On the Capacity of Doubly Correlated MIMO Channels

Hyundong Shin, *Member, IEEE*, Moe Z. Win, *Fellow, IEEE*, Jae Hong Lee, *Senior Member, IEEE*, and Marco Chiani, *Senior Member, IEEE*

***Abstract*— In this paper, we analyze the capacity of multiple-input multiple-output (MIMO) Rayleigh-fading channels in the presence of spatial fading correlation at *both* the transmitter and the receiver, assuming the channel is unknown at the transmitter and perfectly known at the receiver. We first derive the *determinant* representation for the exact characteristic function of the capacity, which is then used to determine the *trace* representations for the mean, variance, skewness, kurtosis, and other higher-order statistics (HOS). These results allow us to exactly evaluate two relevant information-theoretic capacity measures—ergodic capacity and outage capacity—and the HOS of the capacity for such a MIMO channel. The analytical framework presented in the paper is valid for arbitrary numbers of antennas, and generalizes the previously known results for independent and identically distributed or one-sided correlated MIMO channels to the case when fading correlation exists on both sides. We verify our analytical results by comparing them with Monte Carlo simulations for a correlation model based on realistic channel measurements as well as a classical exponential correlation model.**

***Index Terms*— Channel capacity, higher-order statistics (HOS), multiple-input multiple-output (MIMO) system, Rayleigh fading, spatial fading correlation.**

I. Introduction

MULTIPLE-INPUT multiple-output (MIMO) communication systems using multiple transmit and receive antennas promise high spectral efficiency and link reliability for wireless communications [1]–[3]. Although the linear growth of capacity with the number of antennas indicates the potential of MIMO systems, the true benefits of using multiple antennas may be limited by spatial fading correlation due to closely-spaced antenna configurations and poor scattering environments in realistic wireless channels [4], [5].

Since the pioneering work of [1] and [2] in the area of multiple-antenna communications predicted remarkable spectral efficiency of MIMO wireless systems in independent and identically distributed (i.i.d.) Rayleigh fading, much subsequent work has concentrated on characterizing MIMO capacity under correlated fading [4]–[14]. However, the exact analytical results for the capacity, such as ergodic (or mean) capacity, capacity variance, and outage capacity (i.e., capacity versus outage probability),[1] have been known for only a few special cases, largely due to mathematical intractability (see, e.g., [3], [6], [16] for i.i.d. flat Rayleigh fading and [7]–[9] for a one-sided correlated MIMO channel). For a more general case of correlated fading at both the transmitter and the receiver, which we will refer to as *doubly correlated* MIMO channels, some limited results are available: the capacity distribution for a small number of antennas (i.e., $\min\{n_{\mathrm{T}}, n_{\mathrm{R}}\} \leq 3$ where n_{T} and n_{R} are the numbers of transmit and receive antennas, respectively) [10], upper and lower bounds on the ergodic capacity [6], [12], capacity statistics for the case with a large number of antennas [13], and the asymptotic mean and variance of the capacity in the limit as the number of antennas tends to infinity [5], [14]. The temporal behavior of the capacity was analyzed in [11] in terms of level crossing rates and average fade durations.

In this paper, we focus on deriving the exact analytical expressions for capacity statistics of doubly correlated MIMO Rayleigh-fading channels using the methodology developed in [6] and [7], assuming perfect channel knowledge at the receiver and no knowledge at the transmitter with the average input-power constraint. The principal contributions of this paper are as follows.

- We derive a *determinant* representation for the characteristic function (CF) of MIMO capacity, which generalizes the previous results for i.i.d. and one-sided correlated channels [7]–[9] to the doubly correlated case.
- We derive *trace* representations for the mean, variance, and higher-order statistics (HOS) (e.g., cumulants, skewness, and kurtosis) of the capacity using the determinant representation of the CF and the relationship between *polymatrices* and *dimatrices*.[2]

Manuscript received December 17, 2004; revised August 1, 2005; accepted September 24, 2005. The editor coordinating the review of this paper and approving it for publication was S. Aïssa. This research was supported in part by the Korea Research Foundation Grant funded by the Korean Government (KRF-2004-214-D00337), the National Research Laboratory (NRL) Program of Korea, the Charles Stark Draper Endowment, the Office of Naval Research Young Investigator Award N00014-03-1-0489, the National Science Foundation under Grant ANI-0335256, the Institute of Advanced Study Natural Science & Technology Fellowship, and the University of Bologna.

H. Shin was with the Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139 USA. He is now with the School of Electronics and Information, Kyung Hee University, 1 Seocheon, Kihung, Yongin, Kyungki 446-701, Korea (e-mail: hshin@khu.ac.kr).

M. Win is with the Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139 USA (e-mail: moewin@mit.edu).

J. H. Lee is with the School of Electrical Engineering, Seoul National University, Seoul, Korea (e-mail: jhlee@snu.ac.kr).

M. Chiani is with WiLAB, DEIS, University of Bologna, Viale Risorgimento 2, 40136 Bologna, Italy (e-mail: mchiani@deis.unibo.it).

Digital Object Identifier 10.1109/TWC.2006.04850

[1] In general, the capacity distribution is required to determine the outage capacity [2], [3], [15].

[2] The definitions of the polymatrix and the dimatrix will be introduced in Section III.

1536-1276/06$20.00 © 2006 IEEE

- We characterize the effect of fading correlation on the capacity statistics at high signal-to-noise ratio (SNR). We show that at high SNR, the variance, skewness, kurtosis, and other HOS of the capacity depend only on correlation at the side with the larger number of antennas. Moreover, when $n_{\mathrm{T}} = n_{\mathrm{R}}$, these statistics are not affected by fading correlation at any side.

To verify our analytical results, we also compare them with Monte Carlo simulations for doubly correlated MIMO channels using a correlation model based on physical measurements [17], [18], as well as a classical exponential correlation model. It should be noted that alternative derivation of the moment generating function (MGF) of the capacity for doubly correlated MIMO channels can also be found in [19]. In this study, the MGF was obtained *indirectly* from the case of a square channel matrix ($n_{\mathrm{T}} = n_{\mathrm{R}}$) using the *limiting* approach of [20] and then, the first moment (ergodic capacity) was deduced from it in terms of a sum of $\min\{n_{\mathrm{T}}, n_{\mathrm{R}}\}$ determinants.

The remainder of the paper is organized as follows. A brief overview of the distributions of complex random matrices required for our analysis, channel model, and associated channel capacity are presented in Section II. The CF of the capacity is derived and the capacity statistics are analyzed for doubly correlated MIMO channels in Section III. The effect of fading correlation on the capacity statistics is investigated at high SNR in Section IV. In Section V, some numerical and simulation results are provided to illustrate our analytical results. Finally, Section VI concludes the paper.

We shall use the following notation throughout the paper. $\mathbb{N}$ and $\mathbb{C}$ denote the natural numbers and the field of complex numbers, respectively. The superscript $\dagger$ denotes the transpose conjugate. $\mathbf{I}_n$ and $\mathrm{tr}(\mathbf{A})$ represent the $n \times n$ identity matrix and the trace operator of a square matrix $\mathbf{A}$, respectively. By $\mathbf{A} > 0$, we denote that $\mathbf{A}$ is positive definite. For a matrix $\mathbf{A}(t) = [a_{i,j}(t)]$ where $a_{i,j}(t)$ are differentiable functions of t, the nth derivative of $\mathbf{A}(t)$ with respect to t is denoted by

$$\mathbf{A}^{(n)}(t) \triangleq \frac{d^n \mathbf{A}(t)}{dt^n} = \left[\frac{d^n a_{i,j}(t)}{dt^n}\right].$$

II. Preliminaries: Definitions and Models

In this section, we give a brief overview of the distribution theory of complex random matrices (which serves as a central mathematical tool for analyzing MIMO communication systems), channel model, and associated channel capacity.

A. Distributions of Complex Random Matrices

Let us denote a complex Gaussian matrix $\mathbf{X} \in \mathbb{C}^{m\times n}$ with the probability density function (PDF) [6, eq. (1)]

$$p_{\mathbf{X}}(\mathbf{X}) = \frac{\exp\left[-\mathrm{tr}\left\{\mathbf{\Sigma}^{-1}(\mathbf{X}-\mathbf{M})\mathbf{\Psi}^{-1}(\mathbf{X}-\mathbf{M})^{\dagger}\right\}\right]}{\pi^{mn}\det(\mathbf{\Sigma})^{n}\det(\mathbf{\Psi})^{m}} \tag{1}$$

by $\mathbf{X} \sim \tilde{\mathcal{N}}_{m,n}(\mathbf{M}, \mathbf{\Sigma}, \mathbf{\Psi})$ where $\mathbf{\Sigma} \in \mathbb{C}^{m\times m} > 0$ and $\mathbf{\Psi} \in \mathbb{C}^{n\times n} > 0$ are Hermitian. If $\mathbf{X} \sim \tilde{\mathcal{N}}_{m,n}(\mathbf{0}, \mathbf{\Sigma}, \mathbf{I}_n)$, $m \le n$, and $\mathbf{Y} = \mathbf{X}\mathbf{X}^{\dagger}$, then $\mathbf{Y}$ has a complex (central) Wishart density $\tilde{\mathcal{W}}_m(n, \mathbf{\Sigma})$ given by [6, eq. (3)].

Definition 1 (Matrix Quadratic Form [6]): Let

$$\mathbf{X} \sim \tilde{\mathcal{N}}_{m,n}(\mathbf{0}, \mathbf{\Sigma}, \mathbf{\Psi}), \quad m \le n.$$

A positive-definite quadratic form $\mathbf{Y}$ in $\mathbf{X}$ associated with a Hermitian matrix $\mathbf{A} \in \mathbb{C}^{n\times n} > 0$, denoted by $\mathbf{Y} \sim \tilde{Q}_{m,n}(\mathbf{A}, \mathbf{\Sigma}, \mathbf{\Psi})$, is then defined as $\mathbf{Y} = \mathbf{X}\mathbf{A}\mathbf{X}^{\dagger}$.

The PDF of $\mathbf{Y} \sim \tilde{Q}_{m,n}(\mathbf{A}, \mathbf{\Sigma}, \mathbf{\Psi})$ is given by [21, eq. (57)] and can be expressed in an equivalent form

$$p_{\mathbf{Y}}(\mathbf{Y}) = \frac{1}{\tilde{\Gamma}_m(n)}\det(\mathbf{\Sigma})^{-n}\det(\mathbf{A}\mathbf{\Psi})^{-m}\det(\mathbf{Y})^{n-m} \times {}_0\tilde{F}_0^{(n)}\left(-\mathbf{\Sigma}^{-1}\mathbf{Y}, \mathbf{\Psi}^{-1}\mathbf{A}^{-1}\right), \quad \mathbf{Y} > 0 \tag{2}$$

where

$$\tilde{\Gamma}_m(\alpha) = \pi^{m(m-1)/2}\prod_{i=0}^{m-1}\Gamma(\alpha - i), \quad \Re(\alpha) > m-1$$

is the complex multivariate gamma function, $\Gamma(\cdot)$ is Euler's gamma function, and ${}_p\tilde{F}_q^{(n)}(\cdot)$ is the hypergeometric function of two Hermitian matrices, defined by [21, eq. (51)]. Note that the density (2) is a counterpart of the real case in [22, eq. (7.2.5)] and if $\mathbf{A}\mathbf{\Psi} = \mathbf{I}_n$, it reduces to the complex Wishart density $\tilde{\mathcal{W}}_m(n, \mathbf{\Sigma})$.

B. Channel Model and Capacity Random Variable

We consider a point-to-point frequency-flat fading MIMO link with n_{T} transmit and n_{R} receive antennas. Let $\mathbf{x} \in \mathbb{C}^{n_{\mathrm{T}}}$ be a transmitted signal vector with input covariance $\mathbf{Q} = \mathbb{E}\{\mathbf{x}\mathbf{x}^{\dagger}\}$ satisfying the power constraint $\mathrm{tr}(\mathbf{Q}) \le \mathcal{P}$, then the received signal is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{3}$$

where $\mathbf{H} \in \mathbb{C}^{n_{\mathrm{R}}\times n_{\mathrm{T}}}$ is the random channel matrix whose (i,j)th entries H_{ij}, $i = 1, 2, \ldots, n_{\mathrm{R}}$, $j = 1, 2, \ldots, n_{\mathrm{T}}$, are complex propagation coefficients between the jth transmit antenna and the ith receive antenna with $\mathbb{E}\{|H_{ij}|^2\} = 1$, and $\mathbf{n}$ is the complex n_{R}-dimensional zero-mean additive white Gaussian noise (AWGN) vector with covariance $\sigma_n^2\mathbf{I}_{n_{\mathrm{R}}}$. For doubly correlated MIMO channels, the channel matrix $\mathbf{H}$ can be written as [4], [5]

$$\mathbf{H} = \mathbf{\Psi}_{\mathrm{R}}^{1/2}\mathbf{H}_0\mathbf{\Psi}_{\mathrm{T}}^{1/2} \tag{4}$$

where $\mathbf{H}_0 \sim \tilde{\mathcal{N}}_{n_{\mathrm{R}},n_{\mathrm{T}}}(\mathbf{0}, \mathbf{I}_{n_{\mathrm{R}}}, \mathbf{I}_{n_{\mathrm{T}}})$, and $\mathbf{\Psi}_{\mathrm{T}} \in \mathbb{C}^{n_{\mathrm{T}}\times n_{\mathrm{T}}} > 0$ and $\mathbf{\Psi}_{\mathrm{R}} \in \mathbb{C}^{n_{\mathrm{R}}\times n_{\mathrm{R}}} > 0$ are the transmit and receive correlation matrices, respectively. Note that $\mathbf{H} \sim \tilde{\mathcal{N}}_{n_{\mathrm{R}},n_{\mathrm{T}}}(\mathbf{0}, \mathbf{\Psi}_{\mathrm{R}}, \mathbf{\Psi}_{\mathrm{T}})$ and this has been used in various attempts to study correlated MIMO channels [4]–[14]. Recently, this model has also been validated through physical measurements [17].

