin, M. R. McKay, C. Zhong, and K.-K. Wong, "Ergodic capacity analysis of amplify-and-forward MIMO dual-hop systems," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2204–2224, May 2010.
- [24] M. Matthaiou, C. Zhong, and T. Ratnarajah, "Novel generic bounds on the sum rate of MIMO ZF receivers," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4341–4353, Sept. 2011.
- [25] Q. T. Zhang, X. W. Cui and X. M. Li, "Very tight capacity bounds for MIMO-correlated Rayleigh-fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 681–688, Mar. 2005.
- [26] A. Zanella and M. Chiani, "Reduced complexity power allocation strategies for MIMO systems with singular value decomposition," *IEEE Trans. Veh. Technol.*, vol. 61, no. 9, pp. 4031–4041, Nov. 2012.
- [27] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "The multicell multiuser MIMO uplink with very large antenna arrays and a finite-dimensional channel," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2350–2361, June 2013.
- [28] M. Loève, "Probability theory." Springer Verlag, New York, 1977.
- [29] T. S. Ferguson, "Mathematical statistics: A decision theoretic approach," Academic Press, New York, 2014.
- [30] Ö. Oyman, R. U. Nabar, H. Bölcskei, and A. J. Paulraj, "Characterizing the statistical properties of mutual information in MIMO channels," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2784–2795, Nov. 2003.
- [31] 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, Sep. 2005.
- [32] Y. Chen, A. J. Goldsmith, and Y. C. Eldar, "Backing off from infinity: Performance bounds via concentration of spectral measure for random MIMO channels," *IEEE Trans. Inf. Theory*, vol. 61, no. 1, pp. 366–387, Jan. 2015.
- [33] A. Guionnet and O. Zeitouni, "Concentration of the spectral measure for large matrices," *Electronic Communications in Probability*, vol. 5, pp. 119–136, Jun. 2000.
- [34] A. Guntuboyina and H. Leeb, "Concentration of the spectral measure of large Wishart matrices with dependent entries," *Electronic Communications in Probability*, vol. 14, no. 334–342, p. 4, Jun. 2009.

- [35] A. Lozano, A. M. Tulino, and S. Verdú, "High-SNR power offset in multiantenna communication," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4134–4151, Dec. 2005.
- [36] —, "Multiple-antenna capacity in the low-power regime," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2527–2544, Oct. 2003.
- [37] A. H. Stone and J. W. Tukey, "Generalized sandwich theorems," *Duke Math. J.*, vol. 9, no. 2, pp. 356–359, Jun. 1942.
- [38] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical functions with Formulas, Graphs, and Mathematical Tables*. Courier Corporation, Jun. 1964, no. 55.
- [39] R. Askey and A. O. Daalhuis, *NIST handbook of Mathematical functions*. Cambridge University Press, 2010.
- [40] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Uniform power allocation in MIMO channels: A game-theoretic approach," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1707–1727, July 2003.
- [41] G. Alfano, A. M. Tulino, A. Lozano, and S. Verdú, "Capacity of MIMO channels with one-sided correlation," in *Proc. IEEE ISSTA*, Aug. 2004, pp. 515–519.
- [42] H. Shin, M. Z. Win, H. L. Jae, and M. Chiani, "On the capacity of doubly correlated MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 8, pp. 2253–2265, Aug. 2006.
- [43] D. Hammarwall, M. Bengtsson, and B. Ottersten, "Acquiring partial CSI for spatially selective transmission by instantaneous channel norm feedback," *IEEE Trans. Sig. Proc.*, vol. 56, no. 3, pp. 1188–1204, Mar. 2008.
- [44] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products, 7th ed.* New York: Academic Press, 2007.
- [45] "The Wolfram functions site," Available: <http://functions.wolfram.com>, [Online].

