

Outage Capacity of Rayleigh Product Channels: a Free Probability Approach

Zhong Zheng, *Member, IEEE*, Lu Wei, *Member, IEEE*, Roland Speicher,
Ralf Müller, *Senior Member, IEEE*, Jyri Hämäläinen, *Member, IEEE*, and Jukka Corander

Abstract

The Rayleigh product channel model is useful in capturing the performance degradation due to rank deficiency of MIMO channels. In this paper, such a performance degradation is investigated via the channel outage probability assuming slowly varying channel with delay-constrained decoding. Using techniques of free probability theory, the asymptotic variance of channel capacity is derived when the dimensions of the channel matrices approach infinity. In this asymptotic regime, the channel capacity is rigorously proven to be Gaussian distributed. Using the obtained results, a fundamental tradeoff between multiplexing gain and diversity gain of Rayleigh product channels can be characterized by closed-form expression at any finite signal-to-noise ratio. Numerical results are provided to compare the relative outage performance between Rayleigh product channels and conventional Rayleigh MIMO channels.

Index Terms

Central limit theorem; finite-SNR diversity-multiplexing tradeoff; free probability theory; MIMO; outage capacity; Rayleigh product channels.

I. INTRODUCTION

Multi-Input Multi-Output (MIMO) wireless communications have received considerable attention since it is seen as the most promising way to increase link level capacity. Extensive works have focused on the

Z. Zheng and J. Hämäläinen are with the Department of Communications and Networking, Aalto University, Finland (e-mails: {zhong.zheng, jyri.hamalainen}@aalto.fi). L. Wei and J. Corander are with the Department of Mathematics and Statistics, University of Helsinki, Finland (e-mails: {lu.wei, jukka.corander}@helsinki.fi). R. Speicher is with the Faculty of Mathematics, Saarland University, Germany (e-mail: speicher@math.uni-sb.de). R. Müller is with the Institute for Digital Communications, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany (e-mail: mueller@int.de). This work was presented in part at 2014 International Zurich Seminar on Communications.

performance of MIMO channels assuming a rich scattering environment. Therein, the presumed models include full-rank independent Rayleigh or Rician MIMO channels. However, in certain environments the propagation may be subject to structural limits of fading channels caused by either insufficient scattering [1], [2] or the so-called keyhole effect [3]. These channels exhibit rank deficiency compared to the independent Rayleigh and Rician models. The MIMO model that captures these effects is referred to as the double-scattering channel [1]. It is characterized by a matrix product involving three deterministic matrices (i.e., transmit, receiver, and scatterer correlation matrices), and two statistically independent complex Gaussian matrices. In a typical office environment, empirical measurements have been used to demonstrate the validity of the double-scattering channel model [4].

There exist a number of studies concerning the information-theoretic quantities of the double-scattering channels. Shin *et al.* derived an upper bound for the ergodic capacity [5, Th. III.3] and an exact expression for a single keyhole channel [5, Th. III.4]. The diversity-multiplexing tradeoff of the double-scattering channel was obtained in [6]. The authors in [7] investigated the asymptotic Rayleigh-limit when one of the matrix dimensions approaches infinity. In this case, the double-scattering model reduces to an equivalent Rayleigh MIMO channel. Furthermore, if all matrix dimensions are large, the ergodic capacity has been obtained in [2] via numerical integration. Recently, an asymptotic expression for ergodic capacity of the double-scattering channels was derived in [8]. Moreover, authors in [9]–[11] derived the ergodic mutual information for finite dimensional channel matrices. However, all the above results are valid for ergodic channels, where each codeword has infinite length. For many practical communication systems such as WLANs [12], the channels, albeit random, are slowly varying and the encoding/decoding process is subject to a delay constraint with moderate target packet error rates around 10^{-2} – 10^{-1} . The fading channel seen by each codeword are therefore non-ergodic. In this case, the ergodic capacity has no physical significance, whereas the outage capacity is a more relevant performance metric [13]. In literature, the outage capacity has been studied for conventional Rayleigh MIMO channels [14]–[20] as well as for Rician MIMO channels [21]–[23] via various random matrix techniques.

To the best of our knowledge, the outage capacity for the double-scattering channel has not been addressed in the most general form¹. It turns out to be a difficult random matrix theory problem. To gain insights into the outage behavior of the double-scattering channel, we consider a simplified channel

¹Note that authors in [18] derived the outage probability of a similar MIMO model with random steering matrices at antenna arrays. However, these steering matrices are slowly varying compared to multi-path fading and considered as deterministic. Thus, the tools in [18] are not applicable here.

model involving a product of two statistically independent complex Gaussian matrices, also known as the Rayleigh product channel. This channel model corresponds to the scenario where the antenna elements as well as the scattering objects are sufficiently separated and there is no spatial correlation at antenna arrays or between scatterers. To characterize the capacity fluctuations, we use the free probability theory for large dimensional random matrices [24]–[26]. By utilizing the second order Cauchy transform and R -transform machinery, we derive a compact expression for the asymptotic variance of the capacity of the Rayleigh product channel. We further show that the channel capacity distribution is asymptotically Gaussian by proving a Central Limit Theorem (CLT) for the Linear Spectral Statistics (LSS) of the Rayleigh product ensemble. This result generalizes the CLT for correlated Wishart random matrices [27] and the CLT for Rayleigh product ensembles from polynomial LSS to generic analytic functions [28]. The capacity distribution is then utilized to study the corresponding finite Signal-to-Noise-Ratio (SNR) Diversity-Multiplexing Tradeoff (DMT). The derived results in this paper are formally valid when the dimensions of the channel matrices grow to infinity. However, numerical simulations show that they serve as good approximations when the numbers of antennas and scatterers are comparable to practical systems.

The rest of the paper is organized as follows. In Section II, we give the channel model, the signal model as well as the MIMO capacity formulation. In Section III, we study the second order eigenvalue fluctuations and the asymptotic capacity variance. The second order Cauchy transform of Rayleigh product ensembles is derived in Section IV. The CLT of the capacity of Rayleigh product channels is proved in Section V. Based on this result, the approximations for outage probability and the finite-SNR DMT are calculated. In Section VI, we conclude the main findings of the paper. Proofs of the technical results are provided in the Appendices.

Notations. Throughout the paper, vectors are represented by lower-case bold-face letters, and matrices are represented by upper-case bold-face letters. The complex vector field with length n is denoted as \mathbb{C}^n . We use $\mathcal{CN}(0, \mathbf{A})$ to denote the zero-mean complex Gaussian vector with covariance matrix \mathbf{A} and \mathbf{I}_n is an $n \times n$ identity matrix. The superscript $(\cdot)^\dagger$ denotes the matrix conjugate-transpose operation and $(\cdot)^T$ is matrix transpose. We denote $\overline{(\cdot)}$ as the complex conjugate operator. Denote $\text{Tr}(\mathbf{A})$ as the trace of $n \times n$ matrix \mathbf{A} and $\text{tr}(\mathbf{A})$ as the normalized trace $\text{Tr}(\mathbf{A})/n$. The notation $\mathbb{E}[\cdot]$ denotes the expectation, and $\det(\cdot)$ denotes the matrix determinant.

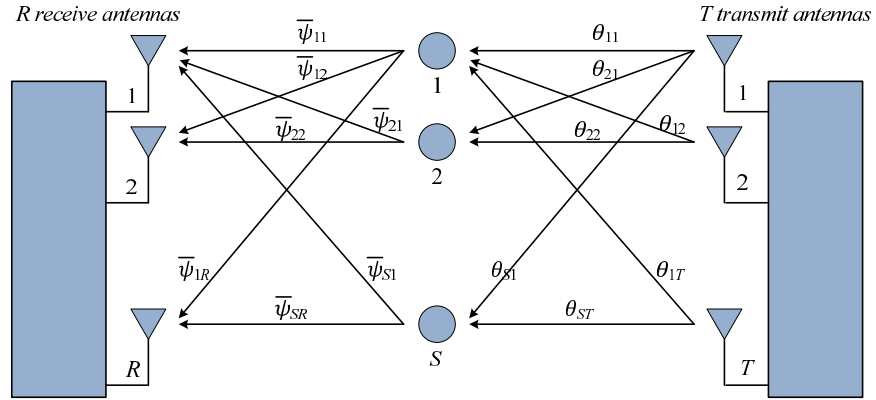


Fig. 1. MIMO communications over the Rayleigh product channel with T transmit antennas, R receive antennas, and S scatterers.

II. RAYLEIGH PRODUCT MIMO CHANNELS

A. Channel Model

Consider a discrete-time, baseband MIMO system with T transmit and R receive antennas. The channel is assumed to follow the Rayleigh product fading with S scattering objects, as shown in Fig. 1. The channels between the s -th scatterer and transmit antennas are denoted by vector $\boldsymbol{\theta}_s = [\theta_{s1}, \dots, \theta_{sT}]$, and the channels between receive antennas and the s -th scatterer are denoted by vector $\boldsymbol{\psi}_s = [\psi_{s1}, \dots, \psi_{sR}]$. The end-to-end equivalent channel matrix \mathbf{H} is given by

$$\mathbf{H} = \frac{1}{\sqrt{RS}} \sum_{s=1}^S \boldsymbol{\psi}_s^\dagger \boldsymbol{\theta}_s = \frac{1}{\sqrt{RS}} \boldsymbol{\Psi}^\dagger \boldsymbol{\Theta}, \quad (1)$$

where $\boldsymbol{\Theta} = [\boldsymbol{\theta}_1^\dagger, \dots, \boldsymbol{\theta}_S^\dagger]^\dagger$ and $\boldsymbol{\Psi} = [\boldsymbol{\psi}_1^\dagger, \dots, \boldsymbol{\psi}_S^\dagger]^\dagger$. We assume $\boldsymbol{\theta}_s \sim \mathcal{CN}(0, \mathbf{I}_T)$ and $\boldsymbol{\psi}_s \sim \mathcal{CN}(0, \mathbf{I}_R)$, where θ_i and ψ_j , $1 \leq i, j \leq S$, are statistically independent. The channel \mathbf{H} is thus modeled as a product of two independent complex Gaussian random matrices. In line with [2], [8]–[11], the channel \mathbf{H} is normalized by the constant $1/\sqrt{RS}$ so that the total energy of the channel is equal to an AWGN channel with an array gain $\mathbb{E}[\text{Tr}(\mathbf{H}\mathbf{H}^\dagger)] = \sum_{i,j} \mathbb{E}[|H_{ij}|^2] = T$.

The presence of independent Gaussian matrices $\boldsymbol{\Theta}$ and $\boldsymbol{\Psi}$ in (1) requires two independent and richly-scattered environments, where the scattering happens between the S scatterers/keyholes and transmit and receive arrays, respectively. This requires the existence of a large number of independently reflected and scattered paths around the antenna arrays [29]. The two environments are connected only via the S scatterers/keyholes. By controlling the number S , the Rayleigh product channel (1) embraces a general

family of MIMO fading channel, spanning from the degenerate keyhole channel $S = 1$ [3] to the full-rank Rayleigh MIMO channel $S \rightarrow \infty$ with fixed R and T [7]. Müller and Hofstetter [4] have shown that the number of significant scatterers is around ten in a typical office building with an 8×8 antenna configuration. Measurement results in [30] indicate that the effective rank of a 6×6 keyhole channel depends on the sizes of scatterer/keyhole at different transmission frequencies. In general, the number of separable scattering objects depends on the number of antenna elements since a larger array increases the spatial resolution [2]. Note that the model (1) also describes the MIMO relay channels when assuming noiseless relays [31].

B. Signal Model and Channel Capacity

The channel output vector $\mathbf{y} \in \mathbb{C}^R$, at a given time instance, equals

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (2)$$

where $\mathbf{x} \in \mathbb{C}^T$ is the transmit vector that follows the complex Gaussian distribution $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$ with $\mathbf{\Sigma} = \mathbb{E}[\mathbf{x}\mathbf{x}^\dagger]$. The additive noise $\mathbf{n} \in \mathbb{C}^R$ is modeled as an *i.i.d.* complex Gaussian vector $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_R)$. In this paper, we have adopted the following assumptions:

- A1) The Channel State Information (CSI) is perfectly known at the receiver but not at the transmitter.
- A2) The channel \mathbf{H} is frequency flat and quasi-static. It remains constant for certain symbol durations and takes independently a new value for each coherence time.
- A3) Delay-constrained encoding/decoding. The encoded transmit message has a finite block length and spread in time over no more than a maximum allowable decoding delay. We assume the length of a coding block is equal to one independently faded interval.