In what follows, we refer to

$$n_{\mathrm{S}} = \min\{n_{\mathrm{T}}, n_{\mathrm{R}}\}, \quad n_{\mathrm{L}} = \max\{n_{\mathrm{T}}, n_{\mathrm{R}}\}$$

and define the random matrix $\mathbf{\Theta} \in \mathbb{C}^{n_{\mathrm{S}}\times n_{\mathrm{S}}} > 0$ as

$$\mathbf{\Theta} \triangleq \begin{cases} \mathbf{H}\mathbf{H}^{\dagger}, & \text{if } n_{\mathrm{R}} \le n_{\mathrm{T}} \\ \mathbf{H}^{\dagger}\mathbf{H}, & \text{otherwise.} \end{cases}$$

Also, let us denote, for convenience,

$$(\mathbf{\Psi}_{\mathrm{S}}, \mathbf{\Psi}_{\mathrm{L}}) = \begin{cases} (\mathbf{\Psi}_{\mathrm{R}}, \mathbf{\Psi}_{\mathrm{T}}), & \text{if } n_{\mathrm{R}} \le n_{\mathrm{T}} \\ (\mathbf{\Psi}_{\mathrm{T}}, \mathbf{\Psi}_{\mathrm{R}}), & \text{otherwise} \end{cases}$$

and let $0 < \lambda_{\mathrm{S},1} < \lambda_{\mathrm{S},2} < \cdots < \lambda_{\mathrm{S},n_{\mathrm{S}}}$ and $0 < \lambda_{\mathrm{L},1} < \lambda_{\mathrm{L},2} < \cdots < \lambda_{\mathrm{L},n_{\mathrm{L}}}$ be distinct ordered eigenvalues of $\mathbf{\Psi}_{\mathrm{S}}$ and $\mathbf{\Psi}_{\mathrm{L}}$, respectively. Then, $\mathbf{\Theta} \sim \tilde{Q}_{n_{\mathrm{S}},n_{\mathrm{L}}}(\mathbf{I}_{n_{\mathrm{L}}}, \mathbf{\Psi}_{\mathrm{S}}, \mathbf{\Psi}_{\mathrm{L}})$ for the doubly correlated MIMO channel.

In general, when the receiver has perfect channel knowledge, the input distribution that maximizes the mutual information between $\mathbf{x}$ and $\mathbf{y}$ is circularly symmetric complex Gaussian for any given input covariance $\mathbf{Q}$. When the transmitter has no channel knowledge, power among transmit antennas cannot be allocated in accordance with the realization of $\mathbf{H}$ to maximize the mutual information, and hence equal power allocation to each transmit antenna is the most reasonable strategy, i.e., choosing $\mathbf{Q} = (\mathcal{P}/n_{\mathrm{T}})\mathbf{I}_{n_{\mathrm{T}}}$.[3] This yields the capacity in nats/s/Hz as [2], [3]

$$C = \ln\det\left(\mathbf{I}_{n_{\mathrm{S}}} + \frac{\eta}{n_{\mathrm{T}}}\mathbf{\Theta}\right) \tag{5}$$

where $\eta = \mathcal{P}/\sigma_n^2$ is the average SNR at each receive antenna. Since the channel matrix $\mathbf{H}$ is random, the associated channel capacity C is also a random variable whose statistics are determined by the statistical properties of the eigenvalues of $\mathbf{\Theta} \sim \tilde{Q}_{n_{\mathrm{S}},n_{\mathrm{L}}}(\mathbf{I}_{n_{\mathrm{L}}}, \mathbf{\Psi}_{\mathrm{S}}, \mathbf{\Psi}_{\mathrm{L}})$.

III. Capacity Statistics

In this section, we will investigate the statistical properties of the capacity random variable C in (5) for doubly correlated MIMO channels. We begin by deriving the CF of C, from which all other functions—such as the PDF, cumulative distribution function (CDF), and cumulant generating function (CGF)—and statistical moments of C can be obtained.

A. Characteristic Function

Theorem 1: Let $\mathbf{H} \sim \tilde{\mathcal{N}}_{n_{\mathrm{R}},n_{\mathrm{T}}}(\mathbf{0}, \mathbf{\Psi}_{\mathrm{R}}, \mathbf{\Psi}_{\mathrm{T}})$, i.e., $\mathbf{\Theta} \sim \tilde{Q}_{n_{\mathrm{S}},n_{\mathrm{L}}}(\mathbf{I}_{n_{\mathrm{L}}}, \mathbf{\Psi}_{\mathrm{S}}, \mathbf{\Psi}_{\mathrm{L}})$. Then, the CF of the capacity C in nats/s/Hz is

$$\begin{aligned}\Phi_C(\jmath\omega) &\triangleq \mathbb{E}\left\{e^{\jmath\omega C}\right\} \\ &= \frac{\Upsilon_{n_{\mathrm{S}}}(\jmath\omega)}{K_{\mathrm{cor}}}\det\mathbf{\Lambda}(\jmath\omega)\end{aligned} \tag{6}$$

where $\jmath = \sqrt{-1}$,

$$\Upsilon_{n_{\mathrm{S}}}(\jmath\omega) = \prod_{\ell=1}^{n_{\mathrm{S}}-1}(\jmath\omega + \ell)^{-\ell}, \tag{7}$$

$$K_{\mathrm{cor}} = \left(\frac{\eta}{n_{\mathrm{T}}}\right)^{n_{\mathrm{S}}(n_{\mathrm{S}}-1)/2}\prod_{i<j}^{n_{\mathrm{S}}}(\lambda_{\mathrm{S},j} - \lambda_{\mathrm{S},i})\prod_{i<j}^{n_{\mathrm{L}}}(\lambda_{\mathrm{L},j} - \lambda_{\mathrm{L},i}) \tag{8}$$

and $\mathbf{\Lambda}(\jmath\omega)$ is the $n_{\mathrm{L}} \times n_{\mathrm{L}}$ matrix whose (i,j)th entry is given in Table I.

[3]It has been shown in [3] that if the channel has i.i.d. Rayleigh fading between antenna pairs, the optimum input covariance matrix is $\mathbf{Q} = (\mathcal{P}/n_{\mathrm{T}})\mathbf{I}_{n_{\mathrm{T}}}$.

Proof: See Appendix B. □

Note that Theorem 1 requires correlation matrices $\mathbf{\Psi}_{\mathrm{T}}$ and $\mathbf{\Psi}_{\mathrm{R}}$ to have distinct eigenvalues. For the case when the correlation matrices have non-distinct eigenvalues (some of $\lambda_{\mathrm{S},i}$'s or $\lambda_{\mathrm{L},i}$'s are equal), we can obtain the CF as a limiting case of (6) [7], [23]. In particular, when $\mathbf{\Psi}_{\mathrm{T}} = \mathbf{I}_{n_{\mathrm{T}}}$ and $\mathbf{\Psi}_{\mathrm{R}} = \mathbf{I}_{n_{\mathrm{R}}}$ (i.i.d. case), $\Phi_C(\jmath\omega)$ is given by [7, eq. (25)]

$$\Phi_C(\jmath\omega) = K_{\mathrm{iid}}^{-1}\det\mathbf{\Omega}(\jmath\omega) \tag{9}$$

where $K_{\mathrm{iid}} = \prod_{\ell=1}^{n_{\mathrm{S}}}(n_{\mathrm{L}} - \ell)!\,(\ell-1)!$ and $\mathbf{\Omega}(\jmath\omega)$ is the $n_{\mathrm{S}} \times n_{\mathrm{S}}$ Hankel matrix whose (i,j)th entry is given in Table I. Theorem 1 generalizes the previous results for i.i.d. and one-sided correlated channels (which are special cases of non-distinct eigenvalues) [7]–[9] to the doubly correlated MIMO channel given by (4). Using the analytical formulas for the CF in (6) and (9), the PDF and CDF of C can be expressed in forms of the inverse Fourier transform of $\Phi_C(\jmath\omega)$, which can be efficiently calculated by using the fast Fourier transform (FFT) method [7], [24].

B. Mean, Variance, and Higher-Order Statistics

From the CF of C in (6) and (9) involving the determinants, we derive the exact closed-form expressions for the mean, variance, and other HOS such as cumulants, skewness, and (excess) kurtosis of the capacity. To do this, we first introduce the logarithmic derivative of a determinant.

1) Logarithmic Derivatives of a Determinant: Let $\mathbf{R}(t)$ be a matrix depending on a parameter t. If each entry of $\mathbf{R}(t)$ is differentiable with respect to t, then so is $\det\mathbf{R}(t)$ because the determinant is a polynomial in the entries of $\mathbf{R}(t)$. If $\mathbf{R}(t)$ is invertible, the first-order logarithmic derivative of $\det\mathbf{R}(t)$ is given by [25]

$$\frac{d\ln\det\mathbf{R}(t)}{dt} = \mathrm{tr}\left\{\mathbf{R}^{-1}(t)\,\mathbf{R}^{(1)}(t)\right\}. \tag{10}$$

We now generalize (10) to the arbitrary order of differentiation.

Definition 2 (Polymatrix and Dimatrix): Let $\mathbf{R}(t)$ be an invertible matrix whose elements are differentiable with respect to t. Then, the nth *polymatrix* of $\mathbf{R}(t)$ with respect to t is defined as

$$\mathbf{R}_{[n]}(t) \triangleq \mathbf{R}^{-1}(t)\,\mathbf{R}^{(n)}(t). \tag{11}$$

In particular, we call $\mathbf{R}_{[1]}(t)$ the *dimatrix* of $\mathbf{R}(t)$.

Lemma 1: The polymatrices and the derivatives of the dimatrix of $\mathbf{R}(t)$ have the following relationship

$$\mathbf{R}_{[n]}(t) = \sum_{\ell=1}^{n}\binom{n-1}{\ell-1}\mathbf{R}_{[n-\ell]}(t)\,\mathbf{R}_{[1]}^{(\ell-1)}(t) \tag{12}$$

and the ℓth logarithmic derivative of $\det\mathbf{R}(t)$ is the trace of the $(\ell-1)$th derivative of the dimatrix $\mathbf{R}_{[1]}(t)$, i.e.,

$$\frac{d^{\ell}\ln\det\mathbf{R}(t)}{dt^{\ell}} = \mathrm{tr}\left\{\mathbf{R}_{[1]}^{(\ell-1)}(t)\right\}. \tag{13}$$

TABLE I

SOME MATRICES INVOLVED IN THE ANALYTICAL EXPRESSIONS FOR THE CAPACITY STATISTICS (SEE APPENDIX A FOR DETAILS ON THE EVALUATION OF INTEGRALS). DENOTE $\bar{\eta} = \eta / n_{\mathrm{T}}$.

