Capacity of Multiple-Antenna Fading Channels: Spatial Fading Correlation, Double Scattering, and Keyhole

Hyundong Shin, *Student Member, IEEE,* and Jae Hong Lee, *Member, IEEE*

***Abstract*—The capacity of multiple-input multiple-output (MIMO) wireless channels is limited by both the spatial fading correlation and rank deficiency of the channel. While spatial fading correlation reduces the diversity gains, rank deficiency due to double scattering or keyhole effects decreases the spatial multiplexing gains of multiple-antenna channels. In this paper, taking into account realistic propagation environments in the presence of spatial fading correlation, double scattering, and keyhole effects, we analyze the ergodic (or mean) MIMO capacity for an arbitrary finite number of transmit and receive antennas. We assume that the channel is unknown at the transmitter and perfectly known at the receiver so that equal power is allocated to each of the transmit antennas. Using some statistical properties of complex random matrices such as Gaussian matrices, Wishart matrices, and quadratic forms in the Gaussian matrix, we present a closed-form expression for the ergodic capacity of independent Rayleigh-fading MIMO channels and a tight upper bound for spatially correlated/double scattering MIMO channels. We also derive a closed-form capacity formula for keyhole MIMO channels. This analytic formula explicitly shows that the use of multiple antennas in keyhole channels only offers the diversity advantage, but provides no spatial multiplexing gains. Numerical results demonstrate the accuracy of our analytical expressions and the tightness of upper bounds.**

***Index Terms*—Channel capacity, distributions of random matrices, double scattering, keyhole, multiple-input multiple-output (MIMO) systems, multiple antennas, spatial fading correlation.**

I. INTRODUCTION

MULTIPLE-input multiple-output (MIMO) communication systems using multiple-antenna arrays at both the transmitter and the receiver have drawn considerable attention in response to the increasing requirements on high spectral efficiency and reliability in wireless communications [1]–[10]. Recent seminal work in [1] and [2] has shown that the use of multiple antennas at both ends significantly increases the information-theoretic capacity far beyond that of single-antenna systems in rich scattering propagation environments. As the number of antennas at both the transmitter and the receiver gets larger, the capacity increases linearly with the minimum of the number of transmit and receive antennas for fixed power and bandwidth, assuming independent and identically distributed (i.i.d.) Rayleigh fading between antenna pairs [1]–[3].

Manuscript received October 29, 2002; revised May 4, 2003. This work was supported in part by the National Research Laboratory (NRL) Program and the Brain Korea 21 Project, Korea.

The authors are with the School of Electrical Engineering, Seoul National University, Seoul 151-742, Korea (e-mail: shd71@snu.ac.kr).

Communicated by T. L. Marzetta, Guest Editor.
Digital Object Identifier 10.1109/TIT.2003.817439

An important open problem in MIMO communication theory is to obtain closed-form analytic formulas for the capacity or mutual information of wireless MIMO channels. However, it is a mathematically challenging task in that calculations of the MIMO capacity require taking expectations with respect to a random channel matrix rather than a scalar random variable (RV) for the single-antenna case. In random matrix theory [28]–[31], it is well known that the eigenvalues of a large class of random matrix ensembles have fewer random fluctuations as the matrix dimension gets larger—that is, the random distribution of eigenvalues converges to a deterministic limiting distribution for a large matrix size. Another useful result of the random matrix theory is a central limit theorem for random determinants [29], [32], which states that the distribution of the random MIMO capacity is asymptotically Gaussian as the number of antennas tends to infinity with a certain limiting ratio between the numbers of transmit and receive antennas. These results of the random matrix theory were applied for the case of uncorrelated channels in [1] and [10]–[13], and for spatially correlated channels in [17] and [18]. Although this asymptotic analysis is only an approximation to the case of finite matrix size, it circumvents the difficult problem in analytical calculation of the MIMO capacity and provides important insights into impacts of the use of multiple antennas on the capacity behavior.

For the finite number of transmit and receive antennas, Telatar [1] derived the analytical expression for the ergodic (or mean) capacity of i.i.d. Rayleigh flat-fading MIMO channels by using the eigenvalue distribution of the Wishart matrix in integral form involving the Laguerre polynomials. In [12], Smith and Shafi further derived the variance of capacity by extending the analysis in [1] and obtained the complementary cumulative distribution function of the capacity using the Gaussian approximation to random MIMO capacity. Similar results are also found in [13] in which the density function of a random mutual information for i.i.d. MIMO channels was derived in the form of the inverse Laplace transform and the same Gaussian approximation result as in [12] was presented.

In realistic propagation environments, rank deficiency of a channel matrix due to *pinhole* or *keyhole* effects [19]–[21] may severely degrade the capacity of MIMO channels as well as spatial fading correlation [14]–[18]. While spatial correlation reduces diversity advantages, rank deficiency of the channel decreases a spatial multiplexing ability, i.e., the slope of a capacity curve over a signal-to-noise ratio (SNR). Recently, Gesbert *et al.* [19] introduced a double scattering MIMO channel model that includes both the fading correlation and rank deficiency, and pointed out the existence of pinhole channels that exhibit uncorrelated spatial fading between antennas but still have a poor rank property. Also, in [20] and [21], the occurrence of a rank-deficient channel, called a keyhole channel, has been proposed and demonstrated through physical examples. In the presence of the keyhole, the channel has only a single degree of freedom although the spatial fading is uncorrelated, and each entry of the channel matrix is a product of two complex Gaussian RVs, in contrast to the complex Gaussian normally assumed in wireless channels. In fact, keyhole channels may be viewed as a special case of double scattering MIMO channels. These degenerate channels significantly deviate from the idealistic capacity behavior of i.i.d. channels and are of interest because of recent validation through physical measurements.

In this paper, taking into account realistic propagation environments such as spatial fading correlation, double scattering, and keyhole phenomena, we provide analytical expressions for the ergodic capacity of MIMO channels with finite transmit and receive antennas. In particular, we derive a closed-form expression for the capacity of i.i.d. Rayleigh-fading MIMO channels. In contrast to Telatar's integral expression, this capacity formula is in terms of a finite sum of the well-known special functions (exponential integral functions, or incomplete gamma functions) and can be calculated without explicit numerical integration. For spatially correlated and double scattering MIMO channels, we develop tight upper bounds on the capacity by using Jensen's inequality and elementary properties of determinants, such as the principal minor determinant expansion for the characteristic polynomial of a matrix and the Binet–Cauchy formula for the determinant of a product matrix.[1] Finally, for keyhole MIMO channels we provide a closed-form solution for the capacity and show that increasing the number of antennas serves only to eliminate the effect of fading, but provides no further benefits (e.g., spatial multiplexing gains).

The rest of the paper is organized as follows. Section II gives, for reference, the definitions of the complex Gaussian matrix, Wishart matrix, and the positive-definite quadratic form in the complex Gaussian matrix, and presents some new results concerning expectations of certain (logarithmic) determinantal forms of them with a finite matrix size. Using some of the results in Section II, we derive a closed-form expression of the ergodic MIMO capacity for the i.i.d. case and upper bounds for spatially correlated and double scattering cases in Section III. We, finally, derive a closed-form capacity formula for keyhole MIMO channels in Section III-D. Section IV concludes the paper.

We shall use the following notations in this paper. The superscripts $*$, T, and $\dagger$ denote the complex conjugate, transpose, and transpose conjugate, respectively. $\otimes$ and $\boldsymbol{I}_n$ denote the Kronecker product of matrices and an $n \times n$ identity matrix. $\mathrm{vec}\,(\boldsymbol{A})$ represents a vector formed by stacking all the columns of $\boldsymbol{A}$ into a column vector. $\mathrm{tr}\,(\boldsymbol{A})$ denotes a trace operator of a square matrix $\boldsymbol{A}$ and $\mathrm{etr}\,(\boldsymbol{A}) = \exp\{\mathrm{tr}\,(\boldsymbol{A})\}$. By $\boldsymbol{A} > 0$ we denote that $\boldsymbol{A}$ is positive definite.

[1]This approach was first introduced by Grant in [10] to obtain the upper bound on the ergodic MIMO capacity for the i.i.d. Rayleigh-fading case.

II. Review and Some Results on Random Matrices

Much attention has been given over the years to the distribution theory of random matrices because they appear in many applications in statistics and communication theory. The distribution of the covariance matrix of samples from a multivariate Gaussian distribution, which is known as the Wishart distribution, was perhaps the beginning of a theory of distributions of random matrices [36], [28]. In this section, by focusing on the complex cases, we briefly review the definitions and distributions of Gaussian matrices, Wishart matrices, and positive-definite quadratic forms in the Gaussian matrix—which generalize the univariate Gaussian RV, central chi-square RV, and central positive-definite quadratic form in the Gaussian RV, respectively—and derive some new results on them, which are used to calculate the ergodic capacity of MIMO channels in the next section.

In deriving the statistics of a certain random matrix, the Jacobians of matrix transformations are needed and functions of a matrix argument are also widely used in calculations involving matrix-variate distributions. The introductions to them are provided in [28], [33], [34], [37], [38], and [40]. In particular, [33] has dealt with a wide class of matrix-variate distributions. Although this textbook concentrates on matrices of real random variates, one can easily develop the corresponding complex cases.

A. Complex Gaussian Matrices

Let us denote the complex p-variate Gaussian distribution with mean vector $\boldsymbol{m} \in \mathbb{C}^p$ and covariance matrix $\boldsymbol{\Sigma} \in \mathbb{C}^{p\times p}$ by $\tilde{\mathcal{N}}_p\,(\boldsymbol{m}, \boldsymbol{\Sigma})$.

Definition II.1 ([33, Definition 2.2.1]): A random matrix $\boldsymbol{X} \in \mathbb{C}^{p\times q}$ is said to have a matrix-variate Gaussian distribution with mean matrix $\boldsymbol{M} \in \mathbb{C}^{p\times q}$ and covariance matrix $\boldsymbol{\Sigma} \otimes \boldsymbol{\Psi}$ where $\boldsymbol{\Sigma} \in \mathbb{C}^{p\times p} > 0$ and $\boldsymbol{\Psi} \in \mathbb{C}^{q\times q} > 0$ are Hermitian, if

$$\mathrm{vec}(\boldsymbol{X}^\dagger) \sim \tilde{\mathcal{N}}_{pq}\big(\mathrm{vec}\big(\boldsymbol{M}^\dagger\big), \boldsymbol{\Sigma} \otimes \boldsymbol{\Psi}\big).$$

We use the notation $\boldsymbol{X} \sim \tilde{\mathcal{N}}_{p,\,q}\,(\boldsymbol{M}, \boldsymbol{\Sigma} \otimes \boldsymbol{\Psi})$ to denote that $\boldsymbol{X}$ is Gaussian distributed. The density function of $\boldsymbol{X}$ is given by

$$p_{\boldsymbol{X}}\,(\boldsymbol{X}) = \pi^{-pq} \det(\boldsymbol{\Sigma})^{-q} \det(\boldsymbol{\Psi})^{-p} \cdot \mathrm{etr}\left\{-\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}-\boldsymbol{M})\,\boldsymbol{\Psi}^{-1}(\boldsymbol{X}-\boldsymbol{M})^\dagger\right\}. \quad (1)$$

In the following lemma, we give a preliminary result on the complex Gaussian matrix.