Under A1, Telatar [32] has shown that the channel capacity is achieved when the transmitted symbols are independent across antennas and the power is equally allocated, i.e. $\mathbf{\Sigma} = \gamma \mathbf{I}_T$ and γ denotes the Signal to Noise Ratio (SNR) per received antenna. The instantaneous capacity of the MIMO channel (2) in nats/sec/Hz is given by

$$\mathcal{I} = \log \det \left(\mathbf{I}_R + \gamma \mathbf{H}\mathbf{H}^\dagger \right) = \sum_{i=1}^R \log(1 + \gamma \lambda_i),$$

where λ_i , $i = 1, \dots, R$, refer to the eigenvalues of $\mathbf{Q} = \mathbf{H}\mathbf{H}^\dagger$. For the Hermitian matrix \mathbf{Q} , we find it convenient to introduce the Empirical Spectral Distribution (ESD) defined as

$$\widetilde{F}_{\mathbf{Q}}(\lambda) = \frac{1}{R} \sum_{i=1}^R \mathbf{1}(\lambda_i \leq \lambda),$$

where $\mathbf{1}(\cdot)$ denotes the indicator function. By letting $\varphi(x) = \log(1 + \gamma x)$, the channel capacity \mathcal{I} can be rewritten in terms of $\widetilde{F}_{\mathbf{Q}}(\lambda)$ as

$$\mathcal{I} = R \int \varphi(\lambda) d\widetilde{F}_{\mathbf{Q}}(\lambda). \quad (3)$$

As the channel matrix \mathbf{H} is random, the instantaneous capacity (3) is also a random variable. Without CSI at the transmitter, there is a non-zero probability, independent of the code length, that the channel capacity (3) falls below any positive rate. Due to the assumptions A2 and A3, the error probability corresponding to this rate cannot be decreased exponentially with the code length [13]. In this case, no reliable transmission is possible and the performance cannot be evaluated using the ergodic capacity. Instead, the fundamental performance limit of such a system is best explained with the capacity versus outage tradeoff, characterized by the Cumulative Distribution Function (CDF) of \mathcal{I} . Given a fixed rate r , the outage probability is defined as the probability that capacity \mathcal{I} is less than r , i.e.

$$P_{\text{out}}(r) = \Pr\{\mathcal{I} \leq r\} = F_{\mathcal{I}}(r), \quad (4)$$

where $F_{\mathcal{I}}(\cdot)$ denotes the CDF of \mathcal{I} . When the CDF $F_{\mathcal{I}}(\cdot)$ is monotonically increasing, the outage capacity for a given probability P_{out} is obtained as

$$\mathcal{I}_{\text{out}} = F_{\mathcal{I}}^{-1}(P_{\text{out}}).$$

The outage probability (4) is achievable [33] in the sense that for any $\epsilon > 0$, there exists a code of sufficiently large block length for which the packet error rate is upper-bounded by $P_{\text{out}}(r) + \epsilon$. Thus, outage capacity provides useful insights on the operational performance of a delay-constrained coded system. Outage probability is also a meaningful metric to characterize the performance of some contemporary communication systems [34], where timely CSI is available at the transmitters. From this viewpoint, the complementary outage probability $1 - P_{\text{out}}(r)$ can be interpreted as the percentage of time that a transmission takes place at given rate r under perfect link adaptation.

III. STATISTICS OF CHANNEL CAPACITY

In this section, we first review the convergence of empirical eigenvalue distribution of the Hermitian matrix $\mathbf{Q} = \mathbf{H}\mathbf{H}^\dagger$ when matrix dimensions grow to infinity. The capacity per receive antenna is shown to converge to a deterministic value and expressed by a known result in [8]. Then, we study the global fluctuation of eigenvalues around the limiting distribution by deriving a closed-form expression for the second order Cauchy transform of \mathbf{Q} . This result is utilized to obtain the asymptotic variance of the channel capacity.

A. First Order Cauchy Transform and Asymptotic Capacity

For an $N \times N$ Hermitian random matrix \mathbf{A} , we assume that its ESD $\widetilde{F}_{\mathbf{A}}(\cdot)$ converges to a non-random limiting distribution $F_{\mathbf{A}}(\cdot)$ as $N \rightarrow \infty$. Such a convergence is alternatively established in [35] via the convergence of resolvent $\widetilde{\mathcal{G}}_{\mathbf{A}}(z)$ to the first order Cauchy transform² $\mathcal{G}_{\mathbf{A}}(z)$, defined as

$$\begin{aligned}\widetilde{\mathcal{G}}_{\mathbf{A}}(z) &= \text{tr}(\mathbf{I}_N z - \mathbf{A})^{-1} = \int \frac{1}{z-t} d\widetilde{F}_{\mathbf{A}}(t), \\ \mathcal{G}_{\mathbf{A}}(z) &= \int_{\mathcal{S}_{\mathbf{A}}} \frac{1}{z-t} dF_{\mathbf{A}}(t).\end{aligned}\quad (5)$$

Here, $z \in \mathbb{C}^+ = \{z : \text{Im}(z) > 0\}$ and $\mathcal{S}_{\mathbf{A}}$ denotes the support of $F_{\mathbf{A}}(\cdot)$. Due to this limiting behavior of eigenvalues, the normalized linear spectral statistics, such as normalized capacity \mathcal{I}/R , converges to a non-random limit as $N \rightarrow \infty$ for a wide class of matrix ensembles [2], [14], [23].

In the following, the limit $\lim_{R \rightarrow \infty}$ denotes the asymptotic regime,

$$T, S, R \rightarrow \infty, \quad \text{with } \rho = \frac{S}{R} \quad \text{and} \quad \zeta = \frac{T}{S} \quad \text{fixed.} \quad (6)$$

In the asymptotic regime (6), Silverstein [36] shows that the ESD $\widetilde{F}_{\mathbf{Q}}(\cdot)$ converges almost surely to a non-random CDF $F_{\mathbf{Q}}(\cdot)$ and its Cauchy transform $\mathcal{G}_{\mathbf{Q}}$ is the solution to

$$z = \frac{1}{\mathcal{G}_{\mathbf{Q}}} + \rho \int_{\mathcal{S}_{\mathbf{P}}} \frac{\lambda dF_{\mathbf{P}}(\lambda)}{1 - \lambda \mathcal{G}_{\mathbf{Q}}}, \quad (7)$$

where $\mathbf{P} = \Theta\Theta^\dagger/S$ and $F_{\mathbf{P}}(\cdot)$ is the well-known Marčenko-Pastur distribution [37]. The integration range in (7) is $[(1 - \sqrt{\zeta})^2, (1 + \sqrt{\zeta})^2]$. Using multiplicative free convolution, Müller has shown in [2] that $\widetilde{\mathcal{G}}_{\mathbf{Q}}(z) \rightarrow \mathcal{G}_{\mathbf{Q}}(z)$ in the asymptotic regime (6) and $\mathcal{G}_{\mathbf{Q}}$ satisfies the cubic equation

$$z^2 \mathcal{G}_{\mathbf{Q}}^3(z) + (\rho\zeta + \rho - 2)z \mathcal{G}_{\mathbf{Q}}^2(z) + ((\rho\zeta - 1)(\rho - 1) - \rho z) \mathcal{G}_{\mathbf{Q}}(z) + \rho = 0. \quad (8)$$

If $|z| \rightarrow \infty$, $\mathcal{G}_{\mathbf{Q}}(\cdot)$ admits the formal power series expansion

$$\mathcal{G}_{\mathbf{Q}}(z) = \sum_{n=0}^{\infty} \alpha_n z^{-n-1}, \quad (9)$$

where $\alpha_0 = 1$ and α_n is the n -th free moment of \mathbf{Q} , defined as

$$\alpha_n = \lim_{R \rightarrow \infty} \mathbb{E}[\text{tr}(\mathbf{Q}^n)].$$

A concept closely related to the Cauchy transform $\mathcal{G}_{\mathbf{Q}}(z)$ is the R -transform $\mathcal{R}(z)$, defined as a functional [24]

$$\mathcal{G}_{\mathbf{Q}}\left(\mathcal{R}(z) + \frac{1}{z}\right) = z. \quad (10)$$

²In what follows, we refer to the first order Cauchy transform simply as Cauchy transform unless otherwise stated.

If $|z| \rightarrow \infty$, $\mathcal{R}(z)$ has the formal power series representation

$$\mathcal{R}(z) = \sum_{n=1}^{\infty} \kappa_n z^{n-1}, \quad (11)$$

where κ_n is the n -th free cumulant of \mathbf{Q} . As will be shown in Section IV, the free cumulant sequence $\{\kappa_n\}_{n \geq 1}$ and its generating function $\mathcal{R}(z)$ serve as the key analytical tools in the proof of Proposition 2. Note that when the matrix dimensions are finite, the computation of α_n involves non-trivial summations over all partitions of integer n [38, Eq. (27)]. This complicated expression makes it challenging to obtain an explicit expression for the free cumulant κ_n . As the matrix dimensions approach infinity, the calculation of κ_n is much simplified, involving only the so-called non-crossing permutations over integers, see Lemma 1 and Appendix C for a detailed discussion.

Using random matrix theory techniques, authors in [8] prove that the capacity per receive antenna of the Rayleigh product channel converges to an asymptotic limit such that $\lim_{R \rightarrow \infty} (\mathcal{I} - \mu_{\mathcal{I}})/R = 0$. The asymptotic capacity $\mu_{\mathcal{I}}/R$ is given by an explicit closed-form expression, which is summarized in the following proposition.

Proposition 1. (*Asymptotic capacity [8]*) *When $R = T$, the asymptotic capacity per receive antenna $\mu_{\mathcal{I}}/R$ in the regime (6) is given by*

$$\frac{\mu_{\mathcal{I}}}{R} = \log \left(\frac{1}{g} + \frac{\gamma}{\rho}(g + \rho - 1) \right) - \rho \log \left(1 + \frac{g-1}{\rho} \right) - 2(1-g),$$

where g is the unique solution to

$$g^3 - (1-\rho)g^2 + \frac{\rho}{\gamma}(g-1) = 0$$

such that $(1-g)/(g(g+\rho-1)) \geq 0$.

Although the asymptotic capacity $\mu_{\mathcal{I}}$ grows to infinity in the asymptotic regime (6), it serves as a tight approximation to the mean capacity $\mathbb{E}[\mathcal{I}]$ with finite matrix dimensions as shown in [8]. In the following, we will also use $\mu_{\mathcal{I}}$ as the approximated $\mathbb{E}[\mathcal{I}]$ whenever it is clear from the context.

B. Second Order Cauchy Transform and Asymptotic Variance

As $\widetilde{F}_{\mathbf{Q}}(\cdot) \rightarrow F_{\mathbf{Q}}(\cdot)$ in the asymptotic regime (6), the asymptotic capacity per-received antenna $\mu_{\mathcal{I}}/R$ can be formulated by replacing $\widetilde{F}_{\mathbf{Q}}(\lambda)$ in (3) with $F_{\mathbf{Q}}(\lambda)$. Using an integral identity³ [27, Eq. (1.14)],

³The definition of Stieltjes transform in [27] is different from the Cauchy transform by a minus sign.

we utilize an amenable form of asymptotic capacity, which is useful in the following discussion, namely

$$\frac{\mu_{\mathcal{I}}}{R} = \int_{\mathcal{S}_{\mathbf{Q}}} \varphi(\lambda) dF_{\mathbf{Q}}(\lambda) = \frac{1}{2\pi i} \oint_{\mathcal{C}} \varphi(z) \mathcal{G}_{\mathbf{Q}}(z) dz. \quad (12)$$

The complex integral on the right hand side of (12) is over any positively oriented closed contour \mathcal{C} enclosing the support $\mathcal{S}_{\mathbf{Q}}$ and on which $\varphi(\cdot)$ is analytic. For the instantaneous channel capacity (3), there exists a similar integral expression as in (12) with Cauchy transform $\mathcal{G}_{\mathbf{Q}}(z)$ replaced with the resolvent $\widetilde{\mathcal{G}}_{\mathbf{Q}}(z)$. To see this, let the contour \mathcal{C} be selected according to (12) and apply Cauchy's integral formula on $\varphi(\lambda)$, it follows that the instantaneous capacity (3) becomes

$$\frac{\mathcal{I}}{R} = \int \varphi(\lambda) d\widetilde{F}_{\mathbf{Q}}(\lambda) = \frac{1}{2\pi i} \int \oint_{\mathcal{C}} \frac{\varphi(x)}{x-\lambda} dx d\widetilde{F}_{\mathbf{Q}}(\lambda).$$

Exchange the integrations and recall the definition of resolvent (5), we obtain

$$\frac{\mathcal{I}}{R} = \frac{1}{2\pi i} \oint_{\mathcal{C}} \varphi(x) \widetilde{\mathcal{G}}_{\mathbf{Q}}(x) dx. \quad (13)$$

Let us now consider the variance of capacity \mathcal{I} , defined as $\sigma_{\mathcal{I}}^2 = \mathbb{E}[(\mathcal{I} - \mathbb{E}[\mathcal{I}])^2]$. Replacing \mathcal{I} with (13) and $\mathbb{E}[\mathcal{I}]$ with (12), the variance $\sigma_{\mathcal{I}}^2$ can be rewritten as

$$\sigma_{\mathcal{I}}^2 = -\frac{1}{4\pi^2} \mathbb{E} \left[\oint_{\mathcal{C}_x} \varphi(x) G_R(x) dx \oint_{\mathcal{C}_y} \varphi(y) G_R(y) dy \right],$$

where $G_R(x) = R \left(\widetilde{\mathcal{G}}_{\mathbf{Q}}(x) - \mathcal{G}_{\mathbf{Q}}(x) \right)$, the contours \mathcal{C}_x and \mathcal{C}_y are non-overlapping and are taken in the same way as in (12). After interchanging the expectation and integrations, we have

$$\sigma_{\mathcal{I}}^2 = -\frac{1}{4\pi^2} \iint_{\mathcal{C}_x, \mathcal{C}_y} \varphi(x) \varphi(y) \text{Cov}(G_R(x), G_R(y)) dx dy, \quad (14)$$

where $\text{Cov}(G_R(x), G_R(y)) = \mathbb{E}[G_R(x) \cdot G_R(y)]$ is the covariance function of matrix resolvent scaled by the matrix dimension R . In the context of free probability theory, this covariance function is known as the second order Cauchy transform [26] and is denoted as $\text{Cov}(G_R(x), G_R(y)) \triangleq \mathcal{G}_{\mathbf{Q}}(x, y)$.