Notation	Dimension	(i,j)th entry (where $\imath = i - n_{\mathrm{L}} + n_{\mathrm{S}}$ and $n \in \mathbb{N}$)
$\mathbf{\Lambda}(\nu)$	$n_{\mathrm{L}} \times n_{\mathrm{L}}$	$\{\mathbf{\Lambda}(\nu)\}_{i=1,2,\ldots,n_{\mathrm{L}}-n_{\mathrm{S}},j=1,2,\ldots,n_{\mathrm{L}}} = \lambda_{\mathrm{L},j}^{i-1}$ $\{\mathbf{\Lambda}(\nu)\}_{i=n_{\mathrm{L}}-n_{\mathrm{S}}+1,n_{\mathrm{L}}-n_{\mathrm{S}}+2,\ldots,n_{\mathrm{L}},j=1,2,\ldots,n_{\mathrm{L}}}$ $= \lambda_{\mathrm{L},j}^{n_{\mathrm{L}}-n_{\mathrm{S}}-1} \int_0^\infty (1+\bar{\eta}\lambda_{\mathrm{S},\imath} z)^{\nu+n_{\mathrm{S}}-1} e^{-z/\lambda_{\mathrm{L},j}} dz$ $= \lambda_{\mathrm{L},j}^{n_{\mathrm{L}}-n_{\mathrm{S}}} \, {}_2F_0\left(1, -\nu-n_{\mathrm{S}}+1; -\bar{\eta}\lambda_{\mathrm{S},\imath}\lambda_{\mathrm{L},j}\right)$
$\mathbf{\Lambda}^{(n)}(\nu)$	$n_{\mathrm{L}} \times n_{\mathrm{L}}$	$\{\mathbf{\Lambda}^{(n)}(\nu)\}_{i=1,2,\ldots,n_{\mathrm{L}}-n_{\mathrm{S}},j=1,2,\ldots,n_{\mathrm{L}}} = 0$ $\{\mathbf{\Lambda}^{(n)}(\nu)\}_{i=n_{\mathrm{L}}-n_{\mathrm{S}}+1,n_{\mathrm{L}}-n_{\mathrm{S}}+2,\ldots,n_{\mathrm{L}},j=1,2,\ldots,n_{\mathrm{L}}}$ $= \lambda_{\mathrm{L},j}^{n_{\mathrm{L}}-n_{\mathrm{S}}-1} \int_0^\infty (1+\bar{\eta}\lambda_{\mathrm{S},\imath} z)^{\nu+n_{\mathrm{S}}-1} \ln^n\left(1+\bar{\eta}\lambda_{\mathrm{S},\imath} z\right) e^{-z/\lambda_{\mathrm{L},j}} dz$ $= n! \left(\bar{\eta}\lambda_{\mathrm{S},\imath}\right)^{\nu+n_{\mathrm{S}}-1} \lambda_{\mathrm{L},j}^{\nu+n_{\mathrm{L}}-1} \exp\left(\frac{1}{\bar{\eta}\lambda_{\mathrm{S},\imath}\lambda_{\mathrm{L},j}}\right)$ $\times G_{n+1,n+2}^{n+2,0}\left(\frac{1}{\bar{\eta}\lambda_{\mathrm{S},\imath}\lambda_{\mathrm{L},j}} \middle\vert \begin{smallmatrix} 1,1,\ldots,1 \\ 0,0,\ldots,0,\nu+n_{\mathrm{S}} \end{smallmatrix}\right)$
$\mathbf{\Omega}(\nu)$	$n_{\mathrm{S}} \times n_{\mathrm{S}}$	$\{\mathbf{\Omega}(\nu)\}_{i,j=1,2,\ldots,n_{\mathrm{S}}}$ $= \int_0^\infty (1+\bar{\eta}z)^{\nu} z^{n_{\mathrm{L}}-n_{\mathrm{S}}+i+j-2} e^{-z} dz$ $= (n_{\mathrm{L}}-n_{\mathrm{S}}+i+j-2)! \, {}_2F_0\left(n_{\mathrm{L}}-n_{\mathrm{S}}+i+j-1, -\nu; -\bar{\eta}\right)$
$\mathbf{\Omega}^{(n)}(\nu)$	$n_{\mathrm{S}} \times n_{\mathrm{S}}$	$\{\mathbf{\Omega}^{(n)}(\nu)\}_{i,j=1,2,\ldots,n_{\mathrm{S}}}$ $= \int_0^\infty (1+\bar{\eta}z)^{\nu} \ln^n(1+\bar{\eta}z) z^{n_{\mathrm{L}}-n_{\mathrm{S}}+i+j-2} e^{-z} dz$ $= \frac{n! e^{1/\bar{\eta}}}{\bar{\eta}^{n_{\mathrm{L}}-n_{\mathrm{S}}+i+j-1}} \sum_{k=0}^{n_{\mathrm{L}}-n_{\mathrm{S}}+i+j-2} \Bigg[(-1)^{n_{\mathrm{L}}-n_{\mathrm{S}}+i+j-k-2}$ $\times \binom{n_{\mathrm{L}}-n_{\mathrm{S}}+i+j-2}{k} \bar{\eta}^{\nu+k+1} G_{n+1,n+2}^{n+2,0}\left(\frac{1}{\bar{\eta}} \middle\vert \begin{smallmatrix} 1,1,\ldots,1 \\ 0,0,\ldots,0,\nu+k+1 \end{smallmatrix}\right)\Bigg]$

Proof: By definition, we have

$$\begin{aligned}\mathbf{R}_{[n]}(t) &= \mathbf{R}^{-1}(t) \frac{d^{n-1}\mathbf{R}^{(1)}(t)}{dt^{n-1}} \\ &\overset{(a)}{=} \mathbf{R}^{-1}(t) \sum_{\ell=1}^{n} \binom{n-1}{\ell-1} \frac{d^{n-\ell}\mathbf{R}(t)}{dt^{n-\ell}} \frac{d^{\ell-1}\mathbf{R}_{[1]}(t)}{dt^{\ell-1}} \\ &\overset{(b)}{=} \sum_{\ell=1}^{n} \binom{n-1}{\ell-1} \mathbf{R}_{[n-\ell]}(t) \mathbf{R}_{[1]}^{(\ell-1)}(t)\end{aligned}$$

where (a) follows from the Leibniz's identity [26, p. 21] and (b) follows from (11). Also, (13) follows immediately from (10), (11), and by interchanging the order of differentiation and trace operators. □

This lemma says that the ℓth logarithmic derivative of the determinant of a matrix can be determined by its first ℓ polymatrices. For example, the second, third, and fourth order logarithmic derivatives of $\det \mathbf{R}(t)$ are given respectively by

$$\frac{d^2 \ln \det \mathbf{R}(t)}{dt^2} = \mathrm{tr}\left\{\mathbf{R}_{[2]}(t) - \mathbf{R}_{[1]}^2(t)\right\} \tag{14}$$

$$\frac{d^3 \ln \det \mathbf{R}(t)}{dt^3} = \mathrm{tr}\left\{2\mathbf{R}_{[1]}^3(t) - 3\mathbf{R}_{[1]}(t)\mathbf{R}_{[2]}(t) + \mathbf{R}_{[3]}(t)\right\} \tag{15}$$

and

$$\begin{aligned}\frac{d^4 \ln \det \mathbf{R}(t)}{dt^4} = \mathrm{tr}\Big\{&-6\mathbf{R}_{[1]}^4(t) + 12\mathbf{R}_{[1]}^2(t)\mathbf{R}_{[2]}(t) \\ &-3\mathbf{R}_{[2]}^2(t) - 4\mathbf{R}_{[1]}(t)\mathbf{R}_{[3]}(t) + \mathbf{R}_{[4]}(t)\Big\}.\end{aligned} \tag{16}$$

Using the explicit determinantal CF's in Section III-A and Lemma 1, we now derive statistical moments of C, which requires determining the polymatrices

$$\mathbf{\Omega}_{[n]}(\nu) = \mathbf{\Omega}^{-1}(\nu)\mathbf{\Omega}^{(n)}(\nu), \quad \mathbf{\Omega}_{[0]}(\nu) = \mathbf{I}_{n_{\mathrm{S}}} \tag{17}$$

$$\mathbf{\Lambda}_{[n]}(\nu) = \mathbf{\Lambda}^{-1}(\nu)\mathbf{\Lambda}^{(n)}(\nu), \quad \mathbf{\Lambda}_{[0]}(\nu) = \mathbf{I}_{n_{\mathrm{L}}} \tag{18}$$

for i.i.d. and doubly correlated MIMO channels, respectively, where the (i,j)th entries of $\mathbf{\Omega}^{(n)}(\nu)$ and $\mathbf{\Lambda}^{(n)}(\nu)$ for $n \in \mathbb{N}$ are given in Table I.

2) Cumulants: The nth cumulant of C is by definition expressed as

$$\kappa_n \triangleq \left.\frac{d^n}{d\nu^n}\mathcal{K}_C(\nu)\right|_{\nu=0} \tag{19}$$

where $\mathcal{K}_C(\nu) \triangleq \ln \Phi_C(\nu)$ is the CGF of C. Note that the first and second cumulants are the mean and variance of the capacity, respectively.

Theorem 2: Let $\mathbf{H} \sim \tilde{\mathcal{N}}_{n_\mathrm{R},n_\mathrm{T}}(\mathbf{0},\mathbf{I}_{n_\mathrm{R}},\mathbf{I}_{n_\mathrm{T}})$, i.e., $\mathbf{\Theta} \sim \tilde{\mathcal{W}}_{n_\mathrm{S}}(n_\mathrm{L},\mathbf{I}_{n_\mathrm{S}})$ ($n_\mathrm{R}\times n_\mathrm{T}$ i.i.d. MIMO channel). Then, the nth cumulant of the capacity C in nats/s/Hz is

$$\kappa_n = \mathrm{tr}\left\{\mathbf{\Omega}_{[1]}^{(n-1)}(0)\right\}. \tag{20}$$

Proof: It follows immediately from (9), (19), and Lemma 1. □

Theorem 3: Let $\mathbf{H} \sim \tilde{\mathcal{N}}_{n_\mathrm{R},n_\mathrm{T}}(\mathbf{0},\mathbf{\Psi}_\mathrm{R},\mathbf{\Psi}_\mathrm{T})$, i.e., $\mathbf{\Theta} \sim \tilde{Q}_{n_\mathrm{S},n_\mathrm{L}}(\mathbf{I}_{n_\mathrm{L}},\mathbf{\Psi}_\mathrm{S},\mathbf{\Psi}_\mathrm{L})$ ($n_\mathrm{R}\times n_\mathrm{T}$ doubly correlated MIMO channel). Then, the nth cumulant of the capacity C in nats/s/Hz is

$$\kappa_n = \mathrm{tr}\left\{\mathbf{\Lambda}_{[1]}^{(n-1)}(0)\right\} + (-1)^n(n-1)!\sum_{\ell=1}^{n_\mathrm{S}-1}\ell^{-n+1}. \tag{21}$$

Proof: It follows immediately from Theorem 1, (19), and Lemma 1. □

Using the relationship (12), the matrices $\mathbf{\Omega}_{[1]}^{(n-1)}(\nu)$ and $\mathbf{\Lambda}_{[1]}^{(n-1)}(\nu)$ in Theorems 2 and 3 can be determined by the polymatrices in (17) and (18), respectively.

3) Raw and Central Moments: The nth raw and central moments (i.e., moments about the origin and the mean, respectively) of C can be obtained from the general relationships between moments and cumulants [27]:

$$m_n \triangleq \mathbb{E}\{C^n\} = \sum_{\ell=1}^{n}\binom{n-1}{\ell-1}m_{n-\ell}\,\kappa_\ell \tag{22}$$

and

$$\mu_n \triangleq \mathbb{E}\{(C-m_1)^n\} = \sum_{\ell=0}^{n}\binom{n}{\ell}m_{n-\ell}(-m_1)^\ell \tag{23}$$

where m_n and μ_n are the nth raw and central moments of C, respectively. In particular, the ergodic capacity is given by $m_1 = \mathbb{E}\{C\}$.

4) Skewness and Kurtosis: The skewness characterizes the degree of asymmetry of a distribution around its mean, and the kurtosis measures the peakedness or flatness of a distribution relative to a Gaussian distribution. The skewness and (excess) kurtosis of C, denoted by β_1 and β_2, respectively, can be obtained from the cumulants in Theorems 2 and 3 for i.i.d. and doubly correlated cases as

$$\beta_1 \triangleq \frac{\mu_3}{\mu_2^{3/2}} = \frac{\kappa_3}{\kappa_2^{3/2}} \tag{24}$$

and

$$\beta_2 \triangleq \frac{\mu_4}{\mu_2^2} - 3 = \frac{\kappa_4}{\kappa_2^2}. \tag{25}$$

Using (14)–(16), (20), (21), (24), and (25), the trace representations for the mean, variance, skewness, and kurtosis of the capacity are tabulated in Tables II and III for i.i.d. and doubly correlated MIMO channels, respectively.[4] Since the skewness and kurtosis of a Gaussian distribution are equal to zero, nonzero values of these quantities indicate the degree of deviation from the Gaussian distribution.[5]

IV. Effect of Correlation at High SNR

The effect of fading correlation on the behavior of capacity is not immediately apparent from the exact analytical expressions such as the CF in (6) and the moments in Table III. Therefore, we resort to the asymptotic analysis to investigate such an effect in the following. In particular, we consider a high-SNR regime, since the benefits of the use of multiple antennas are more pronounced at high SNR. In this case, the capacity (5) can be written as

$$C = n_\mathrm{S}\ln(\eta/n_\mathrm{T}) + \ln\det(\mathbf{\Theta}) + O(1/\eta), \tag{26}$$

which reveals that at high SNR, the capacity is characterized by the logarithmic generalized variance $\ln\det(\mathbf{\Theta})$ of $\mathbf{\Theta} \sim \tilde{Q}_{n_\mathrm{S},n_\mathrm{L}}(\mathbf{I}_{n_\mathrm{L}},\mathbf{\Psi}_\mathrm{S},\mathbf{\Psi}_\mathrm{L})$. Note in (26) that $\min\{n_\mathrm{T},n_\mathrm{R}\}$ (more precisely, the rank of $\mathbf{H}$) determines the spatial multiplexing gain of a MIMO channel, while the logarithmic generalized variance $\ln\det(\mathbf{\Theta})$ determines the diversity gain in capacity point of view. Consequently, if the correlation matrices have full rank, then the channel matrix $\mathbf{H}$ in (4) has full rank with probability one. In this case, antenna correlation does not diminish the spatial multiplexing gain and only decreases the diversity gain [6].