Lemma II.1: If $\boldsymbol{X} \sim \tilde{\mathcal{N}}_{p,\,q}\,(0, \boldsymbol{I}_p \otimes \boldsymbol{I}_q)$, then we have for $k \le p$ and $k \le q$

$$\mathbb{E}\left[\det\left(\boldsymbol{X}^{i_1,\,i_2,\,\ldots,\,i_k}_{j_1,\,j_2,\,\ldots,\,j_k}\right)\det\left(\left(\boldsymbol{X}^\dagger\right)^{u_1,\,u_2,\,\ldots,\,u_k}_{v_1,\,v_2,\,\ldots,\,v_k}\right)\right] = \begin{cases} k!, & \text{if } i_1 = v_1,\ j_1 = u_1,\ \ldots,\ i_k = v_k,\ j_k = u_k \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where $1 \le i_1 < i_2 < \cdots < i_k \le p$, $1 \le j_1 < j_2 < \cdots < j_k \le q$, $1 \le u_1 < u_2 < \cdots < u_k \le q$, $1 \le v_1 < v_2 < \cdots < v_k \le p$, and $\det\left(\boldsymbol{X}^{i_1, i_2, \ldots, i_k}_{j_1, j_2, \ldots, j_k}\right)$ is a minor determinant of $\boldsymbol{X}$, i.e., a determinant of the $k \times k$ matrix lying in the $i_1, i_2, \ldots, i_k$ rows and in the $j_1, j_2, \ldots, j_k$ columns of $\boldsymbol{X}$ [25].

Proof: From the definition of the determinant [25], [26], we get

$$\begin{aligned}
&\mathbb{E}\left[\det\left(\boldsymbol{X}^{i_1, i_2, \ldots, i_k}_{j_1, j_2, \ldots, j_k}\right)\det((\boldsymbol{X}^\dagger)^{u_1, u_2, \ldots, u_k}_{v_1, v_2, \ldots, v_k})\right]\\
&= \mathbb{E}\left[\sum_{\boldsymbol{a}=(a_1, a_2, \ldots, a_k)} (-1)^{\tau(\boldsymbol{a})} x_{i_1a_1}x_{i_2a_2}\cdots x_{i_ka_k}\right.\\
&\qquad \left.\cdot \sum_{\boldsymbol{b}=(b_1, b_2, \ldots, b_k)} (-1)^{\tau(\boldsymbol{b})} x^*_{v_1b_1}x^*_{v_2b_2}\cdots x^*_{v_kb_k}\right]\\
&= \sum_{\boldsymbol{a}=(a_1, a_2, \ldots, a_k)}\sum_{\boldsymbol{b}=(b_1, b_2, \ldots, b_k)} \left\{(-1)^{\tau(\boldsymbol{a})+\tau(\boldsymbol{b})}\right.\\
&\qquad \left.\cdot\, \mathbb{E}\left[x_{i_1a_1}x_{i_2a_2}\cdots x_{i_ka_k}x^*_{v_1b_1}x^*_{v_2b_2}\cdots x^*_{v_kb_k}\right]\right\}\\
&= \begin{cases} k!, & \text{if } i_1 = v_1,\, j_1 = u_1, \ldots, i_k = v_k,\, j_k = u_k\\ 0, & \text{otherwise}\end{cases}
\end{aligned}$$

where x_{ij} is the (i, j)th entry of $\boldsymbol{X}$ and $\boldsymbol{a} = (a_1, a_2, \ldots, a_k)$ varies over all $k!$ permutations of the numbers $j_1, j_2, \ldots, j_k$, and $\tau(\boldsymbol{a})$ indicates the number of inversions in the permutation $\boldsymbol{a}$ from the normal order $j_1, j_2, \ldots, j_k$. Similarly, $\boldsymbol{b} = (b_1, b_2, \ldots, b_k)$ ranges over all $k!$ permutations of the numbers $u_1, u_2, \ldots, u_k$, and $\tau(\boldsymbol{b})$ is the number of inversions of $b_1, b_2, \ldots, b_k$ from the normal order $u_1, u_2, \ldots, u_k$. The expected terms in summation become nonzero only when $i_n = v_n$, $j_n = u_n$ for all $n = 1, 2, \ldots, k$, and the permutations $\boldsymbol{a}$ and $\boldsymbol{b}$ have the same order—otherwise, we always get terms multiplied by other independent zero-mean Gaussian entries—and the corresponding expected values are equal to one because all entries of $\boldsymbol{X}$ are independent with unit variance. There are $k!$ such nonzero terms. The last step follows immediately from these observations. □

B. Complex Wishart Matrices

Wishart distributions, first obtained by Fisher [35] in the bivariate case and generalized by Wishart [36] using a geometrical argument, are of great interest in multivariate statistical analysis, arising naturally in applied research and as a basis for theoretical models (see [28], [30], [33] and references therein).

Definition II.2 ([33, Definition 3.2.1]): A random Hermitian positive-definite matrix $\boldsymbol{Y} \in \mathbb{C}^{p\times p}$ is said to have a complex central Wishart distribution with parameters p, q, and $\boldsymbol{\Sigma} \in \mathbb{C}^{p\times p} > 0$, denoted by $\boldsymbol{Y} \sim \tilde{\mathcal{W}}_p(q, \boldsymbol{\Sigma})$, $p \le q$, if its density function is given by [38], [10]

$$p_{\boldsymbol{Y}}(\boldsymbol{Y}) = \frac{1}{\tilde{\Gamma}_p(q)}\det(\boldsymbol{\Sigma})^{-q}\det(\boldsymbol{Y})^{q-p}\operatorname{etr}\left(-\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}\right), \quad \boldsymbol{Y} > 0 \qquad (3)$$

where $\tilde{\Gamma}_p(\alpha) = \pi^{p(p-1)/2}\prod_{i=0}^{p-1}\Gamma(\alpha - i)$, $\operatorname{Re}(\alpha) > p - 1$, is the complex multivariate gamma function [38, eq. (83)] and $\Gamma(\cdot)$ is the gamma function.

If $\boldsymbol{X} \sim \tilde{\mathcal{N}}_{p,q}(0, \boldsymbol{\Sigma}\otimes\boldsymbol{I}_q)$, $p \le q$, and $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{X}^\dagger$, then $\boldsymbol{Y}$ is complex central Wishart distributed, i.e., $\boldsymbol{Y} \sim \tilde{\mathcal{W}}_p(q, \boldsymbol{\Sigma})$. The next theorem leads to establish a closed-form expression for the ergodic capacity of i.i.d. Rayleigh-fading MIMO channels in Section III-A.

Theorem II.1: If $\boldsymbol{Y} \sim \tilde{\mathcal{W}}_p(q, \boldsymbol{I}_p)$ and μ is an arbitrary positive real-valued constant, then we have

$$\begin{aligned}
&\mathbb{E}[\ln\det(\boldsymbol{I}_p + \mu\boldsymbol{Y})]\\
&\quad = e^{1/\mu}\sum_{i=0}^{p-1}\sum_{j=0}^{i}\sum_{l=0}^{2j}\left\{\frac{(-1)^l(2j)!(q-p+l)!}{2^{2i-l}j!l!(q-p+j)!}\right.\\
&\qquad \left.\cdot\binom{2i-2j}{i-j}\binom{2j+2q-2p}{2j-l}\sum_{k=0}^{q-p+l}E_{k+1}\left(\frac{1}{\mu}\right)\right\} \qquad (4)
\end{aligned}$$

where

$$E_n(z) = \int_1^\infty e^{-zx}x^{-n}dx, \qquad n = 0, 1, 2, \ldots, \operatorname{Re}(z) > 0$$

is the exponential integral function of order n [45].

Proof: See Appendix A. □

C. Positive-Definite Quadratic Forms in Complex Gaussian Matrices

Khatri [40] has first given a representation for the density function of matrix quadratic forms in the complex Gaussian matrix and different types of representation have been developed in [41]–[43]. It reduces to the Wishart density under certain conditions.

Definition II.3 ([33, Theorem 7.2.1]): Let

$$\boldsymbol{X} \sim \tilde{\mathcal{N}}_{p,q}(0, \boldsymbol{\Sigma}\otimes\boldsymbol{\Psi}), \qquad p \le q.$$

Then a positive-definite quadratic form in $\boldsymbol{X}$ associated with a Hermitian matrix $\boldsymbol{A} \in \mathbb{C}^{q\times q} > 0$, denoted by $\boldsymbol{Y} \sim \tilde{\mathcal{Q}}_{p,q}(\boldsymbol{A}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$, is defined as

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{A}\boldsymbol{X}^\dagger.$$

The density function of $\boldsymbol{Y} \sim \tilde{\mathcal{Q}}_{p,q}(\boldsymbol{A}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$ is given by [40]

$$\begin{aligned}
p_{\boldsymbol{Y}}\boldsymbol{Y} &= \frac{1}{\Gamma_p(q)}\det(\boldsymbol{\Sigma})^{-q}\det(\boldsymbol{A}\boldsymbol{\Psi})^{-p}\det(\boldsymbol{Y})^{q-p}\\
&\quad \cdot\operatorname{etr}(-\varphi^{-1}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y})\,{}_0F_0^{(q)}(;\,;\boldsymbol{B}, \varphi^{-1}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}), \quad \boldsymbol{Y} > 0 \qquad (5)
\end{aligned}$$

where $\boldsymbol{B} = \boldsymbol{I}_q - \varphi\boldsymbol{A}^{-1/2}\boldsymbol{\Psi}^{-1}\boldsymbol{A}^{-1/2}$, $\varphi > 0$ is an arbitrary constant and ${}_m\tilde{F}_n^{(q)}(\cdot)$ is the hypergeometric function of two Hermitian matrices defined by (51). Note that if $\boldsymbol{A}\boldsymbol{\Psi} = \boldsymbol{I}_q$, the density (5) reduces to the Wishart density $\tilde{\mathcal{W}}_p(q, \boldsymbol{\Sigma})$. Recently, the determinant representation for (5) has been derived in [43] to settle the computational problem of hypergeometric functions of matrix arguments.

The following two theorems are the generalizations of [10, Lemma A.1] and [10, Theorem A.4] to quadratic forms in the complex Gaussian matrix (of course, Wishart matrices with arbitrary covariance structure), respectively.

Theorem II.2 (Moments of Generalized Variance): If $\boldsymbol{Y} \sim \tilde{\mathcal{Q}}_{p,q}(\boldsymbol{A}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$, then the νth moment of the generalized variance $\det(\boldsymbol{Y})$ is given by

$$\mathbb{E}[\det(\boldsymbol{Y})^\nu] = \varphi^{p(q+\nu)} \det(\boldsymbol{\Sigma})^\nu \det(\boldsymbol{A\Psi})^{-p} \cdot {}_2\tilde{F}_1(q+\nu, p; q; \boldsymbol{B}) \prod_{i=0}^{p-1} (q-i)_\nu \quad (6)$$

where $\varphi > 0$, $(a)_n = a(a+1)\cdots(a+n-1)$, $(a)_0 = 1$, $a \neq 0$, is the Pochhammer symbol and ${}_2\tilde{F}_1(\cdot)$ is the Gauss hypergeometric function of a Hermitian matrix defined by (50).