The rest of this section is devoted to derive the second order Cauchy transform of \mathbf{Q} and the asymptotic variance of capacity $\sigma_{\mathcal{I}}^2$ by using a recent result from free probability theory. Namely, by the framework of the second order freeness [26], [39], the second order Cauchy transform $\mathcal{G}_{\mathbf{Q}}(x, y)$ exists if \mathbf{Q} has a second order limiting distribution according to the following definition:

Definition 1. Let \mathbf{A}_N be an $N \times N$ random matrix. We say that it has a second order limiting distribution if for all $m, n \geq 1$ the moments $\{\alpha_n\}_{n \geq 1}$ and the limits

$$\alpha_{m,n} = \lim_{N \rightarrow \infty} k_2(\text{Tr}(\mathbf{A}_N^m), \text{Tr}(\mathbf{A}_N^n))$$

exist and if for all $r \geq 3$ and all $n(1), \dots, n(r) \geq 1$,

$$\lim_{N \rightarrow \infty} k_r \left(\text{Tr} \left(\mathbf{A}_N^{n(1)} \right), \dots, \text{Tr} \left(\mathbf{A}_N^{n(r)} \right) \right) = 0,$$

where k_r denotes the r -th classic cumulant.

As $\Theta\Theta^\dagger/S$ and $\Psi\Psi^\dagger/R$ are independent complex Wishart matrices, they are unitarily invariant and their second order limiting distributions exist [40, Th. 3.5]. It follows from [26, Eq. (29)] and the cyclic invariant property of matrix trace that \mathbf{Q} also has the second order limiting distribution. The second order Cauchy transform $\mathcal{G}_{\mathbf{Q}}(x, y)$ is given by the functional [26, Eq. (53)]

$$\mathcal{G}_{\mathbf{Q}}(x, y) = \mathcal{G}'_{\mathbf{Q}}(x)\mathcal{G}'_{\mathbf{Q}}(y)\mathcal{R}(\mathcal{G}_{\mathbf{Q}}(x), \mathcal{G}_{\mathbf{Q}}(y)) + \frac{\partial^2}{\partial x \partial y} \log \frac{\mathcal{G}_{\mathbf{Q}}(x) - \mathcal{G}_{\mathbf{Q}}(y)}{x - y}, \quad (15)$$

where $\mathcal{R}(x, y)$ denotes the second order R -transform of \mathbf{Q} . Similar to the first order case, if $|x| \rightarrow \infty$ and $|y| \rightarrow \infty$, $\mathcal{G}_{\mathbf{Q}}(x, y)$ and $\mathcal{R}(x, y)$ have formal power series representations

$$\mathcal{G}_{\mathbf{Q}}(x, y) = \sum_{m, n \geq 1} \alpha_{m, n} x^{-m-1} y^{-n-1}, \quad \text{and} \quad \mathcal{R}(x, y) = \sum_{m, n \geq 1} \kappa_{m, n} x^{m-1} y^{n-1}. \quad (16)$$

In literature, the covariance function $\mathcal{G}_{\mathbf{B}}(x, y)$ for the Wishart type $N \times N$ random matrix $\mathbf{B} = (1/N)\mathbf{X}^\dagger\mathbf{T}\mathbf{X}$ has been studied in [27], where \mathbf{T} is a non-random Hermitian matrix and \mathbf{X} is a Gaussian like⁴ random matrix with *i.i.d.* entries. Therein, the correlation function of \mathbf{B} has the form

$$\mathcal{G}_{\mathbf{B}}(x, y) = \frac{\mathcal{G}'_{\mathbf{B}}(x)\mathcal{G}'_{\mathbf{B}}(y)}{(\mathcal{G}_{\mathbf{B}}(x) - \mathcal{G}_{\mathbf{B}}(y))^2} - \frac{1}{(x - y)^2}, \quad (17)$$

and it is subsequently used to derive an asymptotic variance of Rayleigh MIMO capacity in [16]–[18]. Note that the second term of the right hand side of (15) is exactly the same as (17) by replacing \mathbf{B} with $\mathbf{Q} = (1/R)\Psi^\dagger\mathbf{P}\Psi$ and assuming $\mathbf{P} = \Theta\Theta^\dagger/S$ non-random. Therefore, the fluctuation of capacity $\sigma_{\mathcal{I}}^2$ of Rayleigh product channels has a distinct functional structure from the Rayleigh MIMO channels, see (14) and (15). The increased fluctuation is due to a non-zero R -transform $\mathcal{R}(x, y)$. Closed-form expressions of $\mathcal{G}_{\mathbf{Q}}(x, y)$ and $\mathcal{R}(x, y)$ are summarized in the following proposition:

Proposition 2. *The second order Cauchy transform of \mathbf{Q} is given by (15) with*

$$\mathcal{R}(x, y) = \frac{\mathcal{G}'_{\mathbf{P}}(1/x)\mathcal{G}'_{\mathbf{P}}(1/y)}{x^2 y^2 (\mathcal{G}_{\mathbf{P}}(1/x) - \mathcal{G}_{\mathbf{P}}(1/y))^2} - \frac{1}{(x - y)^2}, \quad (18)$$

where $\mathcal{G}_{\mathbf{P}}(z)$ is the Cauchy transform of a Marčenko-Pastur distribution with the parameter ζ as in (6)

$$\mathcal{G}_{\mathbf{P}}(z) = \frac{1}{2} + \frac{1 - \zeta}{2z} - \sqrt{\frac{1}{4} - \frac{1 + \zeta}{2z} + \frac{(1 - \zeta)^2}{4z^2}}. \quad (19)$$

⁴Each entry of the Gaussian like matrix has the same second and fourth moments as a Gaussian random variable.

Proof. The proof of Proposition 2 depends on the combinatorial structure of cumulants $\{\kappa_n\}_{n \geq 1}$ and $\{\kappa_{m,n}\}_{m,n \geq 1}$ and is given in detail in Section IV. \square

Substitute (15) into (14) and denote $G(x) = \mathcal{G}_{\mathbf{P}}(1/\mathcal{G}_{\mathbf{Q}}(x))$, then the asymptotic variance $\sigma_{\mathcal{I}}^2$ is rewritten as

$$\sigma_{\mathcal{I}}^2 = -\frac{1}{4\pi^2} \iint_{\mathcal{C}_x, \mathcal{C}_y} \frac{\varphi(x)\varphi(y)}{(G(x) - G(y))^2} dG(x)dG(y). \quad (20)$$

In the general setting, it is difficult to further simplify the double integral (20). However, when the transmitter and receiver have equal number of antennas, i.e. $\zeta = 1/\rho$, a compact expression for $\sigma_{\mathcal{I}}^2$ can be obtained. The results are summarized in the following proposition.

Proposition 3. *When $R = T$, the asymptotic variance $\sigma_{\mathcal{I}}^2$ is given by*

$$\sigma_{\mathcal{I}}^2 = \log \frac{\gamma(\omega_r - 1)^2}{\gamma - \omega_r^2(2\omega_r - 2)}, \quad (21)$$

where $\omega_r \leq 0$ is the solution of the cubic equation

$$t^3 - 2t^2 + (1 - \gamma + \gamma\zeta)t + \gamma = 0. \quad (22)$$

Proof. The proof of Proposition 3 is in Appendix A. \square

Using Cardano's formula to solve the cubic equation (22), the explicit expressions for the roots are

$$t_1 = \frac{2}{3} - \frac{3\gamma\zeta - 3\gamma - 1}{3u(\gamma, \zeta)} + \frac{u(\gamma, \zeta)}{3}, \quad (23)$$

$$t_2 = \frac{2}{3} + e^{\frac{i\pi}{3}} \frac{3\gamma\zeta - 3\gamma - 1}{3u(\gamma, \zeta)} - e^{-\frac{i\pi}{3}} \frac{u(\gamma, \zeta)}{3}, \quad (24)$$

$$t_3 = \frac{2}{3} + e^{-\frac{i\pi}{3}} \frac{3\gamma\zeta - 3\gamma - 1}{3u(\gamma, \zeta)} - e^{\frac{i\pi}{3}} \frac{u(\gamma, \zeta)}{3}, \quad (25)$$

where

$$u(\gamma, \zeta) = \left(\sqrt{(3\gamma\zeta - 3\gamma - 1)^3 + \left(1 + \frac{9\gamma}{2} + 9\gamma\zeta\right)^2} - 1 - \frac{9\gamma}{2} - 9\gamma\zeta \right)^{1/3}. \quad (26)$$

For general values of γ and ζ , it is not straightforward to gain insights based on the variance expressions (21), (23)-(26). However, in the high SNR regime with $\gamma \gg 1$, the asymptotic variance $\sigma_{\mathcal{I}}^2$ is characterized by explicit expressions and the behavior of capacity fluctuation can be understood.

Corollary 1. *In the high SNR regime $\gamma \gg 1$, the asymptotic variance $\sigma_{\mathcal{I}}^2$ is given by (21) with ω_r approximated by*

$$\omega_r \approx \begin{cases} \frac{1}{1-\zeta} & \zeta > 1 \\ \frac{2}{3} - \sqrt{(1-\zeta)\gamma} & 0 < \zeta < 1. \end{cases}$$

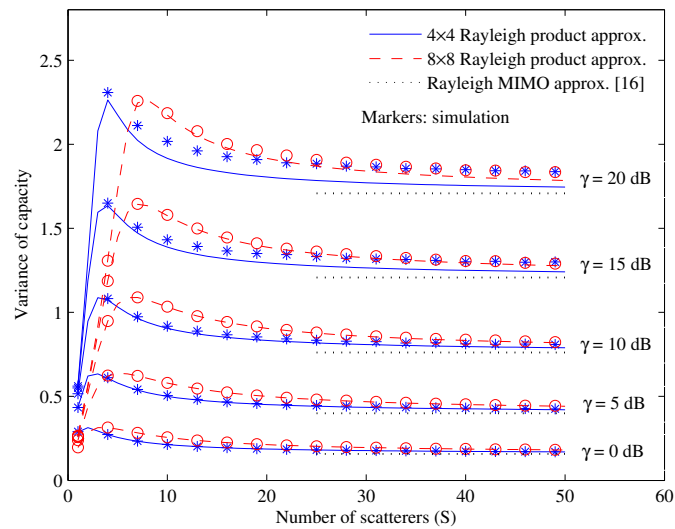
For a fixed SNR, the variance of channel capacity is highest when the number of scattering objects equals to the number of antennas.

Proof. The proof of Corollary 1 is in Appendix B. □

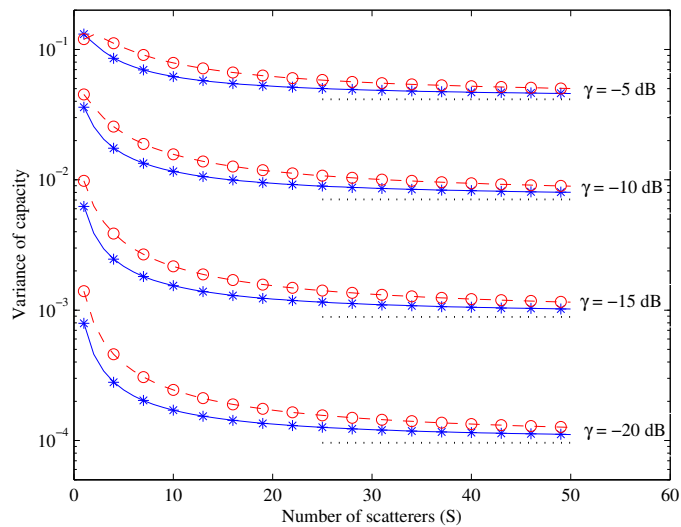
The asymptotic variance $\sigma_{\mathcal{I}}^2$ is derived with the assumption that the dimensions of matrices are large. However, $\sigma_{\mathcal{I}}^2$ serves as a good approximation for the variance of capacity even when the matrix dimensions are comparable to realistic MIMO systems. In Fig. 2, we plot the variance of channel capacity as a function of the number of scattering objects S for 4×4 and 8×8 MIMO systems. The asymptotic variance is calculated by Proposition 3 at SNRs γ ranging from -20 dB to 20 dB with a step size of 5 dB. The analytical calculations are compared with Monte Carlo simulations, where each curve is generated by 10^6 independent channel realizations. We also plot the asymptotic variance of a conventional Rayleigh MIMO channel using [16, Eq. (13)]. Fig. 2 shows that the asymptotic variance achieves a good agreement with the simulations for a wide range of SNRs and numbers of scatterings, especially in the low SNR regime. It is only when $\gamma > 10$ dB that there are observable gaps between analytical and simulation curves. The asymptotic variance for a 8×8 MIMO system remains a better approximation than that of a 4×4 system, as expected. In the high SNR regime, see Fig. 2 (a), there exists a peak value for the variance of capacity when $S > 1$. As the SNR γ increases, the peak of the variance occurs at a fixed value $S = R = T$ ($\zeta = \rho = 1$), which is in line with our prediction in Corollary 1. This is analogous to the observations in [15] that the capacity variance of the conventional Rayleigh MIMO channel is largest when $R = T$. On the other hand, the variance is monotonically decreasing in the low SNR regime, see Fig. 2 (b). As the number of scatterers becomes large, we also observe that the capacity variance of the Rayleigh product channel approaches a limit. This limit is set by the variance of conventional Rayleigh MIMO channel with the same number of antennas. This agrees with the results in [7], where the multi-keyhole channel converges to a Rayleigh MIMO channel when the number of scatterers is large.