Starting with (26) and using similar steps in the proof of Theorem 1, the CF of the capacity in the high-SNR regime can be written as

$$\begin{aligned}\Phi_C(\jmath\omega)\big|_{\eta\ \text{high}} &= A\int_{0<z_1\le\cdots\le z_{n_\mathrm{S}}<\infty}\prod_{\ell=1}^{n_\mathrm{S}}z_\ell^{\jmath\omega}\prod_{i<j}^{n_\mathrm{S}}(z_j-z_i)\\ &\qquad\times\det(\mathbf{\Xi})\,dz_1\cdots dz_{n_\mathrm{S}}\\ &= A\left\{\prod_{\ell=1}^{n_\mathrm{S}}\Gamma(\jmath\omega+\ell)\right\}\det\mathbf{K}(\jmath\omega)\end{aligned} \tag{27}$$

where

$$A = \frac{(\eta/n_\mathrm{T})^{\jmath\omega n_\mathrm{S}}\det(\mathbf{\Psi}_\mathrm{S})^{\jmath\omega}}{\prod_{\ell=1}^{n_\mathrm{S}}\Gamma(\ell)\prod_{i<j}^{n_\mathrm{L}}(\lambda_{\mathrm{L},j}-\lambda_{\mathrm{L},i})} \tag{28}$$

and $\mathbf{K}(\jmath\omega)$ is the $n_\mathrm{L}\times n_\mathrm{L}$ matrix whose (i,j)th entry is given by

$$\{\mathbf{K}(\jmath\omega)\}_{i,j} = \begin{cases}\lambda_{\mathrm{L},j}^{i-1}, & i=1,\ldots,n_\mathrm{L}-n_\mathrm{S},\ j=1,\ldots,n_\mathrm{L}\\ \lambda_{\mathrm{L},j}^{\jmath\omega+i-1}, & i=n_\mathrm{L}-n_\mathrm{S}+1,\ldots,n_\mathrm{L},\ j=1,\ldots,n_\mathrm{L}.\end{cases} \tag{29}$$

[4]Alternative expressions for the ergodic capacity and capacity variance of the i.i.d. case, in terms of integrals involving the Laguerre polynomials, can be found in [3] and [16], respectively.

[5]The skewness and kurtosis of the capacity have been used in conjunction with the Gram-Charlier expansion to estimate the Kullback-Leiblier divergence, as a measure of non-Gaussianity, between the capacity distribution and its Gaussian approximation [24].

TABLE II

MEAN, VARIANCE, SKEWNESS, AND KURTOSIS OF THE CAPACITY IN NATS/S/HZ FOR I.I.D. MIMO CHANNELS, $\mathbf{H} \sim \tilde{\mathcal{N}}_{n_{\mathrm{R}},n_{\mathrm{T}}}\left(\mathbf{0}, \mathbf{I}_{n_{\mathrm{R}}}, \mathbf{I}_{n_{\mathrm{T}}}\right)$. DENOTE $\boldsymbol{\Omega}_n = \boldsymbol{\Omega}_{[n]}(0)$.

m_1 (mean)	$\mathrm{tr}\left(\boldsymbol{\Omega}_1\right)$
μ_2 (variance)	$\mathrm{tr}\left(\boldsymbol{\Omega}_2 - \boldsymbol{\Omega}_1^2\right)$
β_1 (skewness)	$\dfrac{\mathrm{tr}\left(2\,\boldsymbol{\Omega}_1^3 - 3\,\boldsymbol{\Omega}_1\boldsymbol{\Omega}_2 + \boldsymbol{\Omega}_3\right)}{\left[\mathrm{tr}\left(\boldsymbol{\Omega}_2 - \boldsymbol{\Omega}_1^2\right)\right]^{3/2}}$
β_2 (kurtosis)	$\dfrac{\mathrm{tr}\left(-6\,\boldsymbol{\Omega}_1^4 + 12\,\boldsymbol{\Omega}_1^2\,\boldsymbol{\Omega}_2 - 3\,\boldsymbol{\Omega}_2^2 - 4\,\boldsymbol{\Omega}_1\boldsymbol{\Omega}_3 + \boldsymbol{\Omega}_4\right)}{\left[\mathrm{tr}\left(\boldsymbol{\Omega}_2 - \boldsymbol{\Omega}_1^2\right)\right]^2}$

TABLE III

MEAN, VARIANCE, SKEWNESS, AND KURTOSIS OF THE CAPACITY IN NATS/S/HZ FOR DOUBLY CORRELATED MIMO CHANNELS, $\mathbf{H} \sim \tilde{\mathcal{N}}_{n_{\mathrm{R}},n_{\mathrm{T}}}\left(\mathbf{0}, \boldsymbol{\Psi}_{\mathrm{R}}, \boldsymbol{\Psi}_{\mathrm{T}}\right)$. DENOTE $\boldsymbol{\Lambda}_n = \boldsymbol{\Lambda}_{[n]}(0)$.

m_1 (mean)	$\mathrm{tr}\left(\boldsymbol{\Lambda}_1\right) - \left(n_{\mathrm{S}} - 1\right)$
μ_2 (variance)	$\mathrm{tr}\left(\boldsymbol{\Lambda}_2 - \boldsymbol{\Lambda}_1^2\right) + \sum_{\ell=1}^{n_{\mathrm{S}}-1} \ell^{-1}$
β_1 (skewness)	$\dfrac{\mathrm{tr}\left(2\boldsymbol{\Lambda}_1^3 - 3\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2 + \boldsymbol{\Lambda}_3\right) - \sum_{\ell=1}^{n_{\mathrm{S}}-1} 2\ell^{-2}}{\left[\mathrm{tr}\left(\boldsymbol{\Lambda}_2 - \boldsymbol{\Lambda}_1^2\right) + \sum_{\ell=1}^{n_{\mathrm{S}}-1} \ell^{-1}\right]^{3/2}}$
β_2 (kurtosis)	$\dfrac{\mathrm{tr}\left(-6\boldsymbol{\Lambda}_1^4 + 12\boldsymbol{\Lambda}_1^2\boldsymbol{\Lambda}_2 - 3\boldsymbol{\Lambda}_2^2 - 4\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_3 + \boldsymbol{\Lambda}_4\right) + \sum_{\ell=1}^{n_{\mathrm{S}}-1} 6\ell^{-3}}{\left[\mathrm{tr}\left(\boldsymbol{\Lambda}_2 - \boldsymbol{\Lambda}_1^2\right) + \sum_{\ell=1}^{n_{\mathrm{S}}-1} \ell^{-1}\right]^2}$

Also, from Lemma 1, (19), and (27), the nth cumulant of the capacity in nats/s/Hz at high SNR becomes

$$\begin{aligned}\kappa_n\big|_{\eta \text{ high}} &= \mathrm{tr}\left\{\mathbf{K}_{[1]}^{(n-1)}(0)\right\} \\ &\quad + \delta_{1n} \cdot \left[n_{\mathrm{S}} \ln\left(\frac{\eta}{n_{\mathrm{T}}}\right) + \ln\det\left(\boldsymbol{\Psi}_{\mathrm{S}}\right)\right] \\ &\quad + \sum_{\ell=1}^{n_{\mathrm{S}}} \psi^{(n-1)}(\ell) \end{aligned} \tag{30}$$

where

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

is the Kronecker delta and

$$\psi^{(n)}(z) = \frac{d^{n+1} \ln \Gamma(z)}{dz^{n+1}}$$

is the polygamma function.[6] In particular, if $n_{\mathrm{R}} = n_{\mathrm{T}}$, (30) reduces to

$$\begin{aligned}\kappa_n\big|_{\eta \text{ high}} &= \delta_{1n} \cdot \left[n_{\mathrm{T}} \ln\left(\frac{\eta}{n_{\mathrm{T}}}\right) + \ln\det\left(\boldsymbol{\Psi}_{\mathrm{T}}\boldsymbol{\Psi}_{\mathrm{R}}\right)\right] \\ &\quad + \sum_{\ell=1}^{n_{\mathrm{T}}} \psi^{(n-1)}(\ell). \end{aligned} \tag{31}$$

From (30) and (31), we have the following observations.

- At high SNR, the mean capacity decreases by the amount of $\ln\det\left(\boldsymbol{\Psi}_{\mathrm{S}}\right)$ due to correlation at the side with the smaller number of antennas and by the amount of $\mathrm{tr}\left\{\mathbf{K}_{[1]}(0)\right\}$ due to correlation at the side with the larger number of antennas.
- The variance, skewness, kurtosis, and other HOS of the capacity depend only on correlation at the side with the larger number of antennas and they converge to finite quantities determined by (22)–(25) and (30), as $\eta \to \infty$.
- When $n_{\mathrm{R}} = n_{\mathrm{T}}$, fading correlation at any side does not affect the variance and HOS in the high-SNR regime, while the mean capacity decreases by the amount of

[6]For $z \in \mathbb{N}$, the digamma function $\psi^{(0)}(z)$, trigamma function $\psi^{(1)}(z)$, tetragamma function $\psi^{(2)}(z)$, and pentagamma function $\psi^{(3)}(z)$ can be expressed as, respectively,

$$\psi^{(0)}(z) = -\gamma + \sum_{n=1}^{z-1} \frac{1}{n}$$

$$\psi^{(1)}(z) = \frac{\pi^2}{6} - \sum_{n=1}^{z-1} \frac{1}{n^2}$$

$$\psi^{(2)}(z) = -2\zeta(3) + \sum_{n=1}^{z-1} \frac{2}{n^3}$$

$$\psi^{(3)}(z) = \frac{\pi^4}{15} - \sum_{n=1}^{z-1} \frac{6}{n^4}$$

where $\gamma \approx 0.5772156649$ is the Euler-Mascheroni constant and $\zeta(3) \approx 1.2020569$ is Apéry's constant.

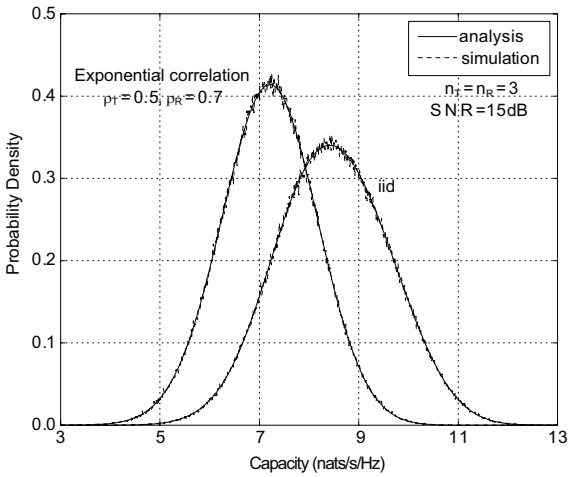

Fig. 1. PDF of the capacity for i.i.d. and exponentially correlated ($\rho_{\mathrm{T}} = 0.5$, $\rho_{\mathrm{R}} = 0.7$) MIMO channels. $n_{\mathrm{T}} = n_{\mathrm{R}} = 3$ and $\eta = 15$ dB.

$\ln\det(\mathbf{\Psi}_{\mathrm{T}}\mathbf{\Psi}_{\mathrm{R}})$. Moreover, it follows from (24), (25), and (31) that

$$-\frac{12\sqrt{6}\cdot\zeta(3)}{\pi^3} \le \beta_1\big|_{\eta\ \text{high}} < 0$$

and

$$0 < \beta_2\big|_{\eta\ \text{high}} \le 2.4$$

which imply that the capacity distribution has an asymmetric tail extending out more to the left of its mean and is *leptokurtic* (i.e., more peaked than a Gaussian distribution) in the high-SNR regime.

- For a single-input single-output (SISO) case ($n_{\mathrm{T}} = n_{\mathrm{R}} = 1$), we have

$$m_1\big|_{\eta\ \text{high}} = \ln(\eta) - \gamma$$
$$\mu_2\big|_{\eta\ \text{high}} = \frac{\pi^2}{6}$$
$$\beta_1\big|_{\eta\ \text{high}} = -\frac{12\sqrt{6}\cdot\zeta(3)}{\pi^3}$$
$$\beta_2\big|_{\eta\ \text{high}} = 2.4$$

which reveal that $-C$ at high SNR follows the *extreme value distribution* [27].