Jide Yuan (S'16) received the B.S. degree in communication engineering from the Ocean University of China, Qingdao, China, in 2011, and the M.S. degree from the Harbin Engineering University, Harbin, China, in 2014. He is currently working towards the Ph.D. degree in Communication and Information Systems, Southeast University, China, under the supervision of Prof. Shi Jin. His current research interests include random matrix theory, design of cellular networks, and ultra-dense DAS.

and Signal Processing, Munich University of Technology (TUM), Germany working as a Postdoctoral Research Associate. He is currently a Senior Lecturer at Queen's University Belfast, U.K. after holding an Assistant Professor position at Chalmers University of Technology, Sweden. His research interests span signal processing for wireless communications, massive MIMO, hardware-constrained communications, and performance analysis of fading channels.

Dr. Matthaiou was the recipient of the 2011 IEEE ComSoc Best Young Researcher Award for the Europe, Middle East and Africa Region and a co-recipient of the 2006 IEEE Communications Chapter Project Prize for the best M.Sc. dissertation in the area of communications. He was co-recipient of the Best Paper Award at the 2014 IEEE International Conference on Communications (ICC) and was an Exemplary Reviewer for IEEE COMMUNICATIONS LETTERS for 2010. In 2014, he received the Research Fund for International Young Scientists from the National Natural Science Foundation of China. He currently serves as Senior Editor for IEEE COMMUNICATIONS LETTERS, an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, and was the Lead Guest Editor of the special issue on "Large-scale multiple antenna wireless systems" of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. He is the chair of the Wireless Communications Symposium (WCS) at IEEE GLOBECOM 2016. He is an associate member of the IEEE Signal Processing Society SPCOM and SAM technical committees.

Shi Jin (S'06–M'07) received the B.S. degree in communications engineering from Guilin University of Electronic Technology, Guilin, China, in 1996, the M.S. degree from Nanjing University of Posts and Telecommunications, Nanjing, China, in 2003, and the Ph.D. degree in communications and information systems from the Southeast University, Nanjing, in 2007. From June 2007 to October 2009, he was a Research Fellow with the Adastral Park Research Campus, University College London, London, U.K. He is currently with the faculty of the National Mobile Communications Research Laboratory, Southeast University. His research interests include space time wireless communications, random matrix theory, and information theory. He serves as an Associate Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and IEEE COMMUNICATIONS LETTERS, and IET COMMUNICATIONS. Dr. Jin and his co-authors have been awarded the 2011 IEEE Communications Society Stephen O. Rice Prize Paper Award in the field of communication theory and a 2010 Young Author Best Paper Award by the IEEE Signal Processing Society.

Feifei Gao (M'09–SM'14) received the B.Eng. degree from Xi'an Jiaotong University, Xi'an, China in 2002, the M.Sc. degree from McMaster University, Hamilton, ON, Canada in 2004, and the Ph.D. degree from National University of Singapore, Singapore in 2007. He was a Research Fellow with the Institute for Infocomm Research (I2R), A*STAR, Singapore in 2008 and an Assistant Professor with the School of Engineering and Science, Jacobs University, Bremen, Germany from 2009 to 2010. In 2011, he joined the Department of

Automation, Tsinghua University, Beijing, China, where he is currently an Associate Professor.

Prof. Gao's research areas include communication theory, signal processing for communications, array signal processing, and convex optimizations, with particular interests in MIMO techniques, multi-carrier communications, cooperative communication, and cognitive radio networks. He has authored/coauthored more than 80 refereed IEEE journal papers and more than 120 IEEE conference proceeding papers, which have been cited more than 4000 times from Google Scholar.

Prof. Gao has served as an Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE COMMUNICATIONS LETTERS, IEEE SIGNAL PROCESSING LETTERS, IEEE WIRELESS COMMUNICATIONS LETTERS, INTERNATIONAL JOURNAL ON ANTENNAS AND PROPAGATIONS, and CHINA COMMUNICATIONS. He has also served as the symposium co-chair for 2015 IEEE Conference on Communications (ICC), 2014 IEEE Global Communications Conference (GLOBECOM), 2014 IEEE Vehicular Technology Conference Fall (VTC), as well as Technical Committee Members for many other IEEE conferences.