IV. SECOND ORDER CUMULANTS AND CAUCHY TRANSFORM

This section is devoted to the proof of Proposition 2, which relies on the knowledge of free cumulants of \mathbf{Q} . Let us recall from (11) and (16) that $\mathcal{R}(z)$ and $\mathcal{R}(x, y)$ are generating functions for the cumulant sequences $\{\kappa_n\}_{n \geq 1}$ and $\{\kappa_{m,n}\}_{m,n \geq 1}$, respectively. We will first deduce the combinatorial descriptions of κ_n and $\kappa_{m,n}$. These results reveal that the cumulant sequences of \mathbf{Q} have the same combinatorial structures as the moment sequences of a Marčenko-Pastur distribution. Namely, $\mathcal{R}(x, y)$ can be obtained based on known results. The notations and terminologies used in the formulation of Lemma 1 and in the



(a)



(b)

Fig. 2. Variance of the Rayleigh product channel capacity. Solid line: asymptotic variance of 4×4 Rayleigh product channel (21); dashed line: asymptotic variance of 8×8 Rayleigh product channel (21); dotted line: asymptotic variance of Rayleigh MIMO channel [16]; markers: simulation. (a) $0 \text{ dB} \leq \gamma \leq 20 \text{ dB}$ (b) $-20 \text{ dB} \leq \gamma \leq -5 \text{ dB}$.

proof of Proposition 2 are given in Appendix C.

Lemma 1. For integers $m, n \geq 1$, the first order free cumulant κ_n of matrix \mathbf{Q} is given by

$$\kappa_n = \rho \sum_{\tau \in \mathcal{S}_{\text{d-nc}}(n)} \zeta^{\#(\tau)}, \quad (27)$$

and the second order free cumulant $\kappa_{m,n}$ is given by

$$\kappa_{m,n} = \sum_{\tau \in \mathcal{S}_{\text{a-nc}}(m,n)} \zeta^{\#(\tau)}, \quad (28)$$

where $\mathcal{S}_{\text{d-nc}}(n)$ and $\mathcal{S}_{\text{a-nc}}(m,n)$ denote the set of non-crossing permutation in disc and annular sense.

Proof. The proof of Lemma 1 is in Appendix D. \square

By comparing (27) with the n -th moment β_n of \mathbf{P} [41, Eq. (7.3)], we have $\kappa_n = \rho \beta_n$ with $n \geq 0$. The cumulant κ_n can be viewed as the scaled version of the moment β_n , where the normalized trace has the normalization factor R instead of the actual matrix dimension S . Similarly, by comparing (28) with the second order moment $\beta_{m,n}$ of \mathbf{P} [41, Eq. (7.5)], we have the second order cumulant $\kappa_{m,n} = \beta_{m,n}$, where no normalization is needed as in Definition 1. Note that the normalization for the first order moments can be arbitrarily chosen without affecting the underlying combinatorial structures provided that the normalization factor grows at the same rate as the matrix dimensions. The functional relations between moments and cumulants as well as their generating functions $\mathcal{G}_{\mathbf{Q}}(\cdot)$, $\mathcal{R}(\cdot)$ are therefore preserved. For notational simplicity, it is convenient to work with the properly scaled moment sequences $\{\beta_n\}$ and $\{\beta_{m,n}\}$.

As β_n is the n -th moment of a Marčenko-Pastur distribution with parameter ζ , by the moment-cumulant relation [42], the corresponding cumulant c_n equals ζ . In addition, it follows from the second order moment-cumulant relation [26] that the moment $\beta_{m,n}$ can be expressed in terms of c_n as well as the corresponding second order cumulant $c_{m,n}$

$$\beta_{m,n} = \sum_{\pi \in \mathcal{S}_{\text{a-nc}}(m,n)} \prod_{i=1}^r c_{n_i} + \sum_{\substack{\pi_1 \in \mathcal{S}_{\text{d-nc}}(m) \\ \pi_2 \in \mathcal{S}_{\text{d-nc}}(n) \\ |\mathcal{V}| = |\pi_1 \times \pi_2| + 1}} c_{m_k, n_l} \prod_{\substack{i=1 \\ i \neq k}}^r c_{m_i} \prod_{\substack{j=1 \\ j \neq l}}^t c_{n_j}. \quad (29)$$

On the right hand side of (29), $\pi \in \mathcal{S}_{\text{a-nc}}(m,n)$ contains $r \geq 1$ orbits and the i -th orbit contains n_i elements with $n_1 + \dots + n_r = m+n$. In the second summation, $\pi_1 \in \mathcal{S}_{\text{d-nc}}(m)$ and $\pi_2 \in \mathcal{S}_{\text{d-nc}}(n)$ contain $r \geq 1$ and $t \geq 1$ orbits, respectively. The i -th orbit of π_1 contains m_i elements with $m_1 + \dots + m_r = m$, and j -th orbit of π_2 contains n_j elements with $n_1 + \dots + n_t = n$. The partition \mathcal{V} is composed of elements from the k -th orbit of π_1 and the l -th orbit of π_2 and corresponds to the second order cumulant c_{m_k, n_l} . Inserting c_n into (29) and comparing with (28), we notice that $\beta_{m,n}$ is entirely determined by the

summation over non-crossing annular permutation and (29) is only valid when the second order cumulant $c_{m,n}$ is zero. In summary, the cumulant sequences are

$$c_n = \zeta \quad \text{and} \quad c_{m,n} = 0, \quad m, n \geq 1. \quad (30)$$

Apply the functional relation [26, Eq. (52)]

$$M(x, y) = C(xM(x), yM(y)) \frac{\frac{d}{dx}(xM(x))}{M(x)} \frac{\frac{d}{dy}(yM(y))}{M(y)} + xy \left(\frac{\frac{d}{dx}(xM(x))}{(xM(x) - yM(y))^2} - \frac{1}{(x - y)^2} \right), \quad (31)$$

where

$$M(x) = 1 + \sum_{n \geq 1} \beta_n x^n, \quad M(x, y) = \sum_{m, n \geq 1} \beta_{m,n} x^m y^n, \quad C(x, y) = \sum_{m, n \geq 1} c_{m,n} x^m y^n. \quad (32)$$

Due to (30), the formal power series $C(x, y) = 0$ and $xM(x) = \mathcal{G}_{\mathbf{P}}(1/x)$, and (31) becomes

$$M(x, y) = \frac{\mathcal{G}'_{\mathbf{P}}(1/x)\mathcal{G}'_{\mathbf{P}}(1/y)}{xy(\mathcal{G}_{\mathbf{P}}(1/x) - \mathcal{G}_{\mathbf{P}}(1/y))^2} - \frac{xy}{(x - y)^2}. \quad (33)$$

Comparing (16) with (32), we obtain $\mathcal{R}(x, y) = M(x, y)/xy$. This completes the proof of Proposition 2.

V. ASYMPTOTIC CAPACITY DISTRIBUTION

In this section, we prove a central limit theorem for the linear spectral statistics of the matrix $\mathbf{Q} = \mathbf{H}\mathbf{H}^\dagger$ and show that the CDF of channel capacity \mathcal{I} is asymptotically Gaussian as the matrix dimensions grow to infinity. This result generalizes the well-known CLT for the correlated Wishart matrix [27]. Together with the asymptotic mean and variance of capacity calculated in Propositions 1 and 3, the Gaussian convergence of capacity \mathcal{I} gives a compact yet accurate approximation for the outage capacity. In addition, the approximative CDF of \mathcal{I} is useful to analyze the diversity-multiplexing tradeoff of Rayleigh product channels in the finite SNR regime.

A. Central Limit Theorem of Linear Spectral Statistics and Outage Probability

Let $H_R(x) = R\left(\widetilde{F}_{\mathbf{Q}}(x) - F_{\mathbf{Q}}(x)\right)$, we are interested in the distribution of random variable

$$\mathcal{I} - \mu_{\mathcal{I}} = \int_{\mathcal{S}_{\mathbf{Q}}} \varphi(x) dH_R(x). \quad (34)$$

Using the integral identities (12) and (13), we can rewrite (34) as

$$\mathcal{I} - \mu_{\mathcal{I}} = \frac{1}{2\pi i} \oint_{\mathcal{C}} \varphi(z) G_R(z) dz, \quad (35)$$

where \mathcal{C} is the closed contour selected as in (12). In the following proposition, we prove that $G_R(z_1)$ and $G_R(z_2)$ with $z_1, z_2 \in \mathcal{C}$ are jointly Gaussian distributed in the asymptotic regime (6).

Proposition 4. *In the asymptotic regime (6), $\{G_R(z)\}_{z \in \mathcal{C}}$ forms a tight sequence (see, e.g. [27]) on a closed contour \mathcal{C} enclosing the support of $F_{\mathbf{Q}}(\cdot)$, and $G_R(z)$ converges weakly to a Gaussian process on the complex plane.*

Proof. The proof of Proposition 4 is in Appendix E. □

By Proposition 4, the asymptotic Gaussianity of (35) follows from the fact that the Riemann sum corresponding to this integral has jointly Gaussian summands and the sum of which can only be Gaussian. Proposition 4 generalizes the CLT of LSS for Wishart type random matrices involving one deterministic correlation matrix and one random matrix with *i.i.d.* entries [27]. When both matrices Ψ and Θ are random and independent, $G_R(z)$ can be decomposed into two random processes, see (66) in Appendix E. Both random processes are asymptotically Gaussian and each is governed by Lemma 3. As already discussed in Section III-B, the induced fluctuation of LSS is characterized by both the first order Cauchy transform and the second order R -transform. This is different from the Wishart random matrices, where the corresponding $G_R(z)$ only involves one asymptotic Gaussian process and the fluctuation of LSS is solely determined by the first order Cauchy transform. This makes the CLT of Rayleigh product ensembles distinct from the one in [27]. Together with the mean $\mu_{\mathcal{I}}$ and variance $\sigma_{\mathcal{I}}^2$ in Propositions 1 and 3, an analytical Gaussian approximation to the capacity distribution of the Rayleigh product channel is available. This result can not be directly derived based on the existing results in [14]–[18]. Note that the CLT of LSS for the biorthogonal ensembles, such as the Rayleigh product ensemble, was proved by Breuer and Duits [28] for polynomial functions $\varphi(x)$. However, it is not clear how to extend this result to generic analytic functions $\varphi(x)$ such as the channel capacity $\varphi(x) = \log(1 + \gamma x)$. Recently, the CLT for the product of two real and square random matrices was proved by Götze, Naumov, and Tikhomirov [43] for smooth function $\varphi(x)$.

Let $\text{erf}(x) = 2/\sqrt{\pi} \int_0^x e^{-t^2} dt$ denote the error function, the Gaussian approximation to the CDF of channel capacity \mathcal{I} is

$$F_{\mathcal{I}}(x) \approx \frac{1}{2} \left(1 + \text{erf} \left(\frac{x - \mu_{\mathcal{I}}}{\sigma_{\mathcal{I}} \sqrt{2}} \right) \right) \quad (36)$$

and thus the outage capacity is

$$\mathcal{I}_{\text{out}} \approx \mu_{\mathcal{I}} + \sigma_{\mathcal{I}} \sqrt{2} \text{erf}^{-1}(2P_{\text{out}} - 1), \quad P_{\text{out}} \in (0, 1). \quad (37)$$

Based on (36) and (37), the outage behavior of the Rayleigh product channel can be understood. In Fig. 3 the impact of the number of scatterers S as well as the received SNR γ is studied, where a 4×4 MIMO system is considered in the presence of $S = 2, 4, 8,$ and 32 scattering objects. We plot the

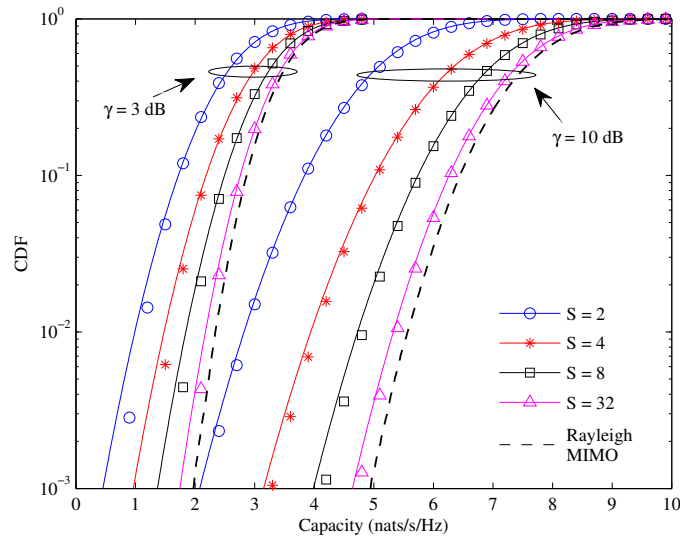


Fig. 3. CDF of channel capacity in the presence of 4×4 Rayleigh product channel. Solid line: Gaussian approximation (36); markers: simulations; dashed line: CDF of 4×4 MIMO capacity with independent Rayleigh fading.

approximative outage probability as well as the empirical one obtained by Monte Carlo simulations. The outage probabilities are evaluated at SNRs $\gamma = 3$ and $\gamma = 10$ dB. As a comparison, we also plot the outage probability of a 4×4 Rayleigh MIMO channel with independent fading entries. As the number of scatterers S increases, the outage capacity at a given probability level rapidly increases until S is equal to the number of antennas, which is especially visible when SNR is large. In this range, the rank of the channel matrix is limited by the number of scatterers and increasing the scatterers effectively improves the rank of channel matrix. When $S > 4$, the matrix rank is limited by the number of antennas and the improvement of outage capacity is relatively slow. Yet, the outage probability curve approaches to a limit, which corresponds to the outage probability of a conventional Rayleigh MIMO channel as predicted by [7].