V. Numerical and Simulation Results

To illustrate our analytical results, we consider the exponential correlation model

$$\mathbf{\Phi}_n^{(\exp)}(\rho) = \left[\rho^{|i-j|}\right]_{i,j=1,2,\ldots,n}, \quad \rho \in [0,1)$$

as well as the multiple element transmit receive antennas (METRA) model [17] in our numerical examples. The former model is reasonable in the case of the equally-spaced linear array. The latter model characterizes the correlation properties of MIMO channels using a reduced set of physical parameters such as antenna spacing, power angular spectrum, azimuth spread, and angle of arrival. This model was validated based on measured data collected in both picocell and microcell environments [17], and also has been proposed recently for mobile broadband wireless access (MBWA) MIMO channels [18].

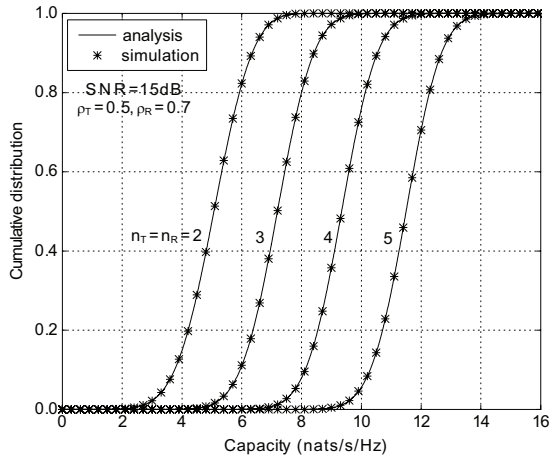

Fig. 2. CDF of the capacity for exponentially correlated MIMO channels with $\rho_{\mathrm{T}} = 0.5$ and $\rho_{\mathrm{R}} = 0.7$. $n_{\mathrm{T}} = n_{\mathrm{R}} = 2$, 3, 4, 5 and $\eta = 15$ dB.

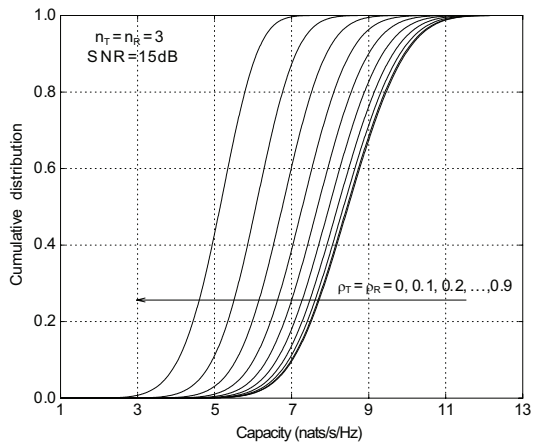

Fig. 3. CDF of the capacity for exponentially correlated MIMO channels with $\rho_{\mathrm{T}} = \rho_{\mathrm{R}} = 0$ (i.i.d.), 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9. $n_{\mathrm{T}} = n_{\mathrm{R}} = 3$ and $\eta = 15$ dB.

A. Exponential Correlation Model

In all examples for exponential correlation, we set $\mathbf{\Psi}_{\mathrm{T}} = \mathbf{\Phi}_{n_{\mathrm{T}}}^{(\exp)}(\rho_{\mathrm{T}})$ and $\mathbf{\Psi}_{\mathrm{R}} = \mathbf{\Phi}_{n_{\mathrm{R}}}^{(\exp)}(\rho_{\mathrm{R}})$. Fig. 1 shows the PDF of C for i.i.d. and exponentially correlated ($\rho_{\mathrm{T}} = 0.5$, $\rho_{\mathrm{R}} = 0.7$) MIMO channels at $\eta = 15$ dB when $n_{\mathrm{T}} = n_{\mathrm{R}} = 3$. The analytical curves are plotted by using (6), (9), and [24, eq. (4.33)]. We also compare our analytical results with the simulated PDF obtained by generating 100 000 realizations of $\mathbf{H}$. It can be seen that analytical and simulated curves match exactly. The figure also shows that the mass of the PDF is mostly above a certain level due to the spatial multiplexing gain (for example, 4 nats/s/Hz for the exponentially correlated case and 5 nats/s/Hz for the i.i.d. case).

Fig. 2 shows the CDF of C for exponentially correlated MIMO channels with $\rho_{\mathrm{T}} = 0.5$ and $\rho_{\mathrm{R}} = 0.7$ at $\eta = 15$ dB when $n_{\mathrm{T}} = n_{\mathrm{R}} = 2$, 3, 4, and 5. The analytical curves are plotted by using (6) and [24, eq. (4.34)], and they agree exactly with the simulated ones. It can be seen that the capacity increases linearly with the number of antennas for the entire range of cumulative probability, despite the presence of correlation. This can be attributed to the spatial multiplexing gain achieved by increasing the number of antennas at both sides. For example, the capacity at the cumulative probability

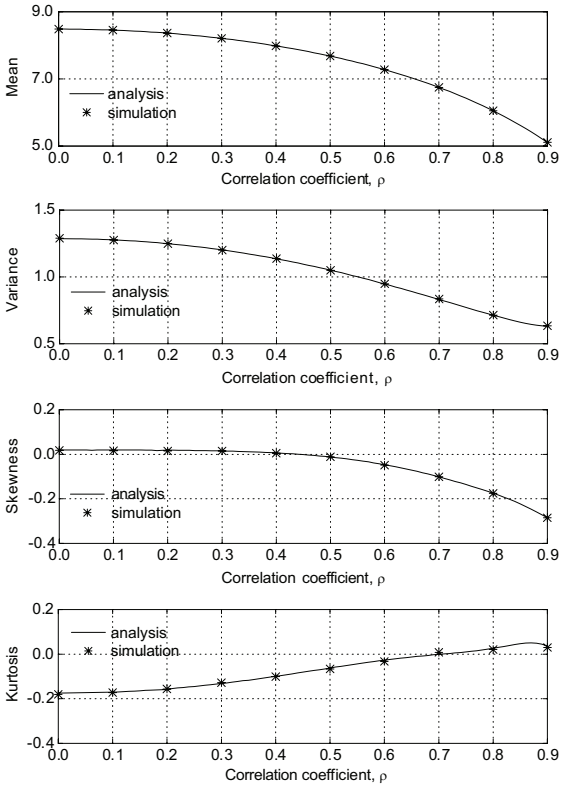

Fig. 4. Mean, variance, skewness, and kurtosis of the capacity (nats/s/Hz) as a function of correlation coefficient ρ for exponentially correlated MIMO channels with $\rho_{\mathrm{T}} = \rho_{\mathrm{R}} = \rho$. $n_{\mathrm{T}} = n_{\mathrm{R}} = 3$ and $\eta = 15$ dB.

of 0.1 (i.e., 10% outage capacity) is about 3.76, 5.95, 8.12, and 10.30 nats/s/Hz for $n_{\mathrm{T}} = n_{\mathrm{R}} = 2, 3, 4$, and 5, respectively; we can gain approximately 2.18 nats/s/Hz of additional capacity for each increase in the number of antennas at both the transmitter and the receiver. Also, the spatial multiplexing gain of MIMO systems guarantees a certain transmission rate at arbitrarily low outage probability (e.g., 2.00, 4.18, 6.36, and 8.54 nats/s/Hz for $n_{\mathrm{T}} = n_{\mathrm{R}} = 2, 3, 4$, and 5, respectively). Fig. 3 illustrates the effect of exponential fading correlation on the capacity distribution for the case of $n_{\mathrm{T}} = n_{\mathrm{R}} = 3$ and $\eta = 15$ dB, where ρ_{T} and ρ_{R} range from 0 to 0.9. It can be seen that the decrease in capacity due to exponential correlation is negligible for a small amount of correlation, but it becomes more significant as the correlation coefficient increases. Moreover, the capacity reduction is more pronounced at high cumulative (or outage) probability.

Fig. 4 shows the analytical and simulated mean, variance, skewness, and kurtosis of C for exponentially correlated MIMO channels as a function of correlation coefficient ρ for $\rho_{\mathrm{T}} = \rho_{\mathrm{R}} = \rho$, $n_{\mathrm{T}} = n_{\mathrm{R}} = 3$, and $\eta = 15$ dB. Again, our analytical results are in excellent agreement with Monte Carlo simulations that are carried out by generating 100 000 realizations of $\mathbf{H}$. Fig. 5 shows the mean, variance, skewness, and kurtosis of C versus SNR for i.i.d. and doubly correlated ($\rho_{\mathrm{T}} = 0.5$, $\rho_{\mathrm{R}} = 0.7$) MIMO channels when $n_{\mathrm{T}} = n_{\mathrm{R}} = 2$. It can be seen that as the SNR increases, the variance, skewness, and kurtosis for both i.i.d. and doubly correlated cases converge to 2.290, -0.810, and 1.333 according to (24), (25), and (31), respectively, and the effect of fading correlation on these statistics diminishes.

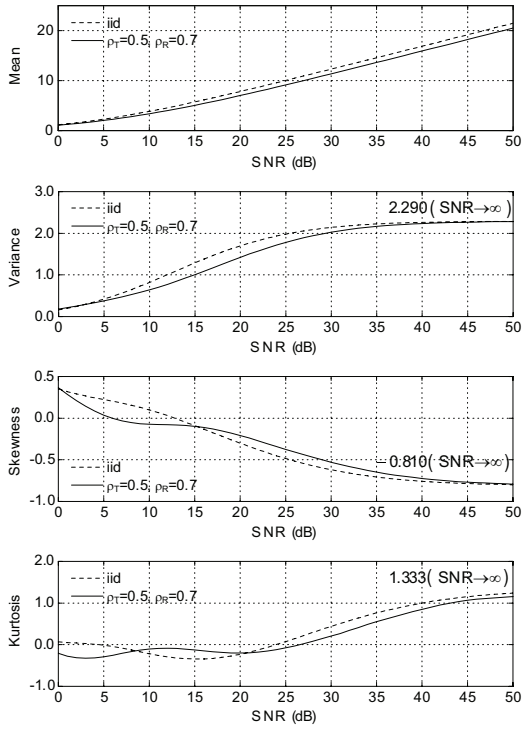

Fig. 5. Mean, variance, skewness, and kurtosis of the capacity (nats/s/Hz) versus SNR for i.i.d. and doubly correlated ($\rho_{\mathrm{T}} = 0.5$, $\rho_{\mathrm{R}} = 0.7$) MIMO channels. $n_{\mathrm{T}} = n_{\mathrm{R}} = 2$.

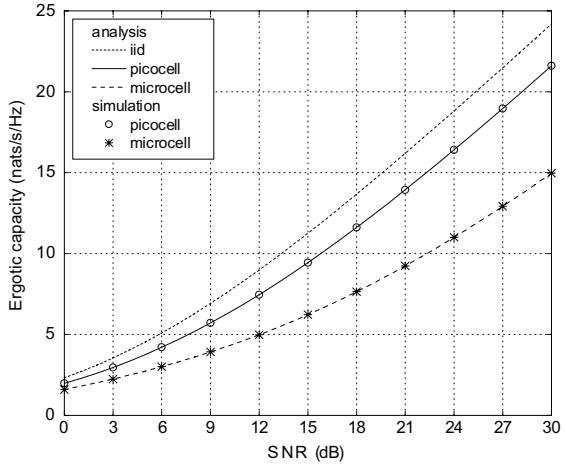

Fig. 6. Ergodic capacity of doubly correlated MIMO channels with $n_{\mathrm{T}} = n_{\mathrm{R}} = 4$ and METRA correlation matrices given in [17, p. 82] for picocell and microcell environments. For comparison, the ergodic capacity of the 4×4 i.i.d. MIMO channel is also plotted.

B. METRA Correlation Model

We now consider 4×4 MIMO channels with correlation matrices obtained by the METRA model [17], [18]. In all examples, it can be observed that our analytical results agree exactly with Monte Carlo simulations. Fig. 6 shows the ergodic capacity for METRA correlation matrices, given in [17, p. 82], for picocell and microcell environments. The picocell example is a partially decorrelated scenario selected from a small office environment, whereas the microcell case corresponds to an environment where the receiver (base station) is highly correlated (see [17] for details on the antenna configurations and environment setups). We can see from Fig. 6 that the ergodic capacity at $\eta = 15$ dB is 11.25, 9.44, and 6.22

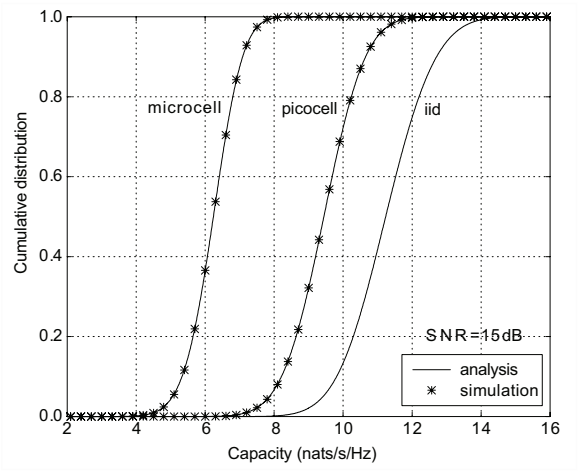

Fig. 7. CDF of the capacity for doubly correlated MIMO channels with $n_{\mathrm{T}} = n_{\mathrm{R}} = 4$ and METRA correlation matrices given in [17, p. 82] for picocell and microcell environments. For comparison, the CDF of capacity for the 4×4 i.i.d. MIMO channel is also plotted. $\eta = 15$ dB.