In particular, if $q = p$, then

$$\mathbb{E}[\det(\boldsymbol{Y})^\nu] = \det(\boldsymbol{\Sigma A \Psi})^\nu \prod_{i=0}^{p-1} (p-i)_\nu. \quad (7)$$

Proof: See Appendix B. □

Theorem II.3: If $\boldsymbol{Y} \sim \tilde{\mathcal{Q}}_{p,q}(\boldsymbol{A}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$ and μ is an arbitrary positive real-valued constant, then we have

$$\mathbb{E}[\det(\boldsymbol{I}_p + \mu\boldsymbol{Y})] = \sum_{k=0}^{p} \left\{ \mu^k k! \sum_{1 \le i_1 < i_2 < \cdots < i_k \le p} \det\left(\boldsymbol{\Sigma}^{i_1, i_2, \ldots, i_k}_{i_1, i_2, \ldots, i_k}\right) \cdot \sum_{1 \le u_1 < u_2 < \cdots < u_k \le q} \det\left(\boldsymbol{Q}^{u_1, u_2, \ldots, u_k}_{u_1, u_2, \ldots, u_k}\right) \right\} \quad (8)$$

where $\boldsymbol{Q} = \boldsymbol{\Psi}^{1/2} \boldsymbol{A} \boldsymbol{\Psi}^{1/2}$.

Proof: The proof of this theorem depends on the following elementary properties of determinants. For any $p \times p$ matrix $\boldsymbol{A}$, $\det(\boldsymbol{I}_p + \boldsymbol{A})$ can be written as a sum of all the principal minor determinants from the theorem of principal minor determinant expansion for the characteristic polynomial of a matrix [10], [25], i.e.,

$$\det(\boldsymbol{I}_p + \boldsymbol{A}) = \sum_{k=0}^{p} \sum_{1 \le i_1 < i_2 < \cdots < i_k \le p} \det\left(\boldsymbol{A}^{i_1, i_2, \ldots, i_k}_{i_1, i_2, \ldots, i_k}\right) \quad (9)$$

where the 0-rowed principal minor determinant (i.e., $k = 0$) is assumed to be 1. Moreover, let $\boldsymbol{A} = \boldsymbol{BCD}$ where $\boldsymbol{B} \in \mathbb{C}^{p \times b}$, $\boldsymbol{C} \in \mathbb{C}^{b \times c}$, and $\boldsymbol{D} \in \mathbb{C}^{c \times p}$, then from the Binet–Cauchy formula for the determinant of a product matrix [25], [26], we can write the k-rowed principal minor determinant of $\boldsymbol{A}$ as

$$\det\left(\boldsymbol{A}^{i_1, i_2, \ldots, i_k}_{i_1, i_2, \ldots, i_k}\right) = \sum_{1 \le j_1 < j_2 < \cdots < j_k \le b} \sum_{1 \le u_1 < u_2 < \cdots < u_k \le c} \left\{ \det\left(\boldsymbol{B}^{i_1, i_2, \ldots, i_k}_{j_1, j_2, \ldots, j_k}\right) \cdot \det\left(\boldsymbol{C}^{j_1, j_2, \ldots, j_k}_{u_1, u_2, \ldots, u_k}\right) \det\left(\boldsymbol{D}^{u_1, u_2, \ldots, u_k}_{i_1, i_2, \ldots, i_k}\right) \right\} \quad (10)$$

where if $k > b$ or $k > c$, then the k-rowed principal minor determinant of the product matrix $\boldsymbol{A}$ becomes zero. Note that by easy induction, the property (10) can be extended to the product of any number of matrices.

Let $\boldsymbol{X} \sim \tilde{\mathcal{N}}_{p,q}(0, \boldsymbol{I}_p \otimes \boldsymbol{I}_q)$, then $\boldsymbol{Y} = \boldsymbol{\Sigma}^{1/2} \boldsymbol{XQX}^\dagger \boldsymbol{\Sigma}^{1/2}$. Using (9), (10), and Lemma II.1, we have the result (11) shown at the bottom of the page. □

We remark that if $\boldsymbol{\Sigma} = \boldsymbol{I}_p$ and $\boldsymbol{A\Psi} = \boldsymbol{I}_q$ in Theorem II.3, we have

$$\sum_{1 \le i_1 < i_2 < \cdots < i_k \le p} \det\left(\boldsymbol{\Sigma}^{i_1, i_2, \ldots, i_k}_{i_1, i_2, \ldots, i_k}\right) = \binom{p}{k}$$

and

$$\sum_{1 \le u_1 < u_2 < \cdots < u_k \le q} \det\left(\boldsymbol{Q}^{u_1, u_2, \ldots, u_k}_{u_1, u_2, \ldots, u_k}\right) = \binom{q}{k}$$

which yield the following result on the Wishart matrix $\boldsymbol{Y} \sim \tilde{\mathcal{W}}_p(q, \boldsymbol{I}_p)$:

$$\mathbb{E}[\det(\boldsymbol{I}_p + \mu\boldsymbol{Y})] = \sum_{k=0}^{p} \mu^k k! \binom{p}{k} \binom{q}{k} = p! \mu^p L_p^{q-p}(-1/\mu) \quad (12)$$

where the last step follows from the expression of the Laguerre polynomial in (40). This result was first derived by Grant [10, Theorem A.4]. Theorem II.3 will be applied to obtain an upper

$$\begin{aligned}
\mathbb{E}[\det(\boldsymbol{I}_p + \mu\boldsymbol{Y})] &= \mathbb{E}\left[\det(\boldsymbol{I}_p + \mu\boldsymbol{\Sigma XQX}^\dagger)\right] \\
&= \mathbb{E}\left[\sum_{k=0}^{p} \sum_{1 \le i_1 < i_2 < \cdots < i_k \le p} \det\left((\mu\boldsymbol{\Sigma XQX}^\dagger)^{i_1, i_2, \ldots, i_k}_{i_1, i_2, \ldots, i_k}\right)\right] \\
&= \sum_{k=0}^{p} \mu^k \sum_{1 \le i_1 < i_2 < \cdots < i_k \le p} \sum_{1 \le j_1 < j_2 < \cdots < j_k \le p} \sum_{1 \le u_1 < u_2 < \cdots < u_k \le q} \sum_{1 \le v_1 < v_2 < \cdots < v_k \le q} \det\Bigg\{ \left(\boldsymbol{\Sigma}^{i_1, i_2, \ldots, i_k}_{j_1, j_2, \ldots, j_k}\right) \det\left(\boldsymbol{Q}^{u_1, u_2, \ldots, u_k}_{v_1, v_2, \ldots, v_k}\right) \\
&\qquad \cdot \mathbb{E}\left[\det\left(\boldsymbol{X}^{j_1, j_2, \ldots, j_k}_{u_1, u_2, \ldots, u_k}\right) \det\left((\boldsymbol{X}^\dagger)^{v_1, v_2, \ldots, v_k}_{i_1, i_2, \ldots, i_k}\right)\right] \Bigg\} \\
&= \sum_{k=0}^{p} \mu^k \sum_{1 \le i_1 < i_2 < \cdots < i_k \le p} \sum_{1 \le u_1 < u_2 < \cdots < u_k \le q} \det\Bigg\{ \left(\boldsymbol{\Sigma}^{i_1, i_2, \ldots, i_k}_{i_1, i_2, \ldots, i_k}\right) \det\left(\boldsymbol{Q}^{u_1, u_2, \ldots, u_k}_{u_1, u_2, \ldots, u_k}\right) \\
&\qquad \cdot \mathbb{E}\left[\det\left(\boldsymbol{X}^{i_1, i_2, \ldots, i_k}_{u_1, u_2, \ldots, u_k}\right) \det\left((\boldsymbol{X}^\dagger)^{u_1, u_2, \ldots, u_k}_{i_1, i_2, \ldots, i_k}\right)\right] \Bigg\} \\
&= \sum_{k=0}^{p} \mu^k k! \sum_{1 \le i_1 < i_2 < \cdots < i_k \le p} \det\left(\boldsymbol{\Sigma}^{i_1, i_2, \ldots, i_k}_{i, i_2, \ldots, i_k}\right) \sum_{1 \le u_1 < u_2 < \cdots < u_k \le q} \det\left(\boldsymbol{Q}^{u_1, u_2, \ldots, u_k}_{u_1, u_2, \ldots, u_k}\right). \quad (11)
\end{aligned}$$

bound on the ergodic capacity of MIMO channels in the next section.

III. Capacity of MIMO Wireless Channels

We consider a point-to-point communication link with t transmit and r receive antennas. In what follows, we refer to $n = \max\{t, r\}$ and $m = \min\{t, r\}$ and restrict our analysis to the frequency flat-fading case. We assume that the channel is perfectly known to the receiver but unknown to the transmitter. The total power of the complex transmitted signal vector $\boldsymbol{x} \in \mathbb{C}^t$ is constrained to $\mathcal{P}$ regardless of the number of antennas, i.e.,

$$\mathbb{E}\left[\boldsymbol{x}^\dagger \boldsymbol{x}\right] \le \mathcal{P}. \tag{13}$$

At each symbol interval, the received signal vector $\boldsymbol{y} \in \mathbb{C}^r$ is given by

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n} \tag{14}$$

where $\boldsymbol{H} \in \mathbb{C}^{r\times t}$ is a random channel matrix and $\boldsymbol{n} \sim \tilde{\mathcal{N}}_r\left(0, \boldsymbol{\Sigma}_n^2 \boldsymbol{I}_r\right)$ is a complex r-dimensional additive white Gaussian noise (AWGN) vector. The entries $h_{ij}, i=1, 2, \ldots, r$ and $j=1, 2, \ldots, t$ of $\boldsymbol{H}$ are the complex channel gains between transmit antenna j, and receive antenna i with $\mathbb{E}\left[\left|h_{ij}\right|^2\right] = 1$. In this case, the average SNR at each receive antenna is equal to $\gamma = \mathcal{P}/\sigma_n^2$.

When the transmitted signal vector $\boldsymbol{x}$ is composed of t statistically independent equal power components each with a circularly symmetric complex Gaussian distribution, the channel capacity under transmit power constraint $\mathcal{P}$ is given by [1], [2]

$$C = \log_2 \det\left(\boldsymbol{I}_r + \frac{\gamma}{t}\boldsymbol{H}\boldsymbol{H}^\dagger\right) \qquad \text{(bits /s/ Hz)}. \tag{15}$$

The ergodic (mean) capacity[2] of the random MIMO channel, which is the Shannon capacity obtained by assuming it is possible to code over many channel realizations of the ergodic fading process, is evaluated by averaging C with respect to the random channel matrix $\boldsymbol{H}$, i.e., [1]

$$\begin{aligned}\langle C\rangle_{t,r} &= \mathbb{E}\left[\log_2 \det\left(\boldsymbol{I}_r + \frac{\gamma}{t}\boldsymbol{H}\boldsymbol{H}^\dagger\right)\right] \\ &= \mathbb{E}\left[\log_2 \det\left(\boldsymbol{I}_m + \frac{\gamma}{t}\boldsymbol{\Xi}\right)\right]\end{aligned} \tag{16}$$

where the random matrix $\boldsymbol{\Xi} \in \mathbb{C}^{m\times m}$ is defined as

$$\boldsymbol{\Xi} = \begin{cases}\boldsymbol{H}\boldsymbol{H}^\dagger, & t \ge r \\ \boldsymbol{H}^\dagger\boldsymbol{H}, & t < r.\end{cases} \tag{17}$$

A. Independent and Identically Distributed MIMO Channels

Now consider an i.i.d. Rayleigh-fading case, i.e., $\boldsymbol{H} \sim \tilde{\mathcal{N}}_{r,t}(0, \boldsymbol{I}_r \otimes \boldsymbol{I}_t)$. The following result gives a closed-form formula for the ergodic capacity of i.i.d. Rayleigh-fading MIMO channels.