In Fig. 4 we examine the impact of number of antennas on the outage capacity. We plot the approximative 1% outage capacity (37) as a function of received SNR γ . Assume the number of transmit and receive antennas $T = R = 2, 4, 8$, and 16 while fixing the number of scatterers $S = 8$. As expected, it is seen that the outage capacity of the Rayleigh product channel is lower than the conventional Rayleigh MIMO channel due to the presence of a finite number of scatterers. In the high SNR regime, the outage capacity curves of both channels attain the same slope when $T \leq S$, which suggests that the capacity scales at the same rate as the limiting Rayleigh MIMO channel. On the other hand, when $S < T$ there is

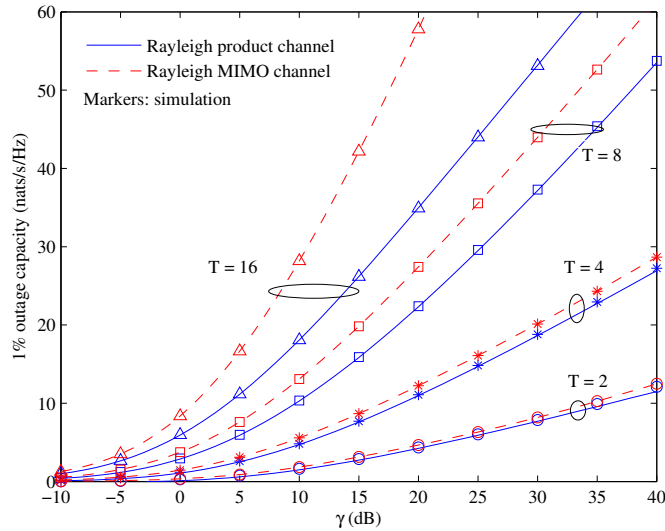


Fig. 4. 1% outage capacity of Rayleigh product channels with $S = 8$ scattering objects and equal numbers of antennas. Solid line: Gaussian quantile approximation (37); markers: simulations; dashed line: outage capacity of conventional Rayleigh MIMO channels with equal number of antennas.

an increasing gap between the two channels as γ increases. Finally, it is observed from Fig. 3 and 4 that the Gaussian approximation (36) and (37) is reasonably accurate for a wide range of parameter settings.

B. Finite-SNR Diversity-Multiplexing Tradeoff

The concept of DMT was originally proposed in [44] to characterize the diversity gain, which is related to link reliability, and the multiplexing gain, which is related to spectral efficiency. The DMT indicates that both types of performance gains can be obtained simultaneously while satisfying a fundamental tradeoff. The operational interpretation of the DMT framework is via the existence of universal codes, which are tradeoff optimal in the high SNR regime [45]. In space-time code design [46], DMT represents a useful analytical tool to characterize the asymptotic performance of codes. However, the asymptotic tradeoff is a too optimistic upper bound to estimate the operational performance at realistic SNRs. Recent works have shown that codes optimized at high SNR may not be optimal at low or moderate SNR. Motivated by these facts, Narasimhan [47] proposed a finite-SNR DMT framework, which characterizes the non-asymptotic DMT. There, he studied the finite DMT for the correlated Rayleigh and Rician MIMO channels at realistic SNR levels.

Under the assumptions of slow fading and capacity achieving codes with rate r , the multiplexing gain

m of a MIMO channel is defined according to [48, Eq. (21)] as

$$m = \frac{n}{\mu_{\mathcal{I}}} r,$$

where $n = \min(R, S, T)$. The multiplexing gain provides an indication of the sensitivity of rate adaptation strategy as the SNR changes. When the applied codes have a higher multiplexing gain, the rate adaptation tends to respond more dramatically to the SNR variations. At a fixed multiplexing gain, the finite-SNR diversity gain $d(m, \gamma)$ is defined as the negative slope of the log-log plot of outage probability $P_{\text{out}}(r)$ at rate $r = m \mu_{\mathcal{I}}/n$ versus SNR γ ,

$$d(m, \gamma) = -\frac{\partial \log P_{\text{out}}(m \mu_{\mathcal{I}}/n)}{\partial \log \gamma}. \quad (38)$$

At a particular SNR γ and multiplexing gain m , the diversity gain (38) provides an estimate of the additional SNR needed to reduce the outage probability by a certain amount. Using the derived outage probability (36), the finite-SNR DMT can be obtained for the Rayleigh product channel.

Proposition 5. *When $R = T$, the finite-SNR DMT of Rayleigh product channels can be approximated by*

$$d(m, \gamma) = \frac{2\gamma \exp(-K(m, \gamma)^2/2)}{\sqrt{\pi} (1 + \text{erf}(-K(m, \gamma)))} \frac{\partial K(m, \gamma)}{\partial \gamma}, \quad (39)$$

where $K(m, \gamma) = \frac{n-m}{\sqrt{2n}} \frac{\mu_{\mathcal{I}}}{\sigma_{\mathcal{I}}}$ with $\mu_{\mathcal{I}}$ and $\sigma_{\mathcal{I}}^2$ calculated by Propositions 1 and 3.

Proof. The proof of Proposition 5 follows by substituting (36) into (38). □

Note that the approximation (39) is tight in the asymptotic regime (6). This is because the approximation error is induced from (36).

Fig. 5 shows the finite-SNR DMT of a 2×2 Rayleigh product channel with $S = 2$ scatterers. The approximated tradeoff curves are generated by (39) at SNRs $\gamma = 0$ dB and $\gamma = 5$ dB. Compared to the Monte Carlo simulations, the proposed approximation yield close estimate for the MIMO diversity gain. As m approaches the maximum multiplexing gain, the discrepancies between the approximation and simulation curves decrease. When $R = T = 4$ antennas are used, the MIMO channel achieves improved channel diversity for a given multiplexing gain as shown in Fig. 6. In both figures, we have also plotted the asymptotic DMT of Rayleigh product channels according to [6, Eq. (8)], when SNR γ approaches infinity. It is clear that the asymptotic results significantly overestimate the channel diversity at the considered operational SNR levels, which justifies the usefulness of the proposed approximation (39).

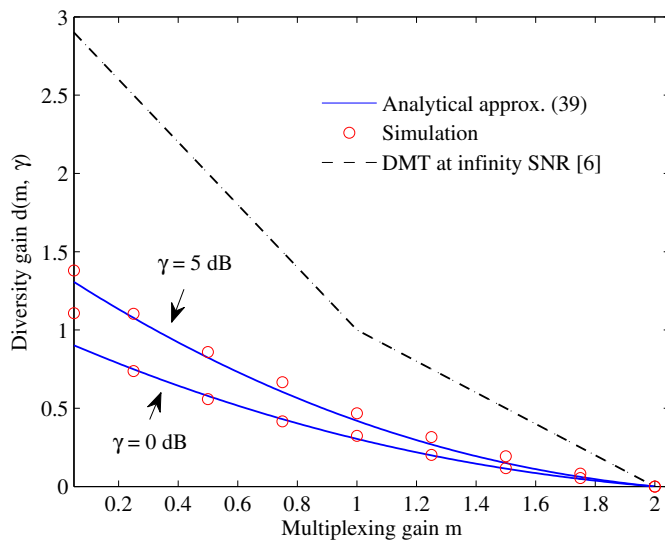


Fig. 5. Finite-SNR DMT of 2×2 Rayleigh product channel with $S = 2$. Solid line: approximation calculated by (39); markers: simulations; dashed line: asymptotic DMT with SNR $\gamma \rightarrow \infty$.

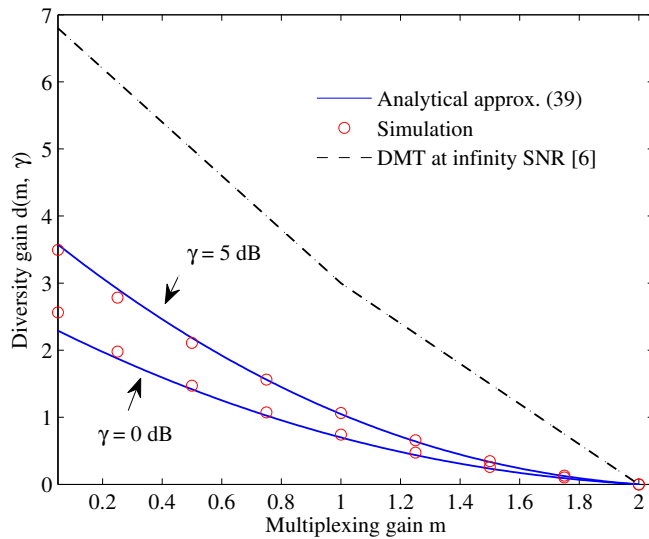


Fig. 6. Finite-SNR DMT of 4×4 Rayleigh product channel with $S = 2$. Solid line: approximation calculated by (39); markers: simulations; dashed line: asymptotic DMT with SNR $\gamma \rightarrow \infty$.

VI. CONCLUSIONS

We studied the outage probability of Rayleigh product channels, which explicitly model the rank deficiency effect. Using free probability theory, the asymptotic variance of channel capacity is calculated for large channel matrix and becomes exact when matrix dimensions approach infinity. Compared to the conventional Rayleigh MIMO channels, the Rayleigh product channels induce a higher capacity fluctuation, which is determined by the second order R -transform of the channel matrix. We have proved that the channel capacity is asymptotically Gaussian by establishing a CLT of a relevant linear spectral statistics. Numerical results show that the proposed Gaussian approximation is reasonably accurate for realistic channel dimensions. Results have been utilized to characterize the tradeoff between diversity and multiplexing of Rayleigh product channels, while the asymptotic tradeoff for large SNR may be an over-optimistic estimate.

ACKNOWLEDGMENT

Z. Zheng and J. Hämäläinen are supported by the Academy of Finland (Grant 284634). L. Wei and J. Corander are supported by the Academy of Finland (Grant 251170). R. Speicher is supported by the ERC Advanced Grant (NCDFP 339760). R. Müller is supported by the Alexander von Humboldt Foundation. We would like to thank O. Arizmendi and A. Soshnikov for helpful discussions.

APPENDIX A

PROOF OF PROPOSITION 3

Let $\mathcal{S}_{\mathbf{Q}}^c \in \mathbb{R}$ denote the complement of the support $\mathcal{S}_{\mathbf{Q}}$ on the real axis. It is shown in [49] that for a given open interval $\mathcal{T} \subset \mathcal{S}_{\mathbf{Q}}^c$, the function $\mathcal{G}_{\mathbf{Q}}(\cdot)$ is continuous, real, and decreasing. This is also true for $\mathcal{G}_{\mathbf{P}}(\cdot)$ with $\mathcal{T} \subset \mathcal{S}_{\mathbf{P}}^c$. Therefore, there exists an inverse function $G^{-1}(t) = \mathcal{G}_{\mathbf{Q}}^{-1}(1/\mathcal{G}_{\mathbf{P}}^{-1}(t))$ continuous, real, and decreasing over $\{t \in \mathbb{R} : t = G(x), x \in \mathcal{S}_{\mathbf{Q}}^c\}$. We choose the contour \mathcal{C}_x to be inside of \mathcal{C}_y such that they both cross real-axis in the intervals $(-1/\gamma, 0)$ and (λ_r, ∞) , where λ_r denotes the right end-point of the support $\mathcal{S}_{\mathbf{Q}}$. By substitutions $t_1 = G(x)$ and $t_2 = G(y)$, the integral (20) can be alternatively integrated over contours \mathcal{C}_1 and \mathcal{C}_2 as

$$\sigma_{\mathcal{I}}^2 = -\frac{1}{4\pi^2} \iint_{\mathcal{C}_1, \mathcal{C}_2} \frac{\varphi(G^{-1}(t_1)) \varphi(G^{-1}(t_2))}{(t_1 - t_2)^2} dt_1 dt_2 = \frac{1}{2\pi i} \oint_{\mathcal{C}_2} \varphi(G^{-1}(t_2)) \mathcal{K}_{\text{inner}}(t_2) dt_2, \quad (40)$$

where

$$\mathcal{K}_{\text{inner}}(t_2) = \frac{1}{2\pi i} \oint_{\mathcal{C}_1} \frac{\log(1 + \gamma G^{-1}(t))}{(t - t_2)^2} dt. \quad (41)$$

The transformed contours \mathcal{C}_1 and \mathcal{C}_2 cross the real-axis in the intervals $(G(0^-), G(-1/\gamma)) = (-\infty, G(-1/\gamma))$ and $(G(\infty), G(\lambda_r)) = (0, G(\lambda_r))$, where

$$G(0^-) = \lim_{x \rightarrow 0^-} G(x), \quad G(\infty) = \lim_{x \rightarrow \infty} G(x).$$

The inverse function $\mathcal{G}_{\mathbf{P}}^{-1}(t)$ is calculated via (19) as

$$\mathcal{G}_{\mathbf{P}}^{-1}(t) = \frac{1}{t} + \frac{\zeta}{1-t}. \quad (42)$$

We obtain the inverse function $\mathcal{G}_{\mathbf{Q}}^{-1}(t)$ by solving the quadratic equation (8) in z as

$$\mathcal{G}_{\mathbf{Q}}^{-1}(t) = \frac{1 - (1 - \zeta)t \pm \sqrt{1 + (1 - \zeta)^2 t^2 - 2(1 + \zeta)t}}{2\zeta t^2}, \quad (43)$$

where the minus sign is taken by the requirement $\lim_{z \rightarrow \infty} z\mathcal{G}_{\mathbf{Q}}(z) = 1$ [49]. Substituting (42) and (43) into (41) and applying integration by parts, we can rewrite $\mathcal{K}_{\text{inner}}(t_2)$ as

$$\begin{aligned} \mathcal{K}_{\text{inner}}(t_2) &= \frac{1}{2\pi i} \oint_{\mathcal{C}_1} \frac{\gamma (G^{-1}(t))'}{(t - t_2)(1 + \gamma G^{-1}(t))} dt \\ &= -\frac{\gamma}{2\pi i} \oint_{\mathcal{C}_1} \frac{2(\zeta - 1)t^2 + 3t - 1}{t(t - 1)(t - t_2)(t - \omega_r)(t - \omega_+)(t - \omega_-)} dt, \end{aligned} \quad (44)$$

where ω_r , ω_+ , and ω_- are the three roots of the cubic equation $t^3 - 2t^2 + (1 - \gamma + \gamma\zeta)t + \gamma = 0$ and ω_r denotes the real solution such that $\omega_r = G(-1/\gamma) < 0$. The integrand of (44) has two simple poles at $t = 0$ and $t = \omega_r$ in \mathcal{C}_1 , and by applying the residue theorem, the integral $\mathcal{K}_{\text{inner}}(t_2)$ becomes

$$\mathcal{K}_{\text{inner}}(t_2) = \frac{1}{t_2} - \frac{1}{t_2 - \omega_r}. \quad (45)$$

Substituting (45) into (40), the variance $\sigma_{\mathcal{I}}^2$ can be therefore expressed as

$$\begin{aligned} \sigma_{\mathcal{I}}^2 &= \frac{1}{2\pi i} \oint_{\mathcal{C}_2} \log(1 + \gamma G^{-1}(t)) \left(\frac{1}{t} - \frac{1}{t - \omega_r} \right) dt \\ &= \frac{1}{2\pi i} \oint_{\mathcal{C}_2} \log \frac{(t - \omega_+)(t - \omega_-)}{(t - 1)^2} \left(\frac{1}{t} - \frac{1}{t - \omega_r} \right) dt + \frac{1}{2\pi i} \oint_{\mathcal{C}_2} \log \frac{t - \omega_r}{t} \left(\frac{1}{t} - \frac{1}{t - \omega_r} \right) dt. \end{aligned} \quad (46)$$

The second integral in (46) has an anti-derivative $(\log \frac{t - \omega_r}{t})^2 / 2$, which is single-valued over \mathcal{C}_2 and therefore vanishes due to Cauchy's theorem. Applying the residue theorem to the first integral in (46), we obtain

$$\sigma_{\mathcal{I}}^2 = \log \frac{(\omega_r - 1)^2 \omega_+ \omega_-}{(\omega_r - \omega_+)(\omega_r - \omega_-)}.$$

The proof is completed by the fact that $\omega_r \omega_+ \omega_- = -\gamma$.