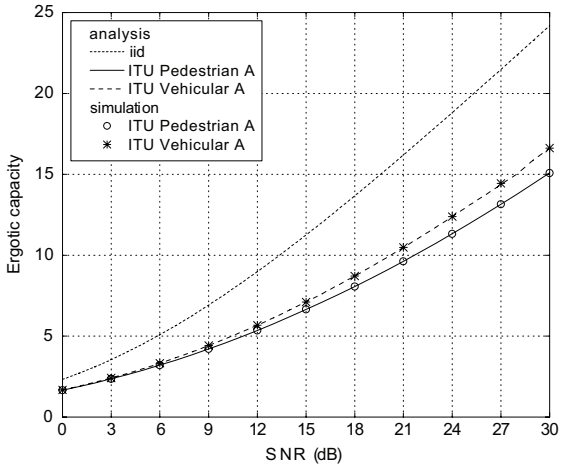

Fig. 8. Ergodic capacity of doubly correlated MIMO channels with $n_{\mathrm{T}} = n_{\mathrm{R}} = 4$ and METRA correlation matrices given in [18] for macrocell ITU Pedestrian A and Vehicular A environments. For comparison, the ergodic capacity of the 4×4 i.i.d. MIMO channel is also plotted.

nats/s/Hz for i.i.d., picocell, and microcell cases respectively. The reduction in ergodic capacity due to spatial correlation is about 16% for picocell and 45% for microcell environments, respectively. Fig. 7 shows the CDF of the capacity at $\eta = 15$ dB in the same environments as in Fig. 6. The 10% outage capacity is 9.82, 8.21, and 5.33 nats/s/Hz for i.i.d., picocell, and microcell scenarios respectively. The reduction in 10% outage capacity due to spatial correlation is about 16% for picocell and 46% for microcell environments, which is similar to the amount of reduction in ergodic capacity.

We next consider the METRA correlation matrices, given in [18], for 4×4 MIMO channels in macrocell Pedestrian A and Vehicular A environments of the international telecommunication union (ITU) standard. For these correlation matrices, the ergodic capacity is shown in Fig. 8 and the CDF of the capacity at $\eta = 15$ dB is shown in Fig. 9. The ergodic capacity at $\eta = 15$ dB is 6.64 nats/s/Hz for the ITU Pedestrian A and 7.11 nats/s/Hz for the ITU Vehicular A, respectively (see Fig. 8). Also, 10% outage capacity is 5.67 and 6.01 nats/s/Hz for each environment (see Fig. 9). The reduction in ergodic capacity and 10% outage capacity due to spatial correlation is about 40% in both environments.

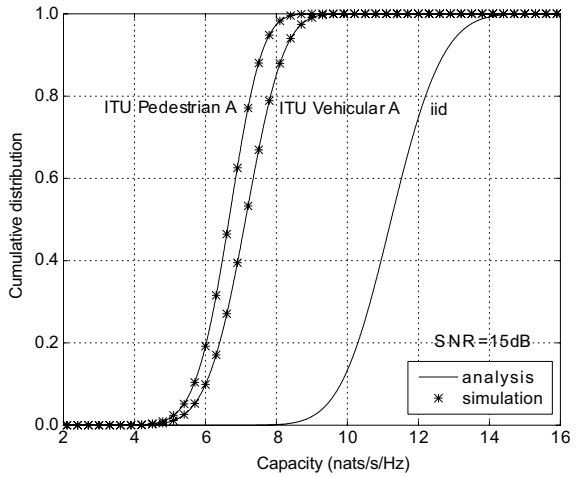

Fig. 9. CDF of the capacity for doubly correlated MIMO channels with $n_{\mathrm{T}} = n_{\mathrm{R}} = 4$ and METRA correlation matrices given in [18] for macrocell ITU Pedestrian A and Vehicular A environments. For comparison, the CDF of capacity for the 4×4 i.i.d. MIMO channel is also plotted. $\eta = 15$ dB.

VI. Conclusions

In this paper, we derived closed-form formulas for the exact capacity statistics of Rayleigh-fading MIMO channels in the presence of spatial fading correlation at both the transmitter and the receiver. In particular, we derived the *determinant* representation for the characteristic function (Theorem 1) and the *trace* representations for the cumulants (Theorems 2 and 3) of MIMO capacity as well as the mean, variance, skewness, and kurtosis (Tables II and III). These results are valid for arbitrary numbers of antennas, enabling us to calculate both the ergodic capacity and the outage capacity without any approximation and generalizing the previous results for i.i.d. and one-sided correlated MIMO channels. We also showed that in a high-SNR regime, the variance, skewness, kurtosis, and other higher-order statistics of the capacity depend only on correlation at the side with the larger number of antennas. Moreover, when the antenna topology is symmetric (i.e., $n_{\mathrm{T}} = n_{\mathrm{R}}$), these statistics are not affected by fading correlation at any side and the capacity distribution has negative skewness greater than or equal to $-12\sqrt{6} \cdot \zeta(3)/\pi^3$, where $\zeta(3) \approx 1.2020569$ is Apéry's constant, and positive (excess) kurtosis less than or equal to 2.4. This implies that the capacity distribution has an asymmetric tail extending out more to the left of the ergodic capacity and a leptokurtic shape more peaked than a Gaussian one. To illustrate our analytical results, we presented numerical examples using the correlation model based on realistic channel measurements as well as the classical exponential correlation model. These examples showed that our analytical results are in excellent agreement with Monte Carlo simulations and that a considerable decrease in capacity, due to spatial fading correlation, can be observed in realistic MIMO channels.

APPENDIX A
INTEGRAL IDENTITIES

Let us define the integrals $\mathcal{G}_n(a,b,\xi)$ and $\mathcal{J}_{n,\ell}(a,b,\xi)$ for $a,b>0$, $n\in\mathbb{N}$ and $\xi\in\mathbb{C}$ as

$$\mathcal{G}_n(a,b,\xi) \triangleq \int_0^\infty (1+ax)^{\xi-1} x^{n-1} e^{-x/b} dx \tag{32}$$

and

$$\begin{aligned}\mathcal{J}_{n,\ell}(a,b,\xi) &\triangleq \frac{\partial^\ell \mathcal{G}_n(a,b,\xi)}{\partial \xi^\ell} \\ &= \int_0^\infty (1+ax)^{\xi-1} \ln^\ell(1+ax)\, x^{n-1} e^{-x/b} dx \end{aligned}\tag{33}$$

which appear in deriving the analytical expressions for the capacity statistics in Section III.

From the integral representation of the confluent hypergeometric function $\Psi(a,b;z)$ in [26, eq. (9.211.4)] and the identity ${}_2F_0\left(a,b;-z^{-1}\right) = z^a \Psi(a,a-b+1;z)$, the integral $\mathcal{G}_n(a,b,\xi)$ can be evaluated as

$$\begin{aligned}\mathcal{G}_n(a,b,\xi) &= a^{-n}(n-1)!\,\Psi\left(n,n+\xi;\frac{1}{ab}\right) \\ &= b^n (n-1)!\,{}_2F_0(n,-\xi+1;-ab)\end{aligned}\tag{34}$$

where ${}_pF_q(a_1,a_2,\ldots,a_p;b_1,b_2,\ldots,b_q;z)$ is the generalized hypergeometric function [26, eq. (9.14.1)]. In particular, for $\xi\in\mathbb{N}$, (34) reduces to a finite sum of elementary functions as

$$\mathcal{G}_n(a,b,\xi) = b^n \sum_{k=0}^{\xi-1}\left\{\binom{\xi-1}{k}(ab)^k (n+k-1)!\right\}. \tag{35}$$

Since the derivatives of the generalized hypergeometric function with respect to its parameters are not known, in general, the integral $\mathcal{J}_{n,\ell}(a,b,\xi)$ cannot be evaluated directly from (34). However, $\mathcal{G}_n(a,b,\xi)$ for $a,b>0$, $n\in\mathbb{N}$ and $\xi\in\mathbb{C}$ can be expressed in an alternate form

$$\begin{aligned}&\mathcal{G}_n(a,b,\xi) \\ &= \frac{e^{1/(ab)}}{a^n}\sum_{k=0}^{n-1}\left\{\binom{n-1}{k}(-1)^{n-k-1}(ab)^{\xi+k}\,\Gamma\left(\xi+k,\frac{1}{ab}\right)\right\}\end{aligned}\tag{36}$$

where $\Gamma(\alpha,z)=\int_z^\infty e^{-x}x^{\alpha-1}dx$ is the complementary incomplete gamma function [26, eq. (8.350.2)]. Then, using (36) and Leibniz's identity [26, p. 21], the integral $\mathcal{J}_{n,\ell}(a,b,\xi)$ defined in (33) can be evaluated as

$$\begin{aligned}&\mathcal{J}_{n,\ell}(a,b,\xi) \\ &= \frac{e^{1/(ab)}}{a^n}\sum_{k=0}^{n-1}\Bigg[(-1)^{n-k-1}\binom{n-1}{k}(ab)^{\xi+k} \\ &\quad\times \sum_{i=0}^{\ell}\left\{\binom{\ell}{i}\ln^{\ell-i}(ab)\cdot\left[\frac{\partial^i}{\partial\alpha^i}\Gamma\left(\alpha,\frac{1}{ab}\right)\right]\Bigg|_{\alpha=\xi+k}\right\}\Bigg] \\ &= \frac{\ell!\,e^{1/(ab)}}{a^n}\sum_{k=0}^{n-1}\Bigg[(-1)^{n-k-1}\binom{n-1}{k}(ab)^{\xi+k} \\ &\quad\times G^{\ell+2,0}_{\ell+1,\ell+2}\left(\frac{1}{ab}\,\middle|\,\begin{matrix}\overbrace{1,1,\ldots,1}^{\ell+1\ 1\text{'s}}\\ \underbrace{0,0,\ldots,0}_{\ell+1\ 0\text{'s}},\xi+k\end{matrix}\right)\Bigg]\end{aligned}\tag{37}$$

where $G^{m,n}_{p,q}(\cdot)$ is the Meijer G-function [26, eq. (9.301)]. In particular, for $\ell=\xi=1$, (37) reduces to

$$\mathcal{J}_{n,1}(a,b,1) = b^n(n-1)!e^{1/(ab)}\sum_{k=0}^{n-1}(ab)^{-k}\,\Gamma\left(-k,\frac{1}{ab}\right). \tag{38}$$

APPENDIX B
PROOF OF THEOREM 1

To proceed with the proof of Theorem 1, we begin by evaluating an integral involving matrix determinants, which is a continuous analogue of the well-known results in multivariate analysis [28]. The next lemma adds a new identity to the list of the generalized Cauchy-Binet formulas derived in [7, Appendix].

Lemma 2: Suppose that f_i and g_j, $i=1,2,\ldots,m$, $j=1,2,\ldots,n$, $m\le n$, are arbitrary integrable functions over $\mathfrak{D}$. Let $\mathbf{F}(\varrho_1,\varrho_2,\ldots,\varrho_m)$ and $\mathbf{G}(\varrho_1,\varrho_2,\ldots,\varrho_m)$ be $m\times m$ and $n\times n$ matrices whose entries depend on $\varrho_1,\varrho_2,\ldots,\varrho_m$, given by

$$\{\mathbf{F}(\varrho_1,\varrho_2,\ldots,\varrho_m)\}_{i,j} = f_j(\varrho_i), \quad i,j=1,2,\ldots,m \tag{39}$$

and

$$\begin{aligned}&\{\mathbf{G}(\varrho_1,\varrho_2,\ldots,\varrho_m)\}_{i,j} \\ &= \begin{cases} c_{i,j}, & i=1,\ldots,n-m,\ j=1,\ldots,n \\ g_j(\varrho_{i-n+m}), & i=n-m+1,\ldots,n,\ j=1,\ldots,n,\end{cases}\end{aligned}\tag{40}$$

where $c_{i,j}$ are scalar constants. Then,

$$\begin{aligned}&\int_{\mathfrak{D}}\cdots\int_{\mathfrak{D}}\det\mathbf{F}(\varrho_1,\varrho_2,\ldots,\varrho_m)\det\mathbf{G}(\varrho_1,\varrho_2,\ldots,\varrho_m) \\ &\qquad\times\prod_{\ell=1}^{m}h(\varrho_\ell)\,d\varrho_1 d\varrho_2\cdots d\varrho_m \\ &= m!\det(\mathbf{\Phi})\end{aligned}\tag{41}$$

where $h(\cdot)$ is an arbitrary function and $\mathbf{\Phi}$ is the $n\times n$ matrix with (i,j)th entry $\phi_{i,j}$ given by

$$\phi_{i,j} = \begin{cases} c_{i,j}, & i=1,\ldots,n-m,\ j=1,\ldots,n \\ \int_{\mathfrak{D}} f_{i-n+m}(\varrho)\,g_j(\varrho)\,h(\varrho)\,d\varrho, \\ \qquad i=n-m+1,\ldots,n,\ j=1,\ldots,n.\end{cases}\tag{42}$$