[2]The capacity for fading channels can be defined in a number of ways, depending on the amount of channel knowledge, delay constraints, signaling constraints, and statistical nature of the channel. The various capacity measures for fading channels can be found in [23].

Theorem III.1: If $\boldsymbol{H} \sim \tilde{\mathcal{N}}_{r,t}(0, \boldsymbol{I}_r \otimes \boldsymbol{I}_t)$, i.e., for an i.i.d. Rayleigh-fading MIMO channel with t transmit and r receive antennas, the ergodic capacity in bits/s/Hz under transmit power constraint $\mathcal{P}$ and equal power allocation is given by

$$\begin{aligned}\langle C\rangle_{t,r} = \log_2(e)\, e^{t/\gamma} \sum_{i=0}^{m-1}\sum_{j=0}^{i}\sum_{l=0}^{2j} &\left\{\frac{(-1)^l (2j)!\,(n-m+l)!}{2^{2i-l} j!\,l!\,(n-m+j)!}\right. \\ &\left.\cdot \binom{2i-2j}{i-j}\binom{2j+2n-2m}{2j-l}\sum_{k=0}^{n-m+l} E_{k+1}\left(\frac{t}{\gamma}\right)\right\}.\end{aligned} \tag{18}$$

Proof: If $\boldsymbol{H} \sim \tilde{\mathcal{N}}_{r,t}(0, \boldsymbol{I}_r \otimes \boldsymbol{I}_t)$, then $\boldsymbol{\Xi} \sim \tilde{\mathcal{W}}_m(n, \boldsymbol{I}_m)$. From (16) and Theorem II.1, we have the theorem. □

The analytical expression of $\langle C\rangle_{t,r}$ for the i.i.d. Rayleigh-fading case was first derived by Telatar [1] in single integral form involving the Laguerre polynomials. In contrast, Theorem III.1 provides a closed-form expression for $\langle C\rangle_{t,r}$ in terms of the exponential integral functions (or incomplete gamma functions). Moreover, it generalizes the previously known result of closed-form capacity formulas for Rayleigh-fading channels with reception diversity [22] to MIMO cases.

Example 1: Consider $r = t$. From (18) with $n = t$ and $m = t$, the ergodic capacity of an i.i.d. MIMO channel with t antennas at both the transmitter and the receiver is given by

$$\begin{aligned}\langle C\rangle_{t,t} = \log_2(e)\, e^{t/\gamma} \sum_{i=0}^{t-1}\sum_{j=0}^{i}\sum_{l=0}^{2j} &\left\{\frac{(-1)^l}{2^{2i-l}}\binom{2i-2j}{i-j}\right. \\ &\left.\cdot \binom{2j}{j}\binom{2j}{l}\sum_{k=0}^{l} E_{k+1}\left(\frac{t}{\gamma}\right)\right\}.\end{aligned} \tag{19}$$

Example 2 (Multiple-Input Single-Output (MISO) Channel): Consider $r = 1$, i.e., a MISO channel. From (18) with $n=t$ and $m=1$, the ergodic capacity of an i.i.d. MISO channel with t transmit antennas is given by

$$\langle C\rangle_{t,1} = \log_2(e)\, e^{t/\gamma}\sum_{k=0}^{t-1} E_{k+1}\left(\frac{t}{\gamma}\right). \tag{20}$$

Example 3 (Single-Input Multiple-Output (SIMO) Channel): Consider $t = 1$, i.e., a SIMO channel. From (18) with $n=r$ and $m=1$, the ergodic capacity of an i.i.d. SIMO channel with r receive antennas is given by

$$\langle C\rangle_{1,r} = \log_2(e)\, e^{1/\gamma}\sum_{k=0}^{r-1} E_{k+1}\left(\frac{1}{\gamma}\right) \tag{21}$$

which is in agreement with [22, eq. (40)] if applying the identity (46).

Applying Jensen's inequality to (16) and using (12), we obtain a simple and tight upper bound to (18), first derived by Grant [10, Theorem 2], as follows:

$$\langle C\rangle_{t,r} \le m\log_2\left(\frac{\gamma}{t}\right) + \log_2(m!) + \log_2\left\{L_m^{n-m}\left(-\frac{t}{\gamma}\right)\right\}. \tag{22}$$

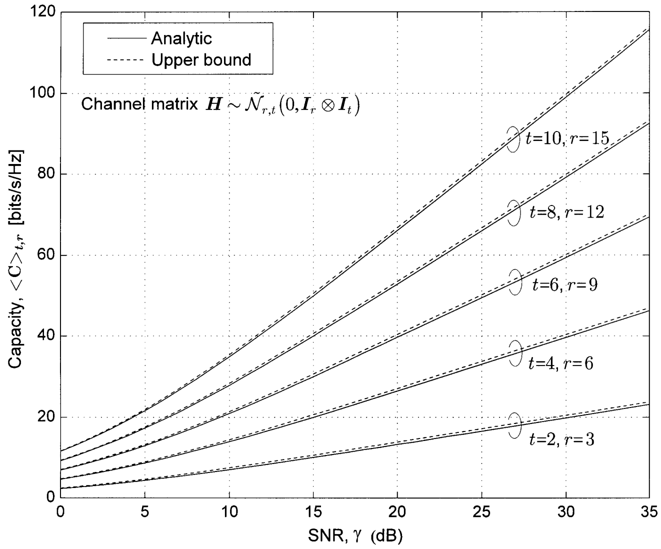

Fig. 1. Ergodic capacity of i.i.d. Rayleigh-fading MIMO channels with t transmit and r receive antennas.

Furthermore, using [10, Lemma A.1], we have at high SNR

$$\begin{aligned}\langle C\rangle_{t,r} &\le \log_2 \mathbb{E}\left[\det\left(\boldsymbol{I}_m + \frac{\gamma}{t}\boldsymbol{\Xi}\right)\right] \\ &\approx \log_2 \mathbb{E}\left[\det\left(\frac{\gamma}{t}\boldsymbol{\Xi}\right)\right] \\ &= m\log_2\left(\frac{\gamma}{t}\right) + \log_2\frac{n!}{(n-m)!} \end{aligned} \tag{23}$$

which can be also obtained from (22) using

$$\lim_{z\to 0} L_n^{\alpha}(z) = \frac{(n+\alpha)!}{\alpha! n!}.$$

Expression (23) implies that at high SNR, the slope of the capacity curve over SNR in decibels is determined by $\min\{t, r\}$—that is, the capacity increases m bits/s/Hz for each 3-dB increase in SNR.

Fig. 1 shows the ergodic capacity of i.i.d. Rayleigh-fading MIMO channels for the following five cases: a) $t=2$, $r=3$; b) $t=4$, $r=6$; c) $t=6$, $r=9$; d) $t=8$, $r=12$; e) $t=10$, $r=15$. The exact capacity and its upper bound are plotted using (18) and (22), respectively. As expected, we see that the slope of the capacity curve over SNR increases with $\min\{t, r\}$ and when the SNR and the ratio between t and r are fixed, the capacity is proportional to $\min\{t, r\}$. For example, the capacity of the channel with $t=10$ and $r=15$ is 115.64 bits/s/Hz at SNR of 35 dB, which is about five times 23.19 bits/s/Hz for $t=2$ and $r=3$.

B. Spatially Correlated MIMO Channels

We consider correlated Rayleigh-fading MIMO channels with the correlation structure of a product form [17], i.e.,

$$\boldsymbol{H} = \boldsymbol{\Phi}_R^{1/2}\boldsymbol{H}_w\boldsymbol{\Phi}_T^{1/2} \tag{24}$$

where $\boldsymbol{H}_w \sim \tilde{\mathcal{N}}_{r,t}(0, \boldsymbol{I}_r \otimes \boldsymbol{I}_t)$, $\boldsymbol{\Phi}_T > 0$ and $\boldsymbol{\Phi}_R > 0$ are transmit and receive correlation matrices, respectively. From (1) and making the transformation from $\boldsymbol{H}_w$ to $\boldsymbol{H}$ with the Jacobian

$$J(\boldsymbol{H}_w \to \boldsymbol{H}) = \det(\boldsymbol{\Phi}_R)^{-t}\det(\boldsymbol{\Phi}_T)^{-r}$$

it is easy to see that $\boldsymbol{H} \sim \tilde{\mathcal{N}}_{r,t}(0, \boldsymbol{\Phi}_R \otimes \boldsymbol{\Phi}_T)$. The following result gives an upper bound on the ergodic capacity of such a channel.

Theorem III.2: If $\boldsymbol{H} \sim \tilde{\mathcal{N}}_{r,t}(0, \boldsymbol{\Phi}_R \otimes \boldsymbol{\Phi}_T)$, i.e., for a correlated Rayleigh-fading MIMO channel with t transmit and r receive antennas, the ergodic capacity in bits/s/Hz is bounded as

$$\begin{aligned}\langle C\rangle_{t,r} \le \log_2\Bigg[\sum_{k=0}^{m}\Bigg\{\left(\frac{\gamma}{t}\right)^k k! &\cdot \sum_{1\le i_1<i_2<\cdots<i_k\le t}\det\left(\boldsymbol{\Phi}_{T_{i_1,i_2,\ldots,i_k}}^{i_1,i_2,\ldots,i_k}\right) \\ &\cdot \sum_{1\le u_1<u_2<\cdots<u_k\le r}\det\left(\boldsymbol{\Phi}_{R_{u_1,u_2,\ldots,u_k}}^{u_1,u_2,\ldots,u_k}\right)\Bigg\}\Bigg].\end{aligned} \tag{25}$$

Proof: If $\boldsymbol{H} \sim \tilde{\mathcal{N}}_{r,t}(0, \boldsymbol{\Phi}_R \otimes \boldsymbol{\Phi}_T)$, then $\boldsymbol{\Xi} \sim \tilde{\mathcal{Q}}_{m,n}(\boldsymbol{I}_n, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$ where if $t \ge r$, $\boldsymbol{\Sigma} = \boldsymbol{\Phi}_R$ and $\boldsymbol{\Psi} = \boldsymbol{\Phi}_T$, and if $t < r$, $\boldsymbol{\Sigma} = \boldsymbol{\Phi}_T$ and $\boldsymbol{\Psi} = \boldsymbol{\Phi}_R$. Applying Jensen's inequality to (16) and using Theorem II.3 yield the desired result. □

Note that the upper bound (25) is the logarithm of a polynomial of degree m in γ and the kth-order coefficient of the polynomial depends only on sums of all k-rowed principal minor determinants of correlation matrices. As the mth-order term becomes dominant at high SNR, the asymptotic slope of the capacity curve over SNR in decibels is determined by $\min\{t, r\}$ for even correlated channels. In particular, if $n = m$, using (7), we have at high SNR

$$\begin{aligned}\langle C\rangle_{t,r} &\le \log_2 \mathbb{E}\left[\det\left(\boldsymbol{I}_m + \frac{\gamma}{t}\boldsymbol{\Xi}\right)\right] \\ &\approx m\log_2\left(\frac{\gamma}{t}\right) + \log_2(m!) \\ &\quad + \log_2\det(\boldsymbol{\Phi}_T) + \log_2\det(\boldsymbol{\Phi}_R)\end{aligned} \tag{26}$$

which can be also obtained by taking only the mth-order term in (25). From (23) and (26), we can see that the capacity reduction due to the spatial fading correlation is $-(\log_2\det(\boldsymbol{\Phi}_T) + \log_2\det(\boldsymbol{\Phi}_R))$ bits/s/Hz at high SNR.