APPENDIX B

PROOF OF COROLLARY 1

When $\zeta > 1$, $u(\gamma, \zeta)$ and $u(\gamma, \zeta)^2$ can be expanded at $\gamma \rightarrow \infty$ as

$$\begin{aligned} u(\gamma, \zeta) &= \sqrt{3(\zeta - 1)}\gamma^{1/2} - \frac{2\zeta + 1}{2(\zeta - 1)} + \frac{\sqrt{3}(4\zeta - 1)}{8(\zeta - 1)^{5/2}}\gamma^{-1/2} + \mathcal{O}(\gamma^{-1}), \\ u(\gamma, \zeta)^2 &= 3(\zeta - 1)\gamma - \frac{\sqrt{3}(2\zeta + 1)}{(\zeta - 1)^{1/2}}\gamma^{1/2} + \mathcal{O}(1). \end{aligned}$$

The real and negative solution ω_r of (22) corresponds to t_1 and it follows from (23) that

$$\omega_r = \frac{2}{3} - \frac{3\gamma\zeta - 3\gamma - 1 - u(\gamma, \zeta)^2}{3u(\gamma, \zeta)} = \frac{1}{1 - \zeta} + \mathcal{O}(\gamma^{-1/2}). \quad (47)$$

When $\zeta < 1$, the asymptotic expansion of $u(\gamma, \zeta)$ at $\gamma \rightarrow \infty$ yields

$$u(\gamma, \zeta) = (-1)^{1/6} \sqrt{3(1 - \zeta)}\gamma^{1/2} + \mathcal{O}(1) = e^{i\pi/6} \sqrt{3(1 - \zeta)}\gamma^{1/2} + \mathcal{O}(1), \quad (48)$$

where we took the principle value $(-1)^{1/6} = e^{i\pi/6}$. In this case, the real and negative solution ω_r of (22) corresponds to t_2 . Inserting (48) into (24), we have

$$\omega_r = \frac{2}{3} - \frac{1}{3} \sqrt{3(1 - \zeta)}\gamma (e^{i\pi/6} + e^{-i\pi/6}) = \frac{2}{3} - \sqrt{(1 - \zeta)}\gamma. \quad (49)$$

Inserting (47) into (21) and taking derivative with respect to ζ , we obtain

$$\frac{\partial}{\partial \zeta} \sigma_{\mathcal{I}}^2 \approx -\frac{2((\zeta - 1)^3\gamma - 2\zeta^2 + \zeta)}{\zeta(\zeta - 1)((\zeta - 1)^3\gamma + 2\zeta)} \approx -\frac{2}{\zeta(\zeta - 1)} < 0.$$

Therefore, when $\zeta > 1$ the asymptotic variance $\sigma_{\mathcal{I}}^2$ is a monotonically decreasing function of ζ . When $\zeta < 1$, the variance $\sigma_{\mathcal{I}}^2$ is a monotonically increasing function of ζ , where the derivative $\partial\sigma_{\mathcal{I}}^2/\partial\zeta > 0$ with ω_r given by (49). To sum up, the asymptotic variance is maximum when ζ approach 1 from both sides of the axis. This completes the proof of Corollary 1.

APPENDIX C

NON-CROSSING PERMUTATIONS

Let us introduce the main combinatorial objects, the non-crossing disc and annular permutations, and the related notations, which are used in Lemma 1 and in the proof of Proposition 2. We refer the readers to [25], [41], [42] for a comprehensive description of the non-crossing permutations.

For a positive integer n , we denote the set $\{1, \dots, n\}$ as $[n]$. Let \mathcal{P}_n denote the set of all partitions of $[n]$. Given a partition $\pi \in \mathcal{P}_n$, we have $\pi = \{B_1, \dots, B_k\}$, where B_1, \dots, B_k , called blocks of π , are non-empty disjoint subset of $[n]$, i.e. $B_1 \cup \dots \cup B_k = [n]$, and $B_i \cap B_j = \emptyset$ for $i \neq j$. Given two

partitions $\pi_1, \pi_2 \in \mathcal{P}_n$, we have $\pi_1 \leq \pi_2$ if and only if every block of π_1 is contained in a block of π_2 and denote $1_n = \{1, \dots, n\}$ the largest partition over $[n]$. We say a partition π is non-crossing in disc sense if there does not exist $1 \leq i, j \leq k$, $i \neq j$, and $1 \leq a < b < c < d \leq n$, such that $a, c \in B_i$ and $b, d \in B_j$. A non-crossing disc partition $\pi \in \mathcal{P}_n$ can be visualized as follows: draw the points $1, \dots, n$ clockwise around the boundary of a disc and connect the points belonging to the same block with a convex hull. The partition π is non-crossing if the convex hulls are pairwise disjoint.

A concept closely related to the partition is the set permutation. Let \mathcal{S}_n denote the set of all permutations over $[n]$. Given a permutation $\tau \in \mathcal{S}_n$, we have $\tau = A_1 \cdots A_k$ such that $[n]$ is decomposed into k orbits and $A_i = (a_i(1), \dots, a_i(s))$ is the i -th orbit of τ containing s elements. For two elements $a_i(p), a_i(q)$ belong to the same orbit A_i , there exists an integer $m \geq 1$ such that $\tau^m(a_i(p)) = a_i(q)$. For instance, if $\tau = (1, 4, 5)(2, 3) \in \mathcal{S}_5$, it maps the elements as $\tau(1) = 4$, $\tau(4) = 5$, $\tau(5) = 1$, $\tau(2) = 3$, and $\tau(3) = 2$. The notion $\#(\tau)$ is used as the number of orbits of τ . We say a permutation is standard in disc sense if for every orbit $A_i = (a_i(1), \dots, a_i(s))$ of τ , there is $a_i(1) < \dots < a_i(s)$. A standard disc permutation τ has an induced partition π , where each block of π contains the same elements as the corresponding orbit of τ . In addition, if the partition induced by standard permutation τ is non-crossing in disc sense, τ is a non-crossing disc permutation and we denote the set of all non-crossing disc permutation on $[n]$ as $\mathcal{S}_{\text{d-nc}}(n)$. Let $\eta = (1, \dots, n)$ the forward cyclic permutation of $[n]$. A permutation $\tau \in \mathcal{S}_n$ satisfy the so-called geodesic condition as

$$\#(\tau) + \#(\tau^{-1}\eta) \leq n + 1, \quad (50)$$

where $\tau^{-1}\eta$ or alternative $\tau^{-1} \circ \eta$ is the composite permutation by first applying η and then τ^{-1} . The equality in (50) only holds when τ is non-crossing disc permutation. The geodesic condition can be intuitively viewed as the triangular inequality for the Cayley graph of permutation group \mathcal{S}_n [50]. Let $\tau_1, \tau_2 \in \mathcal{S}_n$, the distance between τ_1 and τ_2 in Cayley graph of \mathcal{S}_n amounts to $d(\tau_1, \tau_2) = n - \#(\tau_1^{-1}\tau_2)$. The inequality (50) can be rewritten in terms of Cayley distance as $d(id, \tau) + d(\tau, \eta) \geq d(id, \eta)$, where id is the identity permutation. The condition that permutation τ is non-crossing is equivalent that τ lies on the geodesic connecting id and η in the Cayley graph.

Let us consider another set of permutations \mathcal{S}_{m+n} , illustrated via topological drawing in the (m, n) -annular sense. Instead of placing $m+n$ points on the boundary of one disc, we will use two concentric circles. The points $1, \dots, m$ are placed clockwise on the external circle and the points $m+1, \dots, m+n$ are placed counter-clockwise on the internal circle. The annulus between the two circles are referred to as (m, n) -annulus. Given a permutation $\tau \in \mathcal{S}_{m+n}$, it is visualized by drawing curves within the (m, n) -

annulus, which connect the elements of each orbit, respectively. Let $A = (a(1), \dots, a(s))$ an orbit of τ with s elements. The corresponding curve connects $a(1)$ to $a(2)$, then $a(2)$ to $a(3)$, \dots , then $a(s)$ to $a(1)$ such that: 1) it does not intersect with itself; 2) it encloses a region completely contained in (m, n) -annulus; 3) it goes clockwise around the region. We say a permutation τ is (m, n) -connected if there is at least one orbit of τ contains elements on both circles, otherwise τ is (m, n) -disconnected. In addition, $\tau \in \mathcal{S}_{m+n}$ is standard in (m, n) -annular sense if each orbit of τ satisfies either of the following conditions:

- 1) Given an orbit $A \subseteq \tau$ such that $A \subseteq \{1, \dots, m\}$ or $A \subseteq \{m+1, \dots, m+n\}$. The elements of A , upon cyclic permutations, can be sorted in increasing order.
- 2) $A \cap \{1, \dots, m\} \neq \emptyset$ and $A \cap \{m+1, \dots, m+n\} \neq \emptyset$. We have $A = (a(1), \dots, a(k), b(1), \dots, b(l))$, where $a(1), \dots, a(k) \in \{1, \dots, m\}$ and $b(1), \dots, b(l) \in \{m+1, \dots, m+n\}$. Both sequences $\{a(i)\}$ and $\{b(j)\}$, upon cyclic permutations, can be sorted in increasing order, respectively.

We say a permutation $\tau \in \mathcal{S}_{m+n}$ is non-crossing in (m, n) -annular sense if τ is standard and the regions enclosed by every orbits of τ are not overlapping in the annular visualization described above. We denote $\mathcal{S}_{\text{a-nc}}(m, n)$ as the set of non-crossing (m, n) -annular permutations. Finally, according to [41, Th. 6.1], a permutation $\tau \in \mathcal{S}_{m+n}$ and (m, n) -connected satisfies a geodesic condition in the (m, n) -annular sense as

$$\#(\tau) + \#(\tau^{-1}\eta_0) \leq m + n, \quad (51)$$

where $\eta_0 = (1, \dots, m)(m+1, \dots, m+n)$ and the equality only holds when $\tau \in \mathcal{S}_{\text{a-nc}}(m, n)$.

APPENDIX D

PROOF OF LEMMA 1

The proof relies on a known combinatorial identity of the moments of Gaussian random variables, which is stated below.

Lemma 2. (*Wick's Lemma [41]*). *Let Z_1, \dots, Z_t denote i.i.d. complex Gaussian random variables with zero mean and unit variance.*

- 1) *Let m, n be positive integers such that $m \neq n$, and consider two functions $\alpha : [m] \rightarrow [t]$ and $\beta : [n] \rightarrow [t]$. Then*

$$\mathbb{E} [Z_{\alpha(1)} \cdots Z_{\alpha(m)} \bar{Z}_{\beta(1)} \cdots \bar{Z}_{\beta(n)}] = 0.$$

- 2) *Let n be a positive integer and consider two functions $\alpha, \beta : [n] \rightarrow [t]$. Then*

$$\mathbb{E} [Z_{\alpha(1)} \cdots Z_{\alpha(n)} \bar{Z}_{\beta(1)} \cdots \bar{Z}_{\beta(n)}] = \text{card} \{ \tau \in \mathcal{S}_n \mid \alpha = \beta \circ \tau \}, \quad (52)$$

where $\text{card}(\cdot)$ denotes the cardinality.