Proof: Let

$$\mathbf{a}=(a_1,a_2,\ldots,a_n)$$

and

$$\mathbf{b}=(b_1,b_2,\ldots,b_m)$$

be the permutations of integers $1,2,\ldots,n$ and $1,2,\ldots,m$, respectively. Then, the integration of the left-hand side of (41) becomes

$$\begin{aligned}&\sum_{\mathbf{a}}\sum_{\mathbf{b}}\operatorname{sgn}(\mathbf{a})\operatorname{sgn}(\mathbf{b})\prod_{i=1}^{n-m}c_{i,a_i}\prod_{j=1}^{m}\phi_{n-m+b_j,a_{n-m+j}} \\ &= \sum_{\mathbf{a}}\operatorname{sgn}(\mathbf{a})\sum_{\mathbf{b}}\operatorname{sgn}^2(\mathbf{b})\prod_{i=1}^{n}\phi_{i,a_i} \\ &= m!\det(\mathbf{\Phi})\end{aligned}$$

where $\mathrm{sgn}(\cdot)$ denotes the sign of the permutation. □

Proof of Theorem 1: The CF of C can be written as

$$\Phi_C(\jmath\omega) = \int_{\mathbf{\Theta}=\mathbf{\Theta}^\dagger>0} \det(\mathbf{I}_{n_\mathrm{S}} + \bar{\eta}\mathbf{\Theta})^{\jmath\omega} p_{\mathbf{\Theta}}(\mathbf{\Theta})\, d\mathbf{\Theta} \quad (43)$$

where $\bar{\eta} = \eta/n_\mathrm{T}$ and from (2), the PDF of $\mathbf{\Theta} \sim \tilde{Q}_{n_\mathrm{S},n_\mathrm{L}}(\mathbf{I}_{n_\mathrm{L}}, \mathbf{\Psi}_\mathrm{S}, \mathbf{\Psi}_\mathrm{L})$ is given by

$$p_{\mathbf{\Theta}}(\mathbf{\Theta}) = \frac{1}{\tilde{\Gamma}_{n_\mathrm{S}}(n_\mathrm{L})} \det(\mathbf{\Psi}_\mathrm{S})^{-n_\mathrm{L}} \det(\mathbf{\Psi}_\mathrm{L})^{-n_\mathrm{S}} \times \det(\mathbf{\Theta})^{n_\mathrm{L}-n_\mathrm{S}} {}_0\tilde{F}_0^{(n_\mathrm{L})}\left(-\mathbf{\Psi}_\mathrm{S}^{-1}\mathbf{\Theta}, \mathbf{\Psi}_\mathrm{L}^{-1}\right). \quad (44)$$

The typical approach for the evaluation of the integral in (43) is to perform eigenvalue decomposition using the unitary transformation of $\mathbf{\Theta}$ and to exploit the knowledge of the joint eigenvalue distribution of $\mathbf{\Theta}$. However, the correlation matrix $\mathbf{\Psi}_\mathrm{S}$ in the argument of the hypergeometric function in (44) prevents the removal of the unitary matrix from its arguments after the eigenvalue decomposition, which makes it difficult to directly use the joint eigenvalue distribution of $\mathbf{\Theta}$. We alleviate this difficulty by performing two successive transformations as follows. The first transformation is given by

$$\mathbf{Z} = \mathbf{\Psi}_\mathrm{S}^{-1/2}\mathbf{\Theta}\mathbf{\Psi}_\mathrm{S}^{-1/2} \quad (45)$$

with Jacobian

$$d\mathbf{\Theta} = \det(\mathbf{\Psi}_\mathrm{S})^{n_\mathrm{S}}\, d\mathbf{Z}. \quad (46)$$

Using the fact that

$$ {}_0\tilde{F}_0^{(n_\mathrm{L})}\left(-\mathbf{\Psi}_\mathrm{S}^{-1}\mathbf{\Theta}, \mathbf{\Psi}_\mathrm{L}^{-1}\right) = {}_0\tilde{F}_0^{(n_\mathrm{L})}\left(-\mathbf{\Psi}_\mathrm{S}^{-1/2}\mathbf{\Theta}\mathbf{\Psi}_\mathrm{S}^{-1/2}, \mathbf{\Psi}_\mathrm{L}^{-1}\right), \quad (47)$$

we have

$$\Phi_C(\jmath\omega) = \frac{\det(\mathbf{\Psi}_\mathrm{L})^{-n_\mathrm{S}}}{\tilde{\Gamma}_{n_\mathrm{S}}(n_\mathrm{L})} \int_{\mathbf{Z}=\mathbf{Z}^\dagger>0} \det(\mathbf{I}_{n_\mathrm{S}} + \bar{\eta}\mathbf{\Psi}_\mathrm{S}\mathbf{Z})^{\jmath\omega} \times \det(\mathbf{Z})^{n_\mathrm{L}-n_\mathrm{S}} {}_0\tilde{F}_0^{(n_\mathrm{L})}\left(-\mathbf{Z}, \mathbf{\Psi}_\mathrm{L}^{-1}\right) d\mathbf{Z}. \quad (48)$$

Let us denote a unitary manifold of $n_\mathrm{S}\times n_\mathrm{S}$ unitary matrices with real diagonal elements by $\tilde{U}(n_\mathrm{S})$. Since $\mathbf{Z}$ is Hermitian, there exists $\mathbf{U} \in \tilde{U}(n_\mathrm{S})$ such that $\mathbf{Z} = \mathbf{U}\mathbf{D}\mathbf{U}^\dagger$ and $\mathbf{D} = \mathrm{diag}(z_1, z_2, \ldots, z_{n_\mathrm{S}})$ where $0 < z_1 \le z_2 \le \cdots \le z_{n_\mathrm{S}}$ are ordered eigenvalues of $\mathbf{Z}$. We then make the second transformation

$$\mathbf{Z} = \mathbf{U}\mathbf{D}\mathbf{U}^\dagger \quad (49)$$

with Jacobian [29, Theorem 3.1], [30, Theorem 4.4]

$$d\mathbf{Z} = \prod_{i<j}^{n_\mathrm{S}} (z_j - z_i)^2\, d\mathbf{U} d\mathbf{D}, \quad (50)$$

yielding

$$\Phi_C(\jmath\omega) = \frac{\det(\mathbf{\Psi}_\mathrm{L})^{-n_\mathrm{S}}}{\tilde{\Gamma}_{n_\mathrm{S}}(n_\mathrm{L})} \int_{\mathbf{D}} \int_{\mathbf{U}\in\tilde{U}(n_\mathrm{S})} \det(\mathbf{D})^{n_\mathrm{L}-n_\mathrm{S}} \times \det\left(\mathbf{I}_{n_\mathrm{S}} + \bar{\eta}\mathbf{\Psi}_\mathrm{S}\mathbf{U}\mathbf{D}\mathbf{U}^\dagger\right)^{\jmath\omega} {}_0\tilde{F}_0^{(n_\mathrm{L})}\left(-\mathbf{D}, \mathbf{\Psi}_\mathrm{L}^{-1}\right) \times \prod_{i<j}^{n_\mathrm{S}} (z_j - z_i)^2\, d\mathbf{U} d\mathbf{D} \quad (51)$$

where we have used the fact that the hypergeometric function with matrix arguments is invariant under unitary transformations of its arguments.[7] It is now apparent that the above two transformations enable us to remove the dependence of $\mathbf{U}$ on the hypergeometric function. Recall that the total volume of $\tilde{U}(n_\mathrm{S})$ is [30, Corollary 4.3.1]

$$\int_{\mathbf{U}\in\tilde{U}(n_\mathrm{S})} d\mathbf{U} = \frac{\pi^{n_\mathrm{S}(n_\mathrm{S}-1)}}{\tilde{\Gamma}_{n_\mathrm{S}}(n_\mathrm{S})}. \quad (52)$$

We can now carry out the integration with respect to $\mathbf{U}$ using [30, eq. (6.1.19)] and [21, eq. (52)] as

$$\begin{aligned}\int_{\mathbf{U}\in\tilde{U}(n_\mathrm{S})} &\det\left(\mathbf{I}_{n_\mathrm{S}} + \bar{\eta}\mathbf{\Psi}_\mathrm{S}\mathbf{U}\mathbf{D}\mathbf{U}^\dagger\right)^{\jmath\omega} d\mathbf{U} \\ &= \int_{\mathbf{U}\in\tilde{U}(n_\mathrm{S})} {}_1\tilde{F}_0\left(-\jmath\omega; -\bar{\eta}\mathbf{\Psi}_\mathrm{S}\mathbf{U}\mathbf{D}\mathbf{U}^\dagger\right) d\mathbf{U} \\ &= \frac{\pi^{n_\mathrm{S}(n_\mathrm{S}-1)}}{\tilde{\Gamma}_{n_\mathrm{S}}(n_\mathrm{S})} {}_1\tilde{F}_0^{(n_\mathrm{S})}\left(-\jmath\omega; \mathbf{D}, -\bar{\eta}\mathbf{\Psi}_\mathrm{S}\right). \end{aligned} \quad (53)$$

Substituting (53) into (51) gives

$$\begin{aligned}\Phi_C(\jmath\omega) = &\frac{\pi^{n_\mathrm{S}(n_\mathrm{S}-1)} \det(\mathbf{\Psi}_\mathrm{L})^{-n_\mathrm{S}}}{\tilde{\Gamma}_{n_\mathrm{S}}(n_\mathrm{L})\tilde{\Gamma}_{n_\mathrm{S}}(n_\mathrm{S})} \\ &\times \int_{0<z_1\le\cdots\le z_{n_\mathrm{S}}<\infty} \prod_{\ell=1}^{n_\mathrm{S}} z_\ell^{n_\mathrm{L}-n_\mathrm{S}} \\ &\times {}_1\tilde{F}_0^{(n_\mathrm{S})}\left(-\jmath\omega; \mathbf{D}, -\bar{\eta}\mathbf{\Psi}_\mathrm{S}\right) {}_0\tilde{F}_0^{(n_\mathrm{L})}\left(\mathbf{D}, -\mathbf{\Psi}_\mathrm{L}^{-1}\right) \\ &\times \prod_{i<j}^{n_\mathrm{S}} (z_j - z_i)^2\, dz_1 dz_2 \cdots dz_{n_\mathrm{S}}. \end{aligned} \quad (54)$$

Using the results in [23, Lemma 3] and [20], the hypergeometric functions with matrix arguments in the integrand of (54) can be expressed in terms of determinants as (55) and (56), shown at the top of the page, where $(\alpha)_n = \alpha(\alpha+1)\cdots(\alpha+n-1)$, $(\alpha)_0 = 1$, is the Pochhammer symbol and $\mathbf{\Xi}$ is the $n_\mathrm{L}\times n_\mathrm{L}$ matrix whose (i,j)th entry is given by

$$\{\mathbf{\Xi}\}_{i,j} = \begin{cases} \lambda_{\mathrm{L},j}^{i-1}, & i = 1,\ldots,n_\mathrm{L}-n_\mathrm{S},\ j = 1,\ldots,n_\mathrm{L} \\ \lambda_{\mathrm{L},j}^{n_\mathrm{L}-n_\mathrm{S}-1} \exp\left(-\frac{z_{i-n_\mathrm{L}+n_\mathrm{S}}}{\lambda_{\mathrm{L},j}}\right), & \\ \qquad i = n_\mathrm{L}-n_\mathrm{S}+1,\ldots,n_\mathrm{L},\ j = 1,\ldots,n_\mathrm{L}. \end{cases} \quad (57)$$

Combining (54)–(57) together with

$$(-1)^{n_\mathrm{S}(n_\mathrm{S}-1)/2} \prod_{\ell=1}^{n_\mathrm{S}} (-\jmath\omega - n_\mathrm{S} + 1)_{\ell-1} = \prod_{\ell=1}^{n_\mathrm{S}-1} (\jmath\omega + \ell)^\ell, \quad (58)$$

we get the CF as (59), shown at the top of the page, where the last equality follows from the fact that the integrand is symmetric in $z_1, z_2, \ldots, z_{n_\mathrm{S}}$. Finally, applying the integral-type Cauchy-Binet formula in Lemma 2 to (59) and using the identity (34) complete the proof.