Example 4 (Constant Correlation Model): A $d \times d$ correlation matrix is called the dth-order (positive-definite) constant correlation matrix with correlation coefficient $\rho \in [0, 1)$, denoted by $\boldsymbol{\Theta}_d(\rho)$, if it has the following structure:

$$\boldsymbol{\Theta}_d(\rho) = \begin{pmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \cdots & \rho \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \cdots & 1 \end{pmatrix}_{d\times d}.$$

This correlation model may approximate closely spaced antennas and may be used for the worst case analysis or for some rough approximations using the average value of correlation coefficients for all off-diagonal entries of the correlation matrix. Since eigenvalues of $\boldsymbol{\Theta}_d(\rho)$ are $1 + (d-1)\rho$ and $1-\rho$ with $d-1$ multiplicities, its determinant can be written as

$$\det\boldsymbol{\Theta}_d(\rho) = (1-\rho)^{d-1}(1-\rho+d\rho). \tag{27}$$

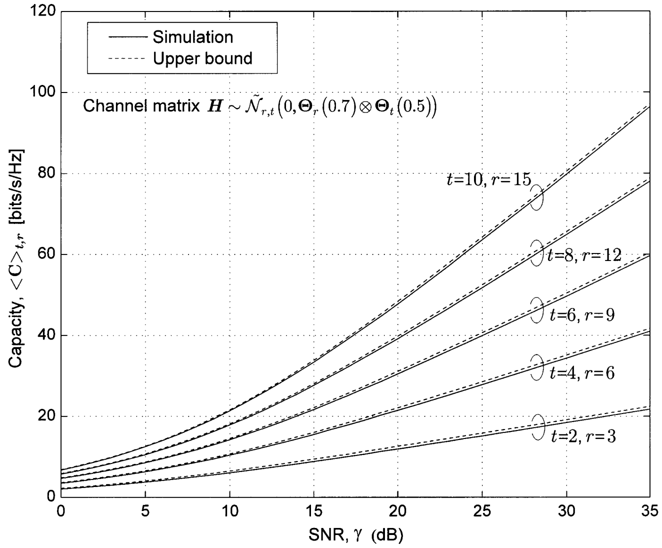

Fig. 2. Ergodic capacity of spatially correlated Rayleigh-fading MIMO channels with t transmit and r receive antennas. The transmit and receive correlations are constant correlations with correlation coefficients $\rho_T = 0.5$ and $\rho_R = 0.7$, i.e., $\mathbf{\Phi}_T = \mathbf{\Theta}_t(0.5)$ and $\mathbf{\Phi}_R = \mathbf{\Theta}_r(0.7)$, respectively.

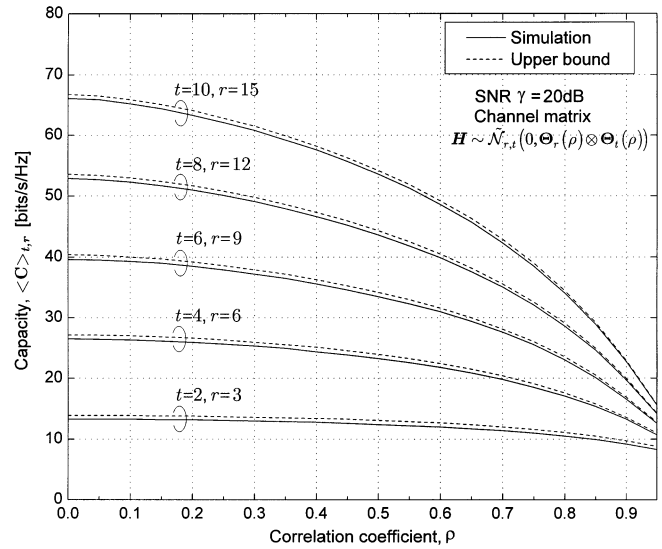

Fig. 3. Ergodic capacity as a function of correlation coefficient ρ for spatially correlated Rayleigh-fading MIMO channels with t transmit and r receive antennas. $\mathbf{\Phi}_T = \mathbf{\Theta}_t(\rho)$ and $\mathbf{\Phi}_R = \mathbf{\Theta}_r(\rho)$.

Furthermore, any k-rowed principal minor matrices of $\mathbf{\Theta}_d(\rho)$ are also the kth-order constant correlation matrices so that their determinants are

$$\det\left(\mathbf{\Theta}_d(\rho)^{i_1, i_2, \ldots, i_k}_{i_1, i_2, \ldots, i_k}\right) = (1-\rho)^{k-1}(1-\rho+k\rho) \quad (28)$$

where $1 \le i_1 < i_2 < \cdots < i_k \le d$.

Let $\mathbf{\Phi}_T = \mathbf{\Theta}_t(\rho_T)$ and $\mathbf{\Phi}_R = \mathbf{\Theta}_r(\rho_R)$ where $\rho_T, \rho_R \in [0, 1)$. Then, from (25) and (28), the ergodic capacity is bounded as

$$\langle C \rangle_{t,r} \le \log_2\left[\sum_{k=0}^{m}\left\{\left(\frac{\gamma}{t}\right)^k k! \binom{t}{k}\binom{r}{k}((1-\rho_T)(1-\rho_R))^{k-1} \cdot (1-\rho_T+k\rho_T)(1-\rho_R+k\rho_R)\right\}\right]. \quad (29)$$

Fig. 2 shows the simulation results and upper bounds for the ergodic capacity of Rayleigh-fading MIMO channels with constant transmit and receive correlations for the same number of antennas as in Fig. 1. The transmit and receive correlation matrices $\mathbf{\Phi}_T$ and $\mathbf{\Phi}_R$ are $\mathbf{\Theta}_t(0.5)$ and $\mathbf{\Theta}_r(0.7)$, respectively. The upper bound is plotted using (29). We can see that upper bounds are quite tight for the entire range of SNRs, regardless of the number of antennas. It can be also shown that the asymptotic slope of capacity curves over SNR is identical with that of i.i.d. cases in Fig. 1, although correlations reduce the diversity advantages (a parallel shift of the capacity curve). Fig. 3 shows the ergodic capacity of uniformly correlated Rayleigh-fading MIMO channels with $\mathbf{\Phi}_T = \mathbf{\Theta}_t(\rho)$ and $\mathbf{\Phi}_R = \mathbf{\Theta}_r(\rho)$ as a function of correlation coefficient ρ at SNR of 20 dB. As expected, the capacity decreases significantly with an increase in correlation coefficient ρ, particularly for larger t and r because the constant correlation is the worst case model.

Example 5 (Jakes Model [15], [27]): In general, the fading correlation depends on both the antenna spacing and the angular spectrum of the incoming radio wave. If we employ a linear array with equally spaced antennas and the classical Jakes correlation model with the uniform angular spectrum [27], the (i, j)th entries of $\mathbf{\Phi}_T$ and $\mathbf{\Phi}_R$ are given by, respectively, [15]

$$\phi_{T,ij} = J_0((i-j) \cdot 2\pi d_T/\lambda)$$
$$\phi_{R,ij} = J_0((i-j) \cdot 2\pi d_R/\lambda)$$

where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind, λ is the wavelength, and d_T and d_R are the interelement distances of the transmit and receive antenna arrays, respectively. More general extension of the Jakes one-ring model of scatterers to MIMO channels has been explored in [14].

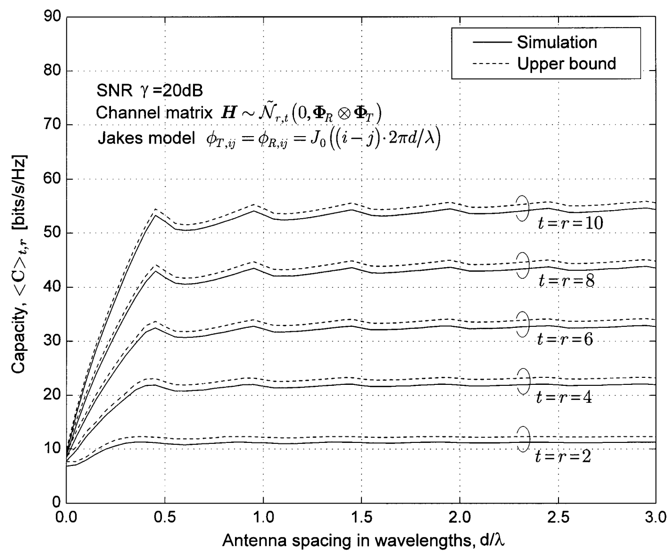

Fig. 4. Ergodic capacity as a function of antenna spacing in wavelengths for spatially correlated Rayleigh-fading MIMO channels with t transmit and r receive antennas. The transmit and receive correlations $\mathbf{\Phi}_T$ and $\mathbf{\Phi}_R$ follow from the Jakes model. $\phi_{T,ij} = \phi_{R,ij} = J_0((i-j) \cdot 2\pi d/\lambda)$ and $\gamma = 20$ dB.

Fig. 4 shows the simulation results and upper bounds for the ergodic capacity of Rayleigh-fading MIMO channels with the Jakes correlation model as a function of antenna spacing at SNR of 20 dB when $t = r = 2, 4, 6, 8$, and 10. The upper bound is plotted using (25). We see that the degradation in capacity due to the fading correlation is small when antenna spacing is greater than 0.5λ, which agrees with the well-known result for spatial diversity systems [27].

C. Double Scattering MIMO Channels [19]

The i.i.d. or correlated Rayleigh fading between antenna elements, which are based on the assumption that only single scattering processes occurred or equivalent single scattering processes could be represented, cannot explain important rank-deficient behavior of MIMO channels. In [19], Gesbert *et al.* proposed a double scattering MIMO channel model that includes rank-deficient effects as well as spatial fading correlation. In double scattering MIMO channels, the channel matrix can be written as [19]

$$\boldsymbol{H} = \frac{1}{\sqrt{s}} \boldsymbol{\Phi}_R^{1/2} \boldsymbol{H}_1 \boldsymbol{\Phi}_S^{1/2} \boldsymbol{H}_2 \boldsymbol{\Phi}_T^{1/2} \tag{30}$$

where $\boldsymbol{H}_1 \sim \tilde{\mathcal{N}}_{r,s}(0, \boldsymbol{I}_r \otimes \boldsymbol{I}_s)$, $\boldsymbol{H}_2 \sim \tilde{\mathcal{N}}_{s,t}(0, \boldsymbol{I}_s \otimes \boldsymbol{I}_t)$, s is the number of effective scatterers on both transmit and receive sides, and correlation matrices $\boldsymbol{\Phi}_T$, $\boldsymbol{\Phi}_R$, and $\boldsymbol{\Phi}_S$ are the transmit, receive, and scatter correlation matrices, respectively. The rank of the MIMO channel (spatial multiplexing ability) is primarily controlled through $\boldsymbol{\Phi}_S$. In this model, it is possible to have uncorrelated fading at both sides but have a rank-deficient MIMO channel and hence poor capacity behavior. This channel is called as a *pinhole* channel.