Denote $\mathbf{Q} = (q_{ij})_{i,j=1}^R$ with the entry q_{ij} given by

$$q_{ij} = \frac{1}{RS} \sum_{a=1}^S \sum_{b=1}^T \sum_{c=1}^S \bar{\psi}_{ai} \theta_{ab} \bar{\theta}_{cb} \psi_{cj}. \quad (53)$$

In light of [26, Th. 2.12], the second order free cumulant of \mathbf{Q} can be expressed in terms of classic joint cumulants of entries q_{ij} as

$$\kappa_{m,n} = \lim_{R \rightarrow \infty} R^{m+n} k_{m+n}(\mathbf{q}_{m,n}), \quad (54)$$

where the vector $\mathbf{q}_{m,n} = [q_{i(1)i(2)}, q_{i(2)i(3)}, \dots, q_{i(m)i(1)}, q_{i(m+1)i(m+2)}, q_{i(m+2)i(m+3)}, \dots, q_{i(m+n)i(m+1)}]$ can be any distinct choice of $i(1), \dots, i(m+n)$. For a partition $\pi \in \mathcal{P}_n$, we define $\pi_i = \{\pi_i(1), \dots, \pi_i(s)\} \subset \pi$ as a block of π with s elements. The expectation over the blocks of partition π is defined as

$$\mathbb{E}_\pi[a_1, \dots, a_n] = \prod_{\pi_i \subset \pi} \mathbb{E}[a_{\pi_i(1)} \cdots a_{\pi_i(s)}].$$

Using the cumulant-moment relations [26, Eq. (10)], k_{m+n} can be written as a sum of $\mathbb{E}_\pi[\mathbf{q}_{m,n}]$ for all $\pi \in \mathcal{P}_{m+n}$, namely

$$k_{m+n}(\mathbf{q}_{m,n}) = \sum_{\pi \in \mathcal{P}_{m+n}} \mathbb{E}_\pi[\mathbf{q}_{m,n}] \text{Möb}_{\mathcal{P}_{m+n}}(\pi, 1_{m+n}), \quad (55)$$

where $\text{Möb}_{\mathcal{P}_{m+n}} : \mathcal{P}_{m+n} \times \mathcal{P}_{m+n} \rightarrow \mathbb{C}$ denotes the Möbius function [42] on \mathcal{P}_{m+n} , which satisfies

$$\sum_{\substack{\eta \in \mathcal{P}_{m+n} \\ \pi \leq \eta}} \text{Möb}_{\mathcal{P}_{m+n}}(\eta, 1_{m+n}) = \begin{cases} 1 & \text{if } \pi = 1_{m+n} \\ 0 & \text{otherwise.} \end{cases} \quad (56)$$

Inserting (53) into (55) and applying Lemma 2, we see that for a given partition π the multiplicative moment $\mathbb{E}_\pi[\mathbf{q}_{m,n}]$ is non-zero only when the partition π takes the forms $\pi^{(1)} = 1_{m+n}$, $\pi^{(2)} = \{\{1, \dots, m\}, \{m+1, \dots, m+n\}\}$. The corresponding Möbius function can be calculated via (56) as

$$\text{Möb}_{\mathcal{P}_{m+n}}(\pi^{(1)}, 1_{m+n}) = 1, \quad \text{Möb}_{\mathcal{P}_{m+n}}(\pi^{(2)}, 1_{m+n}) = -1. \quad (57)$$

It follows from (55) and (57) that the cumulant $k_{m+n}(\mathbf{q}_{m,n})$ equals $k_{m+n}(\mathbf{q}_{m,n}) = \mathbb{E}_{\pi^{(1)}}[\mathbf{q}_{m,n}] - \mathbb{E}_{\pi^{(2)}}[\mathbf{q}_{m,n}]$.

Denote the permutation $\eta_0 = (1, \dots, m)(m+1, \dots, m+n) \in \mathcal{S}_{m+n}$. Substituting $q_{i(t)i(\eta_0(t))}$, $t = 1, \dots, m+n$, into (53), we can express $\mathbb{E}_{\pi^{(1)}}[\mathbf{q}_{m,n}]$ as

$$\frac{1}{(RS)^{m+n}} \sum_{\substack{1 \leq a_1, \dots, a_{m+n} \leq S \\ 1 \leq b_1, \dots, b_{m+n} \leq T \\ 1 \leq c_1, \dots, c_{m+n} \leq S}} \mathbb{E} \left[\prod_{t=1}^{m+n} \bar{\psi}_{a_t i(t)} \theta_{a_t b_t} \bar{\theta}_{c_t b_t} \psi_{c_t i(\eta_0(t))} \right]. \quad (58)$$

It is convenient to introduce the functions $A : [m+n] \rightarrow [S]$, $B : [m+n] \rightarrow [T]$, $C : [m+n] \rightarrow [S]$.

Due to the independence between ψ_{ij} and θ_{kl} , we can rewrite (58) as

$$\frac{1}{(RS)^{m+n}} \sum_{A,B,C} \mathbb{E} \left[\prod_{t=1}^{m+n} \theta_{A(t)B(t)} \prod_{t=1}^{m+n} \bar{\theta}_{C(t)B(t)} \right] \mathbb{E} \left[\prod_{t=1}^{m+n} \psi_{C(\eta_0^{-1}(t))i(t)} \prod_{t=1}^{m+n} \bar{\psi}_{A(t)i(t)} \right]. \quad (59)$$

Since the indexes $i(1), \dots, i(m+n)$ are distinct, by Lemma 2 the second expectation in (59) is non-zero only when $C \circ \eta_0^{-1} = A$. The first expectation in (59) is calculated by (52) with $\alpha(t) = (A(t), B(t))$ and $\beta(t) = (C(t), B(t))$. For a given permutation $\tau \in \mathcal{S}_{m+n}$, the summands of (59) should fulfill $A = C \circ \tau$ and $B = B \circ \tau$ to be able to contribute to the summation. To summarize, $\mathbb{E}_{\pi^{(1)}}[\mathbf{q}_{m,n}]$ is expressed as

$$\mathbb{E}_{\pi^{(1)}}[\mathbf{q}_{m,n}] = \frac{1}{(RS)^{m+n}} \sum_{B,C} \text{card} \{ \tau \in \mathcal{S}_{m+n} \mid C \circ (\tau^{-1}\eta_0) = C, B \circ \tau = B \}. \quad (60)$$

Interchange summation and cardinality operations in (60) and write (60) as a sum over the permutation τ ,

$$\mathbb{E}_{\pi^{(1)}}[\mathbf{q}_{m,n}] = \frac{1}{(RS)^{m+n}} \sum_{\tau \in \mathcal{S}_{m+n}} \text{card} \{ (B, C) \mid C \circ (\tau^{-1}\eta_0) = C, B \circ \tau = B \}. \quad (61)$$

The condition $C \circ (\tau^{-1}\eta_0) = C$ is equivalent to requiring C to be constant on the orbits of $\tau^{-1}\eta_0$. For a given permutation $\tau^{-1}\eta_0$, there are $S^{\#(\tau^{-1}\eta_0)}$ ways to choose indexes C . Similarly, the condition $B = B \circ \tau$ is equivalent to requiring B to be constant on the orbits of τ and there are $T^{\#(\tau)}$ ways to choose indexes B . As a result, (61) equals

$$\mathbb{E}_{\pi^{(1)}}[\mathbf{q}_{m,n}] = \frac{1}{(RS)^{m+n}} \sum_{\tau \in \mathcal{S}_{m+n}} S^{\#(\tau^{-1}\eta_0)} T^{\#(\tau)}. \quad (62)$$

Following the same procedures as in (58)-(62), we obtain $\mathbb{E}_{\pi_1^{(2)}}[\mathbf{q}_{m,n}]$ and $\mathbb{E}_{\pi_2^{(2)}}[\mathbf{q}_{m,n}]$ as

$$\mathbb{E}_{\pi_1^{(2)}}[\mathbf{q}_{m,n}] = \frac{1}{(RS)^m} \sum_{\tau_1 \in \mathcal{S}_m} S^{\#(\tau_1^{-1}\eta_1)} T^{\#(\tau_1)}, \quad (63)$$

$$\mathbb{E}_{\pi_2^{(2)}}[\mathbf{q}_{m,n}] = \frac{1}{(RS)^n} \sum_{\tau_2 \in \mathcal{S}_n} S^{\#(\tau_2^{-1}\eta_2)} T^{\#(\tau_2)}, \quad (64)$$

where the permutations $\eta_1 = (1, \dots, m) \in \mathcal{S}_m$ and $\eta_2 = (1, \dots, n) \in \mathcal{S}_n$. We multiply (63) with (64) and combine the permutations τ_1 and τ_2 to form a new permutation $\tau = \tau_1 \circ \tau_2' \in \mathcal{S}_{m+n}$, where τ_2' is homogeneous to τ_2 with i -th element relabeled as $m+i$. Note that $\pi_\tau \leq \pi_{\eta_0}$, where partitions π_τ and π_{η_0} are induced by τ and η_0 , respectively. The new permutation τ is therefore (m, n) -disconnected, namely

$$\mathbb{E}_{\pi_1^{(2)}}[\mathbf{q}_{m,n}] \mathbb{E}_{\pi_2^{(2)}}[\mathbf{q}_{m,n}] = \frac{1}{(RS)^{m+n}} \sum_{\substack{\tau \in \mathcal{S}_{m+n}, \\ (m,n)\text{-disconnected}}} S^{\#(\tau^{-1}\eta_0)} T^{\#(\tau)}.$$

Inserting $k_{m+n}(\mathbf{q}_{m,n}) = \mathbb{E}_{\pi^{(1)}}[\mathbf{q}_{m,n}] - \mathbb{E}_{\pi^{(2)}}[\mathbf{q}_{m,n}]$ into (54), we obtain

$$\kappa_{m,n} = \lim_{R \rightarrow \infty} \frac{1}{S^{m+n}} \sum_{\substack{\tau \in \mathcal{S}_{m+n}, \\ (m,n)\text{-connected}}} S^{\#(\tau^{-1}\eta_0) + \#(\tau)} \left(\frac{T}{S}\right)^{\#(\tau)}.$$

According to (51), the exponent $\#(\tau) + \#(\tau^{-1}\eta_0) \leq m + n$ for $\tau \in \mathcal{S}_{m+n}$ and is (m, n) -connected. In addition, the equality holds only when τ is non-crossing in the (m, n) -annular sense. Let $S \rightarrow \infty$, all terms in the summation with crossing permutation vanish and for $\tau \in \mathcal{S}_{\text{a-nc}}(m, n)$, $S^{\#(\tau^{-1}\eta_0) + \#(\tau)}$ cancels with S^{m+n} . The derivation for the first order cumulants follows similarly.

APPENDIX E

PROOF OF PROPOSITION 4

Denote $\underline{\mathcal{G}}_{\mathbf{Q}}(z) = \mathbb{E}_{\Psi} \left[\widetilde{\mathcal{G}}_{\mathbf{Q}}(z) \right]$ as the expected resolvent of \mathbf{Q} , which is averaged over the ensembles of Ψ . As matrix \mathbf{P} is random, $\underline{\mathcal{G}}_{\mathbf{Q}}$ is also a random variable and is the solution of (7) with $F_{\mathbf{P}}(\cdot)$ replaced by its empirical version $\widetilde{F}_{\mathbf{P}}(\cdot)$, namely

$$z = \frac{1}{\underline{\mathcal{G}}_{\mathbf{Q}}} + \rho \int \frac{\lambda d\widetilde{F}_{\mathbf{P}}(\lambda)}{1 - \lambda \underline{\mathcal{G}}_{\mathbf{Q}}}. \quad (65)$$

We divide $G_R(z)$ into two parts as

$$G_R(z) = R \left(\widetilde{\mathcal{G}}_{\mathbf{Q}}(z) - \underline{\mathcal{G}}_{\mathbf{Q}}(z) \right) + R \left(\underline{\mathcal{G}}_{\mathbf{Q}}(z) - \mathcal{G}_{\mathbf{Q}}(z) \right) = G_R^1(z) + G_R^2(z). \quad (66)$$

The proof of asymptotic Gaussianity of $G_R(z)$ follows in two steps, showing the asymptotic Gaussianity of $G_R^1(z)$ and $G_R^2(z)$, respectively. The proof then boils down to a direct application of Bai and Silverstein's lemma [27, Lemma 1.1]:

Lemma 3. (CLT of Wishart type ensembles [27]). Consider an $N \times N$ Hermitian matrix $\mathbf{B} = \mathbf{X}^\dagger \mathbf{T} \mathbf{X} / N$ and assume:

- 1) \mathbf{X} is an $n \times N$ complex random matrix with i.i.d. entries, $\mathbb{E}[X_{i,j}] = 0$, $\mathbb{E}[|X_{i,j}|^2] = 1$, and $\mathbb{E}[|X_{i,j}|^4] = 2$;
- 2) \mathbf{T} is a non-random Hermitian nonnegative definite matrix and its ESD $\widetilde{F}_{\mathbf{T}}(\cdot)$ converges weakly to a non-random limiting distribution $F_{\mathbf{T}}(\cdot)$.

Let $M_n(z) = n \left(\widetilde{\mathcal{G}}_{\mathbf{B}}(z) - \mathcal{G}_{\mathbf{B}}(z) \right)$ and \mathcal{C}_z a positive contour enclosing the support of \mathbf{B} , then the sequence $\{M_n(z)\}$ is tight on the contour \mathcal{C}_z , and $M_n(z)$ converges weakly to a Gaussian process on the complex plane with $\mathbb{E}[M_n(z)] = 0$ and $\text{Cov}(M_n(z_1), M_n(z_2))$ given by (17).