[7]Note that $\prod_{i<j}^{n_\mathrm{S}}(z_j - z_i)$ is the $n_\mathrm{S}\times n_\mathrm{S}$ Vandermonde determinant of $z_1, z_2, \ldots, z_{n_\mathrm{S}}$, i.e.,

$$\prod_{i<j}^{n_\mathrm{S}} (z_j - z_i) = \det\left(\left[z_j^{i-1}\right]\right).$$

$$ {}_1\tilde{F}_0^{(n_{\rm S})}\left(-\jmath\omega;\mathbf{D},-\bar{\eta}\mathbf{\Psi}_{\rm S}\right)=\frac{\left(-\bar{\eta}\pi\right)^{-n_{\rm S}(n_{\rm S}-1)/2}\tilde{\Gamma}_{n_{\rm S}}\left(n_{\rm S}\right)\det\left[\left(1+\bar{\eta}\lambda_{{\rm S},j}z_i\right)^{\jmath\omega+n_{\rm S}-1}\right]_{i,j=1,2,\ldots,n_{\rm S}}}{\prod_{\ell=1}^{n_{\rm S}}\left(-\jmath\omega-n_{\rm S}+1\right)_{\ell-1}\prod_{i<j}^{n_{\rm S}}\left(z_j-z_i\right)\prod_{i<j}^{n_{\rm S}}\left(\lambda_{{\rm S},j}-\lambda_{{\rm S},i}\right)} \tag{55} $$

$$ {}_0\tilde{F}_0^{(n_{\rm L})}\left(\mathbf{D},-\mathbf{\Psi}_{\rm L}^{-1}\right)=\frac{\pi^{-n_{\rm S}(n_{\rm S}-1)/2}\tilde{\Gamma}_{n_{\rm S}}\left(n_{\rm L}\right)\det\left(\mathbf{\Psi}_{\rm L}\right)^{n_{\rm S}}\det\left(\mathbf{\Xi}\right)}{\prod_{\ell=1}^{n_{\rm S}}z_\ell^{n_{\rm L}-n_{\rm S}}\prod_{i<j}^{n_{\rm S}}\left(z_j-z_i\right)\prod_{i<j}^{n_{\rm L}}\left(\lambda_{{\rm L},j}-\lambda_{{\rm L},i}\right)} \tag{56} $$

$$ \begin{aligned}\Phi_C\left(\jmath\omega\right)&=\frac{\Upsilon_{n_{\rm S}}\left(\jmath\omega\right)}{K_{\rm cor}}\int_{0<z_1\leq\cdots\leq z_{n_{\rm S}}<\infty}\det\left[\left(1+\bar{\eta}\lambda_{{\rm S},j}z_i\right)^{\jmath\omega+n_{\rm S}-1}\right]_{i,j=1,2,\ldots,n_{\rm S}}\det\left(\mathbf{\Xi}\right)dz_1dz_2\cdots dz_{n_{\rm S}}\\&=\frac{\Upsilon_{n_{\rm S}}\left(\jmath\omega\right)}{n_{\rm S}!K_{\rm cor}}\int_0^\infty\cdots\int_0^\infty\det\left[\left(1+\bar{\eta}\lambda_{{\rm S},j}z_i\right)^{\jmath\omega+n_{\rm S}-1}\right]_{i,j=1,2,\ldots,n_{\rm S}}\det\left(\mathbf{\Xi}\right)dz_1dz_2\cdots dz_{n_{\rm S}}\end{aligned} \tag{59} $$

ACKNOWLEDGMENT

The authors would like to thank Irene Keliher for her careful reading of the manuscript.

REFERENCES

[1] J. H. Winters, "On the capacity of radio communication systems with diversity in Rayleigh fading environment," *IEEE J. Select. Areas Commun.*, vol. 5, no. 5, pp. 871-878, June 1987.

[2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311-335, Mar. 1998.

[3] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, no. 6, pp. 585-595, Nov./Dec. 1999.

[4] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502-513, Mar. 2000.

[5] C.-N. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 637-650, Mar. 2002.

[6] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering, and keyhole," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2636-2647, Oct. 2003.

[7] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of spatially correlated MIMO Rayleigh-fading channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2363-2371, Oct. 2003.

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

[9] M. Kang and M.-S. Alouini, "Impact of correlation on the capacity of MIMO channels," in *Proc. IEEE Int. Conf. Communications*, May 2003, pp. 2623-2627.

[10] P. J. Smith, L. M. Garth, and S. Loyka, "Exact capacity distributions for MIMO systems with small numbers of antennas," *IEEE Commun. Lett.*, vol. 7, no. 10, pp. 481-483, Oct. 2003.

[11] A. Giorgetti, P. J. Smith, M. Shafi, and M. Chiani, "MIMO capacity, level crossing rates and fades: The impact of spatial/temporal channel correlation," *J. Communications and Networks*, vol. 5, no. 2, pp. 104-115, June 2003.

[12] Ö. Oyman, R. U. Nabar, H. Bolcskei, and A. J. Paulraj, "Characterizing the statistical properties of mutual information in MIMO channels," *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2784-2795, Nov. 2003.

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

[14] C. Martin and B. Ottersten, "Asymptotic eigenvalue distributions and capacity for MIMO channels under correlated fading," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1350-1359, July 2004.

[15] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2619-2692, Oct. 1998.

[16] P. J. Smith and M. Shafi, "An approximate capacity distribution for MIMO systems," *IEEE Trans. Commun.*, vol. 52, no. 6, pp. 887-890, June 2004.

[17] J. P. Kermoal, "Measurement, modelling and performance evaluation of the MIMO radio channel," Ph.D. dissertation, Aalborg University, Aalorg, Denmark, Aug. 2002.

[18] C802.20-03/50. (2003, May) Overview of METRA Model for MBWA MIMO Channel. This document has been prepared to assist the IEEE 802.20 Working Group, IEEE 802.20 Session #2. [Online]. Available: http://www.ieee802.org/20/Contribs/C802.20-03-50.doc

[19] M. Kiessling, "Unifying analysis of ergodic MIMO capacity in correlated Rayleigh fading environments," *European Trans. Telecommun.*, vol. 16, no. 1, pp. 17-35, Jan./Feb. 2005.

[20] H. Gao and P. J. Smith, "A determinant representation for the distribution of quadratic forms in complex normal vectors," *J. Multivariate Analysis*, vol. 73, no. 2, pp. 155-165, May 2000.

[21] C. G. Khatri, "On certain distribution problems based on positive definite quadratic functions in normal vectors," *Ann. Math. Statistics*, vol. 37, no. 2, pp. 468-479, Apr. 1966.

[22] A. K. Gupta and D. K. Nagar, *Matrix Variate Distributions*. Boca Raton, FL: Chapman & Hall/CRC, 2000.

[23] C. G. Khatri, "On the moments of traces of two matrices in three situations for complex multivariate normal populations," *Sankhya, The Indian J. Statistics, Series A*, vol. 32, pp. 65-80, Mar. 1970.

[24] H. Shin, "Capacity and error exponents for multiple-input multiple-output wireless channels," Ph.D. dissertation, Seoul National University, Seoul, Korea, Aug. 2004.

[25] M. A. Golberg, "The derivative of a determinant," *American Math. Monthly*, vol. 79, no. 10, pp. 1124-1126, Dec. 1972.

[26] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 6th ed. San Diego, CA: Academic Press, 2000.

[27] M. Evans, N. Hastings, and B. Peacock, *Statistical Distributions*, 3rd ed. New York: Wiley, 2000.

[28] P. R. Krishnaiah, "Some recent developments on complex multivariate distributions," *J. Multivariate Analysis*, vol. 6, no. 1, pp. 1-30, Mar. 1976.

[29] A. Edelman, "Eigenvalues and condition numbers of random matrices," Ph.D. dissertation, M.I.T., Cambridge, MA, May 1989.

[30] A. M. Mathai, *Jacobians of Matrix Transformations and Functions of Matrix Arguments*. Singapore: World Scientific, 1997.

Hyundong Shin (S'01-M'04) received the B.S. degree in electronics engineering from Kyung Hee University, Korea, in 1999, and the M.S. and Ph.D. degrees in electrical engineering from Seoul National University, Korea, in 2001 and 2004, respectively.

From 2004 to 2006, Dr. Shin was the Postdoctoral Associate at the Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology. In 2006 he joined the faculty of the School of Electronics and Information, Kyung Hee University, Korea. His research interests include wireless communications, information and coding theory, cooperative/collaborative communications, and multiple-antenna wireless communication systems and networks.

Dr. Shin served on the Technical Program Committees for the IEEE International Conference on Communications and the IEEE International Conference on Ultra Wideband in 2006.

Moe Z. Win (S'85-M'87-SM'97-F'04) received the B.S. degree (*magna cum laude*) from Texas A&M University, College Station, in 1987 and the M.S. degree from the University of Southern California (USC), Los Angeles, in 1989, both in Electrical Engineering. As a Presidential Fellow at USC, he received both an M.S. degree in Applied Mathematics and the Ph.D. degree in Electrical Engineering in 1998.

Dr. Win is an Associate Professor at the Laboratory for Information & Decision Systems (LIDS), Massachusetts Institute of Technology. Prior to joining LIDS, he spent five years at AT&T Research Laboratories and seven years at the Jet Propulsion Laboratory. His main research interests are the application of mathematical and statistical theories to communication, detection, and estimation problems. Specific current research topics include measurement and modeling of time-varying channels, design and analysis of multiple antenna systems, ultra-wide bandwidth (UWB) communications systems, optical communications systems, and space communications systems.

Professor Win has been involved actively in organizing and chairing a number of international conferences. He served as the Technical Program Chair for the IEEE Communication Theory Symposia of ICC-2004 and Globecom-2000, as well as for the IEEE Conference on Ultra Wideband Systems and Technologies in 2002, Technical Program Vice-Chair for the IEEE International Conference on Communications in 2002, and the Tutorial Chair for the IEEE Semiannual International Vehicular Technology Conference in Fall 2001. He is the current chair and past secretary (2002-2004) for the Radio Communications Committee of the IEEE Communications Society. Dr. Win currently serves as Area Editor for Modulation and Signal Design and Editor for Wideband Wireless and Diversity, both for *IEEE Transactions on Communications*. He served as the Editor for Equalization and Diversity from July 1998 to June 2003 for the *IEEE Transactions on Communications*, and as a Guest-Editor for the 2002 *IEEE Journal on Selected Areas in Communications* (Special Issue on Ultra-Wideband Radio in Multiaccess Wireless Communications). Professor Win received the International Telecommunications Innovation Award from Korea Electronics Technology Institute in 2002, a Young Investigator Award from the Office of Naval Research in 2003, and the IEEE Antennas and Propagation Society Sergei A. Schelkunoff Transactions Prize Paper Award in 2003. In 2004, Dr. Win was named Young Aerospace Engineer of the Year by AIAA, and garnered the Fulbright Foundation Senior Scholar Lecturing and Research Fellowship, the Institute of Advanced Study Natural Sciences and Technology Fellowship, the Outstanding International Collaboration Award from the Industrial Technology Research Institute of Taiwan, and the Presidential Early Career Award for Scientists and Engineers from the White House. He was honored with the 2006 IEEE Eric E. Sumner Award "for pioneering contributions to ultra-wide band communications science and technology." Professor Win is an IEEE Distinguished Lecturer and elected Fellow of the IEEE, cited "for contributions to wideband wireless transmission."

Jae Hong Lee (M'86-SM'03) received B.S. and M.S. degrees in electronics engineering from Seoul National University (SNU), Korea, in 1976 and 1978, respectively, and a Ph.D. degree in electrical engineering from the University of Michigan, Ann Arbor, in 1986.

From 1978 to 1981, he was with the faculty of the Republic of Korea Naval Academy, Jinhae. He was a member of technical staff at AT&T Bell Laboratories, Whippany, New Jersey, from 1991 to 1992. In 1987, he joined the faculty of SNU where currently he is a professor in the School of Electrical Engineering. His current research interests include communication and coding theory, space–time code, multiple-input multiple-output (MIMO), and orthogonal frequency-division multiplexing (OFDM), and their application to wireless communications. He is a Vice President of the Institute of Electronics Engineers of Korea and the Korea Society of Broadcasting Engineers.

Dr. Lee is a member of the Tau Beta Pi.

Marco Chiani (M'94–SM'02) was born in Rimini, Italy, in April 1964. He received the Dr. Ing. degree (magna cum laude) in Electronic Engineering and the Ph.D. degree in Electronic and Computer Science from the University of Bologna in 1989 and 1993, respectively. Dr. Chiani is a Full Professor at the II Engineering Faculty, University of Bologna, Italy, where he is the Chair in Telecommunication. During the summer of 2001 he was a Visiting Scientist at AT&T Research Laboratories in Middletown, NJ. He presently holds a Research Affiliate appointment at the Massachusetts Institute of Technology (MIT).

Prof. Chiani's research interests include wireless communication systems, MIMO systems, wireless multimedia, low density parity check codes (LDPCC) and UWB. Under the European research program PROMETHEUS he has worked on short-range millimeter wave communication systems for Advanced Road Transport Telematics. He is also leading the research unit of CNIT/University of Bologna on Joint Source and Channel Coding for wireless video (IST FP6, project Phoenix) and is a consultant to the European Space Agency (ESA-ESOC) for the design and evaluation of error correcting codes based on LDPCC for space CCSDS applications.

Prof. Chiani is the past chair (2002-2004) of the Radio Communications Committee of the IEEE Communication Society and the current Editor of Wireless Communication for the *IEEE Transactions on Communications*.