Theorem III.3: Let $\boldsymbol{H}$ be selected according to (30) at each symbol interval, then the ergodic capacity in bits/s/Hz for such channels is bounded as

$$\begin{aligned} \langle C \rangle_{t,r} \le \log_2 \Bigg[\sum_{k=0}^{\min\{t,r,s\}} \Bigg\{ \left(\frac{\gamma}{st}\right)^k (k!)^2 & \cdot \sum_{1 \le i_1 < i_2 < \cdots < i_k \le t} \det\left(\boldsymbol{\Phi}_{T\,i_1,i_2,\ldots,i_k}^{\;i_1,i_2,\ldots,i_k}\right) \\ & \cdot \sum_{1 \le j_1 < j_2 < \cdots < j_k \le r} \det\left(\boldsymbol{\Phi}_{R_{j_1,j_2,\ldots,j_k}}^{j_1,j_2,\ldots,j_k}\right) \\ & \cdot \sum_{1 \le u_1 < u_2 < \cdots < u_k \le s} \det\left(\boldsymbol{\Phi}_{S_{u_1,u_2,\ldots,u_k}}^{u_1,u_2,\ldots,u_k}\right) \Bigg\} \Bigg]. \end{aligned} \tag{31}$$

Proof: Apply Jensen's inequality to (16) and take the same steps in the proof of Theorem II.3. □

Similar to correlated MIMO channels, the upper bound (31) is the logarithm of a polynomial of degree $\min\{t, r, s\}$ in γ. Therefore, the asymptotic slope of the capacity curve over SNR is determined by $\min\{t, r, s\}$ in double scattering MIMO channels.

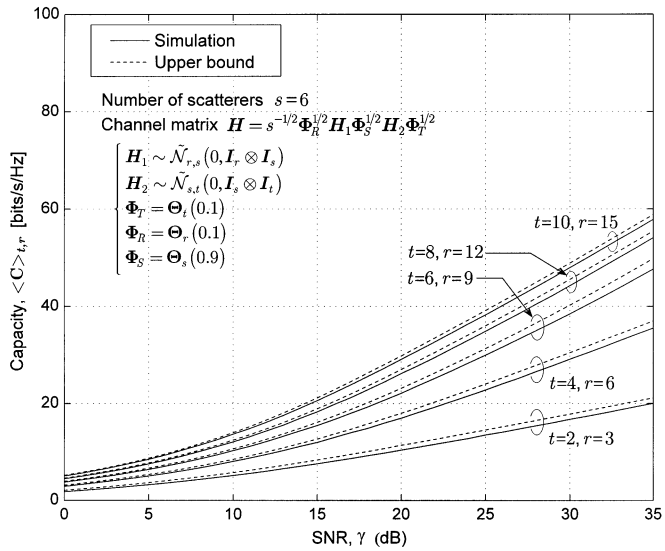

Fig. 5. Ergodic capacity of double scattering MIMO channels with t transmit and r receive antennas. $s = 6$, $\boldsymbol{\Phi}_T = \boldsymbol{\Theta}_t(0.1)$, $\boldsymbol{\Phi}_R = \boldsymbol{\Theta}_r(0.1)$, and $\boldsymbol{\Phi}_S = \boldsymbol{\Theta}_s(0.9)$.

Example 6: If $\boldsymbol{\Phi}_T = \boldsymbol{\Theta}_t(\rho_T)$, $\boldsymbol{\Phi}_R = \boldsymbol{\Theta}_r(\rho_R)$, and $\boldsymbol{\Phi}_S = \boldsymbol{\Theta}_s(\rho_S)$ where $\rho_T, \rho_R, \rho_S \in [0, 1)$, then using (28) and Theorem III.3, we have an upper bound on the ergodic capacity as shown in (32) at the bottom of the page.

Fig. 5 shows the simulation results and upper bounds for the ergodic capacity of double scattering MIMO channels with $s = 6$, $\boldsymbol{\Phi}_T = \boldsymbol{\Theta}_t(0.1)$, $\boldsymbol{\Phi}_R = \boldsymbol{\Theta}_r(0.1)$, and $\boldsymbol{\Phi}_S = \boldsymbol{\Theta}_s(0.9)$. The upper bound is plotted using (32). This example serves to demonstrate the effect of rank deficiency of the channel on the MIMO capacity behavior. We can see that the asymptotic slopes of capacity curves for $t = 8$, $r = 12$ and $t = 10$, $r = 15$ do not increase beyond the slope for $t = 6$, $r = 9$ because $\min\{t, r, s\} = 6$ in all three cases. In other words, since the rank of $\boldsymbol{\Phi}_S$ is 6, there are no further spatial multiplexing benefits from increasing t and r beyond 6. It serves only to provide additional diversity gains.

Example 7: If spatial correlation does not exist, i.e., $\boldsymbol{\Phi}_T = \boldsymbol{I}_t$, $\boldsymbol{\Phi}_R = \boldsymbol{I}_r$, and $\boldsymbol{\Phi}_S = \boldsymbol{I}_s$, we have that

$$\langle C \rangle_{t,r} \le \log_2 \left[\sum_{k=0}^{\min\{t,r,s\}} \left\{ \left(\frac{\gamma}{st}\right)^k (k!)^2 \binom{t}{k} \binom{r}{k} \binom{s}{k} \right\} \right]. \tag{33}$$

Two extreme cases of (33) are the keyhole channel (see Section III-D) and the i.i.d. MIMO channel when $s = 1$ and $s \to \infty$, respectively. If $s = 1$, (33) reduces to

$$\langle C \rangle_{t,r} \le \log_2 (1 + r\gamma) \tag{34}$$

$$\begin{aligned} \langle C \rangle_{t,r} \le \log_2 \Bigg[\sum_{k=0}^{\min\{t,r,s\}} \Bigg\{ \left(\frac{\gamma}{st}\right)^k (k!)^2 \binom{t}{k} \binom{r}{k} \binom{s}{k} \Big((1-\rho_T)(1-\rho_R)(1-\rho_S) \Big)^{k-1} \\ \cdot (1 - \rho_T + k\rho_T)(1 - \rho_R + k\rho_R)(1 - \rho_S + k\rho_S) \Bigg\} \Bigg]. \end{aligned} \tag{32}$$

and if $s \to \infty$, from the fact that

$$\lim_{s\to\infty} \sum_{k=0}^{\min\{t,r,s\}} \left(\frac{\gamma}{st}\right)^k (k!)^2 \binom{t}{k}\binom{r}{k}\binom{s}{k} = \sum_{k=0}^{m} \left(\frac{\gamma}{t}\right)^k k! \binom{t}{k}\binom{r}{k} \quad (35)$$

(33) becomes (22).

D. Keyhole MIMO Channels [20], [21]

In certain MIMO propagation environments, a degenerate channel with only a single degree of freedom (i.e., one-rank channel matrix) may arise due to the keyhole effect [20], [21]. In such a channel, the only way for the radio wave to propagate from the transmitter to the receiver is to pass through the keyhole, and each entry of $\boldsymbol{H}$ is a product of two independent complex Gaussian RVs rather than the complex Gaussian. Then, the channel matrix $\boldsymbol{H}$ for keyhole MIMO channels is given by [20], [21]

$$\boldsymbol{H} = \boldsymbol{\beta}\boldsymbol{\alpha}^T = \begin{pmatrix} \alpha_1\beta_1 & \alpha_2\beta_1 & \cdots & \alpha_t\beta_1 \\ \alpha_1\beta_2 & \alpha_2\beta_2 & \cdots & \alpha_t\beta_2 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1\beta_r & \alpha_2\beta_r & \cdots & \alpha_t\beta_r \end{pmatrix} \quad (36)$$

where $\boldsymbol{\alpha} \sim \tilde{\mathcal{N}}_t(0, \boldsymbol{I}_t)$ and $\boldsymbol{\beta} \sim \tilde{\mathcal{N}}_r(0, \boldsymbol{I}_r)$ describe the rich scattering at the transmit and receive antenna arrays, respectively, and the keyhole is assumed to ideally reradiate the captured energy, like the transmit and receive scatterers. In fact, the keyhole channel is a special case of the double scattering MIMO channel in Section III-C (see Example 7). Note that as all components of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are independent, all entries of $\boldsymbol{H}$ are uncorrelated but $\mathrm{rank}(\boldsymbol{H}) = 1$. Since $\boldsymbol{H}$ is of one rank, we need not use the random matrix results in analysis for keyhole channels. The following theorem provides a closed-form solution for the ergodic capacity of keyhole MIMO channels.

Theorem III.4: If $\boldsymbol{H} = \boldsymbol{\beta}\boldsymbol{\alpha}^T$ where $\boldsymbol{\alpha} \sim \tilde{\mathcal{N}}_t(0, \boldsymbol{I}_t)$ and $\boldsymbol{\beta} \sim \tilde{\mathcal{N}}_r(0, \boldsymbol{I}_r)$, then the ergodic capacity in bits/s/Hz for such channels is given by

$$\langle C \rangle_{t,r} = \log_2\left(\frac{\gamma}{t}\right) + \log_2(e)\{\psi(t) + \psi(r)\} + \frac{\log_2(e)}{\Gamma(t)\Gamma(r)} G_{2,4}^{3,2}\left(\frac{t}{\gamma}\,\middle|\, \begin{matrix} 1, 1 \\ r, t, 1, 0 \end{matrix}\right) \quad (37)$$

where $\psi(z) = \Gamma'(z)/\Gamma(z)$ is the Euler's *digamma* function [45, eq. (8.360.1)] and $G_{p,q}^{m,n}(\cdot)$ is the Meijer's G-function[3] [45, eq. (9.301)].

Proof: See Appendix C. □

We remark that as $\gamma \to \infty$, the last term in (37) vanishes and the capacity becomes asymptotically

$$\langle C \rangle_{t,r} \to \log_2\left(\frac{\gamma}{t}\right) + \log_2(e)\{\psi(t) + \psi(r)\} \quad (38)$$

which shows that the use of multiple antennas in keyhole channels cannot provide the spatial multiplexing gain and only offers the diversity gain determined by $\psi(t) + \psi(r)$ in ergodic capacity point of view. Using Jensen's inequality and (59), we have an upper bound to (37) as

$$\langle C \rangle_{t,r} \le \log_2(1 + r\gamma)$$

which agrees with (34).

[3] The Meijer's G-function is provided as the built-in function in common mathematical software packages such as MAPLE and MATHEMATICA.

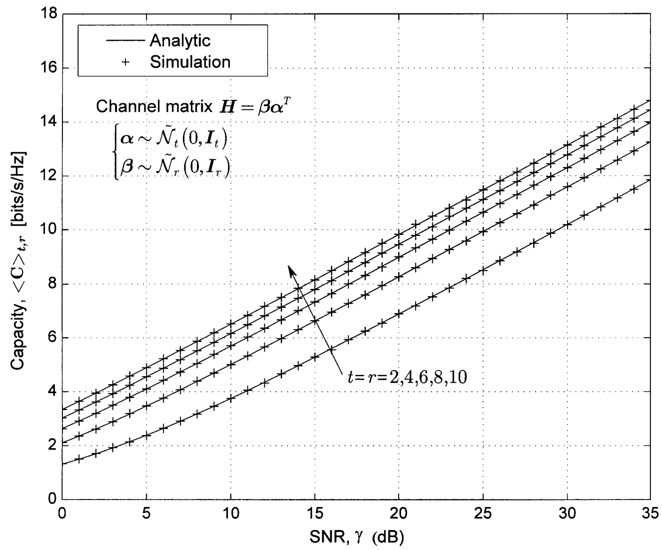

Fig. 6. Ergodic capacity of keyhole MIMO channels with t transmit and r receive antennas.