Conditioned on \mathbf{P} , it is straightforward to verify that the complex Gaussian matrix Ψ fulfills the assumption 1). The ESD $\widetilde{F}_{\mathbf{P}}(\cdot)$ converges to the Marčenko-Pastur distribution and therefore fulfills the assumption 2). Furthermore, $\underline{\mathcal{G}}_{\mathbf{Q}}$ is, by definition, the average of $\widetilde{\mathcal{G}}_{\mathbf{Q}}$. It thus follows from Lemma 3 that $G_R^1(z)$ given \mathbf{P} converges to a Gaussian process on the complex plane with $\mathbb{E}[G_R^1(z)] = 0$ and

$$\text{Cov}(G_R^1(z_1), G_R^1(z_2)) = \frac{\mathcal{G}'_{\mathbf{Q}}(z_1)\mathcal{G}'_{\mathbf{Q}}(z_2)}{(\underline{\mathcal{G}}_{\mathbf{Q}}(z_1) - \underline{\mathcal{G}}_{\mathbf{Q}}(z_2))^2} - \frac{1}{(z_1 - z_2)^2}. \quad (67)$$

By (65), we have

$$z = \frac{1}{\underline{\mathcal{G}}_{\mathbf{Q}}} + \rho \int \left(\frac{\lambda d\widetilde{F}_{\mathbf{P}}(\lambda)}{1 - \lambda \underline{\mathcal{G}}_{\mathbf{Q}}} - \frac{\lambda d\widetilde{F}_{\mathbf{P}}(\lambda)}{1 - \lambda \mathcal{G}_{\mathbf{Q}}} \right) + \rho \int \frac{\lambda d\widetilde{F}_{\mathbf{P}}(\lambda)}{1 - \lambda \mathcal{G}_{\mathbf{Q}}}. \quad (68)$$

Subtracting (7) from (68) yields

$$\begin{aligned} 0 &= \frac{\underline{\mathcal{G}}_{\mathbf{Q}} - \mathcal{G}_{\mathbf{Q}}}{\underline{\mathcal{G}}_{\mathbf{Q}}\mathcal{G}_{\mathbf{Q}}} - \rho \int \frac{(\underline{\mathcal{G}}_{\mathbf{Q}} - \mathcal{G}_{\mathbf{Q}}) \lambda^2 d\widetilde{F}_{\mathbf{P}}(\lambda)}{(1 - \lambda \underline{\mathcal{G}}_{\mathbf{Q}})(1 - \lambda \mathcal{G}_{\mathbf{Q}})} - \rho \int \left(\frac{\lambda d\widetilde{F}_{\mathbf{P}}(\lambda)}{1 - \lambda \underline{\mathcal{G}}_{\mathbf{Q}}} - \frac{\lambda dF_{\mathbf{P}}(\lambda)}{1 - \lambda \mathcal{G}_{\mathbf{Q}}} \right) \\ \underline{\mathcal{G}}_{\mathbf{Q}} - \mathcal{G}_{\mathbf{Q}} &= \frac{\rho \underline{\mathcal{G}}_{\mathbf{Q}}\mathcal{G}_{\mathbf{Q}}}{C} \int \left(\frac{\lambda d\widetilde{F}_{\mathbf{P}}(\lambda)}{1 - \lambda \underline{\mathcal{G}}_{\mathbf{Q}}} - \frac{\lambda dF_{\mathbf{P}}(\lambda)}{1 - \lambda \mathcal{G}_{\mathbf{Q}}} \right), \end{aligned} \quad (69)$$

where $C = 1 - \rho \underline{\mathcal{G}}_{\mathbf{Q}}\mathcal{G}_{\mathbf{Q}} \int \frac{\lambda^2 d\widetilde{F}_{\mathbf{P}}(\lambda)}{(1 - \lambda \underline{\mathcal{G}}_{\mathbf{Q}})(1 - \lambda \mathcal{G}_{\mathbf{Q}})}$. By definition of Cauchy transform, we have

$$\int \frac{\lambda dF_{\mathbf{P}}(\lambda)}{1 - \lambda \mathcal{G}_{\mathbf{Q}}} = -\frac{1}{\mathcal{G}_{\mathbf{Q}}} + \frac{1}{\mathcal{G}_{\mathbf{Q}}^2} \mathcal{G}_{\mathbf{P}} \left(\frac{1}{\mathcal{G}_{\mathbf{Q}}} \right), \quad \text{and} \quad \int \frac{\lambda d\widetilde{F}_{\mathbf{P}}(\lambda)}{1 - \lambda \mathcal{G}_{\mathbf{Q}}} = -\frac{1}{\mathcal{G}_{\mathbf{Q}}} + \frac{1}{\mathcal{G}_{\mathbf{Q}}^2} \widetilde{\mathcal{G}}_{\mathbf{P}} \left(\frac{1}{\mathcal{G}_{\mathbf{Q}}} \right). \quad (70)$$

By inserting (70) into (69) and multiplying R on both sides of (69), we obtain $G_R^2 = S \left(\widetilde{\mathcal{G}}_{\mathbf{P}}(1/\mathcal{G}_{\mathbf{Q}}) - \mathcal{G}_{\mathbf{P}}(1/\mathcal{G}_{\mathbf{Q}}) \right) / C$. In the asymptotic regime (6), $\underline{\mathcal{G}}_{\mathbf{Q}}$ converges to $\mathcal{G}_{\mathbf{Q}}$ and C converges to $1 - \rho \mathcal{G}_{\mathbf{Q}}^2 \int \frac{\lambda^2 dF_{\mathbf{P}}(\lambda)}{(1 - \lambda \mathcal{G}_{\mathbf{Q}})^2}$. It follows from Lemma 3 (with an $S \times S$ matrix $\mathbf{P} = \mathbf{X}^\dagger \mathbf{X}$ and \mathbf{T} being an identity matrix) that G_R^2 converges to a centered Gaussian process. Note that the covariance (67) of $G_R^1(z)$ is independent of \mathbf{P} and the randomness of $G_R^2(z)$ only comes from \mathbf{P} , which makes $G_R^1(z)$ and $G_R^2(z)$ independent of each other. Combining the above arguments, $G_R(z) = G_R^1(z) + G_R^2(z)$ is asymptotically a sum of two independent Gaussian processes and therefore $G_R(z)$ is also a Gaussian process.

REFERENCES

- [1] D. Gesbert, H. Bölcskei, D. A. Gore, and A. J. Paulraj, "Outdoor MIMO wireless channels: models and performance prediction," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 1926-1934, Dec. 2002.
- [2] R. R. Müller, "A random matrix model of communication via antenna arrays," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2495-2506, Sept. 2002.
- [3] D. Chizhik, G. J. Foschini, M. J. Gans, and R. A. Valenzuela, "Keyholes, correlations, and capacities of multielement transmit and receive antennas," *IEEE Trans. Wireless Commun.*, vol. 1, no. 2, pp. 361-368, Apr. 2002.

- [4] R. R. Müller and H. Hofstetter, "Confirmation of random matrix model for the antenna array channel by indoor measurements," in *Proc. IEEE Antennas Propagat. Soc. Int. Symp.*, vol. 1, 2001, pp. 472-475.
- [5] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels: spatial fading correlation, double scattering, and keyhole," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2636-2647, Oct. 2003.
- [6] S. Yang and J.-C. Belfiore, "Diversity-multiplexing tradeoff of double scattering MIMO channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2027-2034, Apr. 2011.
- [7] G. Levin and S. Loyka, "From multi-keyholes to measure of correlation and power imbalance in MIMO channels: outage capacity analysis," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3515-3529, June 2011.
- [8] J. Hoydis, R. Couillet, and M. Debbah, "Asymptotic analysis of double-scattering channels," in *Proc. ASILOMAR'11*, Nov. 2011, pp. 1935-1939.
- [9] L. Wei, Z. Zheng, O. Tirkkonen, and J. Hämäläinen, "On the ergodic mutual information of multiple cluster scattering MIMO channels," *IEEE Commun. Letters*, vol. 17, no. 9, pp. 1700-1703, Sept. 2013.
- [10] G. Akemann, M. Kieburg, and L. Wei, "Singular value correlation functions for products of Wishart random matrices," *J. Phys. A: Math. Theor.*, vol. 46, no. 27, 2013.
- [11] G. Akemann, J. Ipsen, and M. Kieburg, "Products of rectangular random matrices: singular values and progressive scattering," *Phys. Rev. E*, vol. 88, Nov. 2013.
- [12] IEEE 802.11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High-Speed Physical Layer in the 5 GHz Band, Supplement to IEEE 802.11, Sept. 1999.
- [13] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 359-378, May 1994.
- [14] 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, no. 10, pp. 2545-2561, Oct. 2003.
- [15] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1893-1909, Sept. 2004.
- [16] M. A. Kamath and B. L. Hughes, "The asymptotic capacity of multiple-antenna Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4325-4333, Dec. 2005.
- [17] A. M. Tulino and S. Verdú, "Asymptotic outage capacity of multiantenna channels," in *Proc. ICASSP'05*, Mar. 2005, pp. 825-828.
- [18] M. Debbah and R. R. Müller, "MIMO channel modeling and the principle of maximum entropy," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1667-1690, May 2005.
- [19] 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, no. 4, pp. 1984-2007, Apr. 2011.
- [20] Y. Chen and M. R. McKay, "Coulumb fluid, Painlevé transcendents, and the information theory of MIMO systems," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4594-4634, July 2012.
- [21] S. H. Simon, A. L. Moustakas, and L. Marinelli, "Capacity and character expansions: moment-generating function and other exact results for MIMO correlated channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5336-5351, Dec. 2006.
- [22] M. Kang and M.-S. Alouini, "Capacity of MIMO Rician channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 112-122, Jan. 2006.

- [23] E. Riegler and G. Taricco, "Asymptotic statistics of the mutual information for spatially correlated Rician fading MIMO channels with interference," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1542-1559, Apr. 2010.
- [24] D. Voiculescu, "Addition of certain non-commuting random variables," *J. Funct. Anal.*, vol. 66, pp. 323-346, 1986.
- [25] R. Speicher, "Free probability and random matrices," to appear in *Proc. International Congress of Mathematicians*, Seoul, Aug. 2014.
- [26] B. Collins, J. A. Mingo, P. Śniady, and R. Speicher, "Second order freeness and fluctuations of random matrices: III. Higher order freeness and free cumulants," *Doc. Math.*, vol. 12, pp. 1-70, 2007.
- [27] Z. D. Bai and J. W. Silverstein, "CLT for linear spectral statistics of large-dimensional sample covariance matrices," *Ann. Probab.*, vol. 32, no. 1A, pp. 553-605, 2004.
- [28] J. Breuer and M. Duits, "Central limit theorems for biorthogonal ensembles and asymptotics of recurrence coefficients," arXiv: 1309.6224v2
- [29] D. N. C. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge: Cambridge University Press, 2005.
- [30] P. Almers, F. Tufvesson, and A. F. Molisch, "Keyhole effect in MIMO wireless channels: measurements and theory," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3596-3064, Dec. 2006.
- [31] N. Fawaz, K. Zarifi, M. Debbah, and D. Gesbert, "Asymptotic capacity and optimal precoding in MIMO multi-hop relay networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2050-2069, Apr. 2011.
- [32] Ī. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, no. 6, pp. 585-595, Nov. 1999.
- [33] N. Prasad and M. K. Varanasi, "Outage theorems for MIMO block-fading channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5284-5296, Dec. 2006.
- [34] A. Lozano and N. Jindal, "Are yesterday's information-theoretic fading models and performance metrics adequate for the analysis of today's wireless systems?" *IEEE Commun. Mag.*, vol. 50, no. 11, pp. 210-217, Nov. 2012.
- [35] Z. D. Bai, "Convergence rate of expected spectral distributions of large random matrices. Part II. Sample covariance matrices," *Ann. Probab.*, vol. 21, no. 2, pp. 649-672, 1993.
- [36] J. W. Silverstein, "Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices," *J. Multivariate Anal.*, vol. 55, pp. 331-339, 1995.
- [37] V. A. Marčenko and L. A. Pastur, "Distributions of eigenvalues for some sets of random matrices," *Math. USSR-Sbornik*, vol. 1, pp. 457-483, 1967.
- [38] L. Wei, Z. Zheng, J. Corander, and G. Taricco, "On the outage capacity of orthogonal space-time block codes over multi-cluster scattering MIMO channels," arXiv: 1403.5571v2.
- [39] O. Arizmendi and J. A. Mingo, "Second order even and R-diagonal operators," in preparation.
- [40] J. A. Mingo and R. Speicher, "Second order freeness and fluctuations of random matrices: I. Gaussian and Wishart matrices and cyclic Fock spaces," *J. Funct. Anal.*, vol. 235, pp. 226-270, 2006.
- [41] J. Mingo and A. Nica, "Annular noncrossing permutations and partitions, and second-order asymptotics for random matrices," *Int. Math. Res. Not.*, vol. 28, pp. 1413-1460, 2004.
- [42] A. Nica and R. Speicher, *Lectures on the Combinatorics of Free Probability*. Cambridge: Cambridge University Press, 2006.
- [43] F. Götze, A. Naumov, and Tikhomirov, "Distribution of linear statistics of singular values of the product of random matrices," arXiv: 1412.3314v1.

- [44] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073-1096, May 2003.
- [45] S. Tavildar and P. Viswanath, "Approximately universal codes over slow-fading channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3233-3258, July 2006.
- [46] P. Elia, K. R. Kumar, S. A. Pawar, P. V. Kumar, and H.-F. Lu, "Explicit space-time codes achieving the diversity-multiplexing gain tradeoff," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3869-3884, Sept. 2006.
- [47] R. Narasimhan, "Finite-SNR diversity-multiplexing tradeoff for correlated Rayleigh and Rician MIMO channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3965-3979, Sept. 2006.
- [48] S. Loyka and G. Levin, "Finite-SNR diversity-multiplexing tradeoff via asymptotic analysis of large MIMO systems," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4781-4792, Oct. 2010.
- [49] J. W. Silverstein and S. I. Choi, "Analysis of the limiting spectral distribution of large dimensional random matrices," *J. Multivariate Anal.*, vol. 54, no. 2, pp. 295-309, 1995.
- [50] P. Biane, "Some properties of crossings and partitions," *Discrete Math.*, vol. 175, pp. 41-53, 1997.