Fig. 6 shows the ergodic capacity of keyhole MIMO channels when $t = r = 2, 4, 6, 8,$ and 10. Analytical curves are computed using (37) and simulation results are also plotted to verify our analysis. This example demonstrates the effect of keyholes on the MIMO capacity. We can see that for any t and r, the slope of capacity curves over SNR in decibels remains constant since the channel has only a single degree of freedom regardless of the number of antennas. Increasing t and r serves only to provide diversity gains.

IV. Conclusion

In this paper, we studied the capacity of multiple-antenna systems in realistic propagation environments in the presence of spatial fading correlation, double scattering, and keyhole effects. Double scattering can describe the rank-deficient effect as well as spatial fading correlation and the keyhole makes the MIMO channel exhibit uncorrelated spatial fading between antennas but have a one-rank transfer matrix. We obtained the closed-form formula for the ergodic capacity of i.i.d. Rayleigh-fading MIMO channels and upper bounds for correlated and double scattering channels. The upper bounds are in the form of the logarithm of a polynomial in SNR—the degree of the polynomial is equal to the rank of the MIMO channel and the kth-order coefficient depends only on sums of all k-rowed principal minor determinants of correlation matrices—and are quite tight for the entire range of SNRs. In particular, we derived simple and closed-form capacity bounds for constant correlation cases. The closed-form solution for the ergodic capacity of keyhole MIMO channels was also derived.

APPENDIX A
PROOF OF THEOREM II.1

The proof of Theorem II.1 requires the following result on the eigenvalue density of Wishart matrices.

Lemma A.1 (Bronk [39]): If $\boldsymbol{Y} \sim \tilde{\mathcal{W}}_p(q, \boldsymbol{I}_p)$, then the density function of an unordered eigenvalue λ of $\boldsymbol{Y}$ is given by

$$p_\lambda(\lambda) = \frac{1}{p}\sum_{i=0}^{p-1}\frac{i!\lambda^{q-p}e^{-\lambda}}{(q-p+i)!}\left[L_i^{q-p}(\lambda)\right]^2, \qquad \lambda > 0 \quad (39)$$

where $L_n^\alpha(x)$ is the Laguerre polynomial of order n defined as [45, eq. (8.970.1)]

$$L_n^\alpha(x) = \frac{1}{n!}e^x x^{-\alpha}\frac{d^n}{dx^n}\left(e^{-x}x^{n+\alpha}\right) = \sum_{l=0}^{n}(-1)^l\binom{n+\alpha}{n-l}\frac{x^l}{l!}. \quad (40)$$

The above result can also be found in [1]. From (40) and the identity of [45, eq. (8.976.3)]

$$[L_i^\alpha t(x)]^2 = \frac{\Gamma(\alpha+i+1)}{2^{2i}i!}\sum_{j=0}^{i}\frac{\binom{2i-2j}{i-j}(2j)!}{j!\Gamma(\alpha+j+1)}L_{2j}^{2\alpha}(2x) \quad (41)$$

the eigenvalue density of the Wishart matrix $\boldsymbol{Y} \sim \tilde{\mathcal{W}}_p(q, \boldsymbol{I}_p)$ in (39) can be rewritten as

$$p_\lambda(\lambda) = \frac{1}{p}\sum_{i=0}^{p-1}\sum_{j=0}^{i}\sum_{l=0}^{2j}\left\{\frac{(-1)^l(2j)!}{2^{2i-l}j!l!(q-p+j)!}\cdot\binom{2i-2j}{i-j}\binom{2j+2q-2p}{2j-l}\lambda^{q-p+l}e^{-\lambda}\right\}, \qquad \lambda > 0. \quad (42)$$

Using (42), we have

$$\mathbb{E}\left[\ln\det(\boldsymbol{I}_p + \mu\boldsymbol{Y})\right] = \sum_{i=0}^{p-1}\sum_{j=0}^{i}\sum_{l=0}^{2j}\left\{\frac{(-1)^l(2j)!}{2^{2i-l}j!l!(q-p+j)!}\cdot\binom{2i-2j}{i-j}\binom{2j+2q-2p}{2j-l}\cdot\underbrace{\int_0^\infty \ln(1+\mu\lambda)\lambda^{q-p+l}e^{-\lambda}\,d\lambda}_{=I}\right\}. \quad (43)$$

To evaluate the integral I in (43), we use the following result from [22, Appendix B]:

$$\mathcal{I}_n(\nu) \triangleq \int_0^\infty \ln(1+x)x^{n-1}e^{-\nu x}dx, \qquad \nu > 0 \text{ and } n = 1, 2, \ldots$$
$$= (n-1)!e^\nu\sum_{k=1}^{n}\frac{\Gamma(-n+k,\nu)}{\nu^k} \quad (44)$$

where $\Gamma(a, z) = \int_z^\infty e^{-x}x^{a-1}dx$ is the complementary incomplete gamma function [45, eq. (8.350.2)]. Using (44), we get

$$I = (q-p+l)!\mu^{-q+p-l-1}e^{1/\mu}\cdot\sum_{k=1}^{q-p+l+1}\mu^k\Gamma\left(-q+p-l-1+k, \mu^{-1}\right). \quad (45)$$

Furthermore, the exponential integral function $E_n(z)$ is the special case of the complementary incomplete gamma function, i.e.,

$$E_n(z) = z^{n-1}\Gamma(1-n, z). \quad (46)$$

From (45) and (46), we obtain

$$I = (q-p+l)!e^{1/\mu}\sum_{k=0}^{q-p+l}E_{k+1}\left(\frac{1}{\mu}\right). \quad (47)$$

Substituting (47) into (43), we complete the proof of the theorem. □

APPENDIX B
PROOF OF THEOREM II.2

Before proceeding to prove the theorem, it is necessary to give the definitions of hypergeometric functions of matrix arguments.

1) Complex multivariate hypergeometric coefficient for a partition κ [38, eq. (84)]:

$$(a)_\kappa = \prod_{i=1}^{p}(a-i+1)_{k_i} = \frac{\tilde{\Gamma}_p(a,\kappa)}{\tilde{\Gamma}_p(a)} \quad (48)$$

where $(a)_n = a(a+1)\cdots(a+n-1)$, $(a)_0 = 1$, $a \neq 0$ is the Pochhammer symbol, $\kappa = (k_1, k_2, \ldots, k_p)$ denotes a partition of the nonnegative integer k such that $k_1 \geq k_2 \geq \cdots \geq k_p \geq 0$ and $\sum_{i=1}^p k_i = k$, and

$$\tilde{\Gamma}_p(a, \kappa) = \pi^{p(p-1)/2}\prod_{i=1}^{p}\Gamma(a+k_i-i+1) = \tilde{\Gamma}_p(a)\prod_{i=1}^{p}(a-i+1)_{k_i}. \quad (49)$$

2) Hypergeometric functions of matrix arguments [38, eqs. (87) and (88)]:

$${}_m\tilde{F}_n(a_1, \ldots, a_m; b_1, \ldots, b_n; \boldsymbol{A}) = \sum_{k=0}^{\infty}\sum_{\kappa}\frac{(a_1)_\kappa\cdots(a_m)_\kappa}{(b_1)_\kappa\cdots(b_n)_\kappa}\frac{\tilde{C}_\kappa(\boldsymbol{A})}{k!} \quad (50)$$

$${}_m\tilde{F}_n^{(p)}(a_1, \ldots, a_m; b_1, \ldots, b_n; \boldsymbol{A}, \boldsymbol{B}) = \sum_{k=0}^{\infty}\sum_{\kappa}\frac{(a_1)_\kappa\cdots(a_m)_\kappa}{(b_1)_\kappa\cdots(b_n)_\kappa}\frac{\tilde{C}_\kappa(\boldsymbol{A})\tilde{C}_\kappa(\boldsymbol{B})}{\tilde{C}_\kappa(\boldsymbol{I}_p)k!} \quad (51)$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are Hermitian matrices and $\tilde{C}_\kappa(\cdot)$ is the zonal polynomial of a Hermitian matrix [38, eq. (85)].

From the density of $\boldsymbol{Y} \sim \tilde{\mathcal{Q}}_{p,q}(\boldsymbol{A}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$ in (5), we have

$$\mathbb{E}\left[\det(\boldsymbol{Y})^\nu\right] = \frac{1}{\tilde{\Gamma}_p(q)}\det(\boldsymbol{\Sigma})^{-q}\det(\boldsymbol{A\Psi})^{-p}\phi_\nu(\boldsymbol{B}) \quad (52)$$

where

$$\phi_\nu(\boldsymbol{B}) = \int_{\boldsymbol{Y}>0} \det(\boldsymbol{Y})^{q-p+\nu}\, \operatorname{etr}\left(-\varphi^{-1}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}\right) \cdot {}_0\tilde{F}_0^{(q)}\left(;;\boldsymbol{B},\, \varphi^{-1}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}\right) d\boldsymbol{Y}. \quad (53)$$

Using the expansion of the hypergeometric function in terms of zonal polynomials in (51) and the following properties of zonal polynomials [34]:

$$\frac{\tilde{C}_\kappa(\boldsymbol{I}_p)}{\tilde{C}_\kappa(\boldsymbol{I}_q)} = \frac{(p)_\kappa}{(q)_\kappa} \quad (54)$$

and, for Hermitian matrices $\boldsymbol{S} \in \mathbb{C}^{p\times p} > 0$, $\boldsymbol{A} \in \mathbb{C}^{p\times p} > 0$, and $\operatorname{Re}(\alpha) > p-1$

$$\int_{\boldsymbol{S}>0} \operatorname{etr}(-\boldsymbol{S}\boldsymbol{A}) \det(\boldsymbol{S})^{\alpha-p}\, \tilde{C}_\kappa(\boldsymbol{S}\boldsymbol{B})\, d\boldsymbol{S} = \tilde{\Gamma}_p(\alpha,\kappa)\det(\boldsymbol{A})^{-\alpha}\tilde{C}_\kappa\left(\boldsymbol{B}\boldsymbol{A}^{-1}\right) \quad (55)$$

(53) can be evaluated as

$$\begin{aligned}\phi_\nu(\boldsymbol{B}) &= \sum_{k=0}^{\infty}\sum_\kappa \frac{\tilde{C}_\kappa(\boldsymbol{B})}{\tilde{C}_\kappa(\boldsymbol{I}_q)\,k!} \int_{\boldsymbol{Y}>0} \det(\boldsymbol{Y})^{q-p+\nu} \cdot \operatorname{etr}\left(-\varphi^{-1}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}\right)\tilde{C}_\kappa\left(\varphi^{-1}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}\right) d\boldsymbol{Y}\\ &= \det(\varphi\boldsymbol{\Sigma})^{q+\nu}\sum_{k=0}^{\infty}\sum_\kappa \frac{\tilde{\Gamma}_p(q+\nu,\kappa)\,\tilde{C}_\kappa(\boldsymbol{B})\,\tilde{C}_\kappa(\boldsymbol{I}_p)}{\tilde{C}_\kappa(\boldsymbol{I}_q)\,k!}\\ &= \varphi^{p(q+\nu)}\tilde{\Gamma}_p(q+\nu)\det(\boldsymbol{\Sigma})^{q+\nu}\,{}_2\tilde{F}_1(q+\nu,\,p;q;\boldsymbol{B}).\end{aligned} \quad (56)$$

Substituting (56) into (52) yields the result (6). For the special case that $q = p$, using ${}_2\tilde{F}_1(a,\,c;\,c;\,\boldsymbol{A}) = \det(\boldsymbol{I}_p - \boldsymbol{A})^{-a}$ for a Hermitian matrix $\boldsymbol{A} \in \mathbb{C}^{p\times p}$ [34], we can show that (6) reduces to (7). □

Appendix C
Proof of Theorem III.4

Note that

$$\det\left(\boldsymbol{I}_r + \frac{\gamma}{t}\boldsymbol{H}\boldsymbol{H}^\dagger\right) = \det\left(\boldsymbol{I}_r + \frac{\gamma}{t}\|\boldsymbol{\alpha}\|^2\boldsymbol{\beta}\boldsymbol{\beta}^\dagger\right) = 1 + (\gamma/t)\|\boldsymbol{\alpha}\|^2\|\boldsymbol{\beta}\|^2. \quad (57)$$

Let $Z = \|\boldsymbol{\alpha}\|^2\|\boldsymbol{\beta}\|^2$, $U = \|\boldsymbol{\alpha}\|^2$, and $V = \|\boldsymbol{\beta}\|^2$. Since U and V are sums of t and r independent exponential RVs, respectively, they are central chi-square distributed with $2t$ and $2r$ degrees of freedom. The density function of Z is, therefore, given by [24]

$$p_Z(z) = \int_{-\infty}^{\infty} \frac{1}{|u|} p_U(u)\cdot p_V\left(\frac{z}{u}\right) du = \frac{2z^{(t+r)/2-1}}{\Gamma(t)\Gamma(r)} K_{r-t}\left(2\sqrt{z}\right), \qquad z \ge 0 \quad (58)$$

where $K_\nu(\cdot)$ is the νth-order modified Bessel function of the second kind, and the kth moment of Z is given by

$$\mathbb{E}\left[Z^k\right] = \int_0^\infty z^k p_Z(z)\,dz = \frac{\Gamma(t+k)\Gamma(r+k)}{\Gamma(t)\Gamma(r)} = (t)_k(r)_k. \quad (59)$$

From (16), (57), and (58), we have

$$\begin{aligned}\langle C\rangle_{t,r} &= \int_0^\infty \log_2\left(1+\frac{\gamma z}{t}\right) p_Z(z)\,dz\\ &= \underbrace{\int_0^\infty \log_2\left(\frac{\gamma z}{t}\right) p_Z(z)\,dz}_{=I_1} + \underbrace{\int_0^\infty \log_2\left(1+\frac{t}{\gamma z}\right) p_Z(z)\,dz}_{=I_2}.\end{aligned} \quad (60)$$

The integral I_1 in (60) is evaluated as

$$I_1 = \log_2\left(\frac{\gamma}{t}\right) + \log_2(e)\{\psi(t)+\psi(r)\}. \quad (61)$$

Expressing $\log_2\{1+t/(\gamma z)\}$ in terms of the Meijer's G-function, namely [46, eq. (8.4.6.6)]

$$\log_2\left(1+\frac{t}{\gamma z}\right) = \log_2(e)\, G_{2,2}^{2,1}\left(\frac{\gamma z}{t}\,\middle|\,\begin{matrix}0,\,1\\0,\,0\end{matrix}\right) \quad (62)$$

and using the integral table [45, eq. (7.821.3)]

$$\begin{aligned}&\int_0^\infty x^{-\sigma}K_\nu\left(2\sqrt{x}\right) G_{p,q}^{m,n}\left(\alpha x\,\middle|\,\begin{matrix}a_1,\,\ldots,\,a_p\\ b_1,\,\ldots,\,b_q\end{matrix}\right) dx\\ &= \frac{1}{2} G_{p+2,q}^{m,n+2}\left(\alpha\,\middle|\,\begin{matrix}\sigma-\frac{1}{2}\nu,\,\sigma+\frac{1}{2}\nu,\,a_1,\,\ldots,\,a_p\\ b_1,\,\ldots,\,b_q\end{matrix}\right)\\ &\quad p+q < 2(m+n),\ \ |\arg\alpha| < \left(m+n-\tfrac{1}{2}p-\tfrac{1}{2}q\right)\pi,\\ &\quad \operatorname{Re}(\sigma) < 1-\tfrac{1}{2}|\operatorname{Re}(\nu)| + \min_{1\le j\le m}\operatorname{Re}(b_j)\end{aligned} \quad (63)$$

we can evaluate the second integral I_2 in (60) as

$$\begin{aligned}I_2 &= \frac{2\log_2(e)}{\Gamma(t)\Gamma(r)}\int_0^\infty G_{2,2}^{2,1}\left(\frac{\gamma z}{t}\,\middle|\,\begin{matrix}0,\,1\\0,\,0\end{matrix}\right) z^{(t+r)/2-1} \cdot K_{r-t}\left(2\sqrt{z}\right) dz\\ &= \frac{\log_2(e)}{\Gamma(t)\Gamma(r)} G_{4,2}^{2,3}\left(\frac{\gamma}{t}\,\middle|\,\begin{matrix}1-r,\,1-t,\,0,\,1\\0,\,0\end{matrix}\right)\\ &= \frac{\log_2(e)}{\Gamma(t)\Gamma(r)} G_{2,4}^{3,2}\left(\frac{t}{\gamma}\,\middle|\,\begin{matrix}1,\,1\\r,\,t,\,1,\,0\end{matrix}\right)\end{aligned} \quad (64)$$

where the last equality follows from [45, eq. (9.31.2)]. Substituting (61) and (64) into (60) gives the result (37). □

References

[1] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, no. 6, pp. 586–595, Nov./Dec. 1999.

[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] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, no. 3, pp. 357–366, Mar. 1998.

[4] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs. Tech. J.*, pp. 41–59, Autumn 1996.

[5] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.

[6] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, July 1999.

[7] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 46, pp. 543–564, Mar. 2000.

[8] B. L. Hughes, "Differential space-time modulation," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2567–2578, Nov. 2000.

[9] B. M. Hochwald, T. L. Marzetta, and B. Hassibi, "Space-time autocoding," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2761–2781, Nov. 2001.

[10] A. Grant, "Rayleigh fading multiple-antenna channels," *EURASIP J. Appl. Signal Processing (Special Issue on Space-Time Coding (Part I))*, vol. 2002, no. 3, pp. 316–329, Mar. 2002.

[11] S. Verdú and S. Shamai (Shitz), "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 622–640, Mar. 1999.

[12] P. J. Smith and M. Shafi, "On a Gaussian approximation to the capacity of wireless MIMO systems," in *Proc. IEEE Int. Conf. Communications (ICC'02)*, New York, Apr. 2002, pp. 406–410.

[13] Z. Wang and G. B. Giannakis, "Outage mutual information of space-time MIMO channels," in *Proc. 40th Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Oct. 2002, pp. 885–894.

[14] 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, pp. 502–513, Mar. 2000.

[15] D. Chizhik, F. Rashid-Farrokhi, J. Ling, and A. Lozano, "Effect of antenna separation on the capacity of BLAST in correlated channels," *IEEE Commun. Lett.*, vol. 4, pp. 337–339, Nov. 2000.

[16] S. L. Loyka and A. Kouki, "New compound upper bound on MIMO channel capacity," *IEEE Commun. Lett.*, vol. 6, pp. 96–98, Mar. 2002.

[17] 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, pp. 637–650, Mar. 2002.

[18] 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," preprint, May 2002.

[19] D. Gesbert, H. Bölcskei, D. A. Gore, and A. J. Paulraj, "Outdoor MIMO wireless channels: Models and performance prediction," *IEEE Trans. Commun.*, vol. 50, pp. 1926–1934, Dec. 2002.

[20] D. Chizhik, G. J. Foschini, and R. A. Valenzuela, "Capacities of multi-element transmit and receive antennas: Correlations and keyholes," *Electron. Lett.*, vol. 36, no. 13, pp. 1099–1100, June 2000.

[21] 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, pp. 361–368, Apr. 2002.

[22] M.-S. Alouini and A. J. Goldsmith, "Capacity of Rayleigh fading channels under different adaptive transmission and diversity-combining techniques," *IEEE Trans. Veh. Technol.*, vol. 48, pp. 1165–1181, July 1999.

[23] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2619–2692, Oct. 1998.

[24] H. Shin and J. H. Lee, "Effect of keyholes on the symbol error rate of space-time block codes," *IEEE Commun. Lett.*, vol. 7, pp. 27–29, Jan. 2003.

[25] E. T. Browne, *Introduction to the Theory of Determinants and Matrices*. Chapel Hill, NC: Univ. North Carolina Press, 1958.

[26] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, 2nd ed. San Diego, CA: Academic, 1985.

[27] W. C. Jakes, Jr., *Microwave Mobile Communications*. New York: Wiley, 1974.

[28] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. New York: Wiley, 1982.

[29] V. L. Girko, *Theory of Random Determinants*. Dordrecht, The Netherlands: Kluwer, 1990.

[30] A. Edelman, "Eigenvalues and condition numbers of random matrices," Ph.D. dissertation, MIT, Cambridge, MA, May 1989.

[31] 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.

[32] V. L. Girko, "A refinement of the central limit theorem for random determinants," *Theory Probab. Its Applic.*, vol. 42, no. 1, pp. 121–129, 1998.

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

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

[35] R. A. Fisher, "Frequency distribution of the values of the correlation coefficient in samples from indefinitely large population," *Biometrika*, vol. 10, pp. 507–521, 1915.

[36] J. Wishart, "The generalized product moment distribution in samples from a normal multivariate population," *Biometrika*, vol. A20, pp. 32–52, 1928.

[37] N. R. Goodman, "Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction)," *Ann. Math. Statist.*, vol. 34, pp. 152–177, 1963.

[38] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Ann. Math. Statist.*, vol. 35, pp. 475–501, 1964.

[39] B. V. Bronk, "Exponential ensembles for random matrices," *J. Math. Phys.*, vol. 6, pp. 228–237, 1965.

[40] C. G. Khatri, "On certain distribution problems based on positive definite quadratic functions in normal vectors," *Ann. Math. Statist.*, vol. 37, pp. 468–479, 1966.

[41] B. K. Shah, "Distribution theory of a positive definite quadratic form with matrix argument," *Ann. Math. Statist.*, vol. 41, pp. 692–697, 1970.

[42] C. G. Khatri, "Series representations of distributions of quadratic form in the normal vectors and generalized variance," *J. Multivariate Anal.*, vol. 1, pp. 199–214, 1971.

[43] H. Gao and P. J. Smith, "A determinant representation for the distribution of quadratic forms in complex normal vectors," *J. Multivariate Anal.*, vol. 73, pp. 155–165, 2000.

[44] I. Olkin, "The 70th anniversary of the distribution of random matrices: A survey," *Linear Algebra Appl.*, vol. 354, pp. 231–243, Oct. 2002.

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

[46] A. P. Prudnikov, Y. A. Brychkov, and O. I. Marichev, *Integrals and Series*. New York: Gordon and Breach, 1990, vol. 3.