MIMO Diversity in the Presence of Double Scattering

Hyundong Shin, *Member, IEEE*, and Moe Z. Win, *Fellow, IEEE*

***Abstract*—The potential benefits of multiple-antenna systems may be limited by two types of channel degradations—*rank deficiency* and spatial fading correlation of the channel. In this paper, we assess the effects of these degradations on the diversity performance of multiple-input multiple-output (MIMO) systems, with an emphasis on orthogonal space–time block codes (OSTBC), in terms of the symbol error probability (SEP), the effective fading figure (EFF), and the capacity at low signal-to-noise ratio (SNR). In particular, we consider a general family of MIMO channels known as *double-scattering* channels—i.e., *Rayleigh product* MIMO channels—which encompasses a variety of propagation environments from independent and identically distributed (i.i.d.) Rayleigh to degenerate keyhole or pinhole cases by embracing both rank-deficient and spatial correlation effects. It is shown that a MIMO system with n_{T} transmit and n_{R} receive antennas achieves the diversity of order $n_{\mathrm{T}}n_{\mathrm{S}}n_{\mathrm{R}}/\max(n_{\mathrm{T}},n_{\mathrm{S}},n_{\mathrm{R}})$ in a double-scattering channel with n_{S} effective scatterers. We also quantify the combined effect of the spatial correlation and the lack of scattering richness on the EFF and the low-SNR capacity in terms of the *correlation figures* of transmit, receive, and scatterer correlation matrices. We further show the monotonicity properties of these performance measures with respect to the strength of spatial correlation, characterized by the eigenvalue majorization relations of the correlation matrices.**

***Index Terms*—Channel capacity, diversity, double scattering, fading figure, keyhole, multiple-input multiple-output (MIMO) system, orthogonal space–time block code (OSTBC), spatial fading correlation, symbol error probability (SEP).**

I. Introduction

RECENT rapid advances in multiple-input multiple-output (MIMO) communication theory and growing cognizance of the tremendous performance gains achieved by MIMO techniques [1]–[9] have spurred efforts to integrate this technology into future wireless systems such as wireless local area networks (WLANs) and 4G cellular systems. One of the approaches to exploiting diversity capability of MIMO channels is the use of orthogonal space–time block codes (OSTBCs), which have drawn considerable attention because they attain full diversity with scalar maximum-likelihood (ML) decoding [7]–[9].[1]

In general, the potential benefits of multiple-antenna systems may be limited by rank deficiency of the channel due to double scattering or the keyhole effect, for example, as well as spatial fading correlation due, for instance, to insufficient spacing between antenna elements [17]–[30]. Some mechanism rendering a MIMO channel rank deficient cannot be explained by the archetypal model based on single-scattering processes [26], [27]. To address this issue, a double-scattering MIMO model has been proposed recently in [24] wherein the channel matrix is characterized by a product of two statistically independent complex Gaussian matrices, in contrast to the common single complex Gaussian matrix characterization for wireless MIMO channels.[2] This double-scattering model can capture both rank-deficient and spatial correlation effects of MIMO channels and encompass a variety of propagation environments, bridging the gap between an independent and identically distributed (i.i.d.) Rayleigh case and a degenerate one-rank channel known as a keyhole or pinhole channel. There are other recent attempts to modeling MIMO channels for more realistic scattering environments (e.g., double or multibounce diffuse scattering) beyond single scattering [31]–[34].

The effects of rank deficiency and spatial correlation on the capacity of MIMO channels are relatively well understood (see, e.g., [17]–[30]). From a capacity point of view, it has been known that at high signal-to-noise ratio (SNR), the spatial fading correlation reduces the diversity advantage—a parallel shift of the capacity curve over SNR in decibels (dB)—offered by multiple antennas, whereas the rank deficiency decreases the spatial multiplexing benefit—a slope of the capacity curve over SNR—of multiple-antenna channels [21]. Previously, the performance of space–time coding in the presence of spatial fading correlation has been extensively studied for the most popular Rayleigh, Rician, and Nakagami-m fading [35]–[40]. Also, the effect of rank deficiency has been investigated in [41]–[44] for a special case of the keyhole channel.

The objective of this paper is to assess the effects of double scattering on the diversity performance of MIMO systmes in a communication link with n_{T} transmit antennas, n_{R} receive antennas, and n_{S} effective scatterers on each of the transmit and receive sides, which is referred to as a "double-scattering $(n_{\mathrm{T}},n_{\mathrm{S}},n_{\mathrm{R}})$-MIMO channel." Due to the channel decoupling property, the OSTBC converts a MIMO fading channel into identical single-input single-output (SISO) subchannels, each

Manuscript received November 6, 2005; revised February 7, 2007. This work was supported in part by the Korea Research Foundation Grant funded by the Korean Government (KRF-2004-214-D00337), the Office of Naval Research Young Investigator Award N00014-03-1-0489, the National Science Foundation under Grants ANI-0335256 and ECS-0636519, DoCoMo USA Labs, and the Charles Stark Draper Laboratory Robust Distributed Sensor Networks Program.

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

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

Communicated by A. Høst-Madsen, Associate Editor for Detection and Estimation.

Digital Object Identifier 10.1109/TIT.2008.924672

[1]However, OSTBCs with arbitrary complex constellation cannot provide the full diversity and full transmission rate simultaneously for more than two transmit antennas [8, Theorem 5.4.2] (see also [10]–[13]). A new class of quasi-orthogonal codes has been proposed in [14]–[16] with the tradeoff between the decoding complexity, transmission rate, and/or diversity.

[2]In [24], the model was validated by simulations using ray tracing techniques.

for a different transmitted symbol, with a path gain given by the Frobenius norm[3] of the channel matrix $\boldsymbol{H}$ [38]–[42]. As a result, the maximum achievable diversity performance of MIMO systems can be characterized by the statistical property of $\|\boldsymbol{H}\|_{\mathrm{F}}$. Therefore, using the OSTBC as a pivotal MIMO diversity technique[4] (particularly, in the absence of channel knowledge at the transmitter), we analyze the relevant performance measures in double-scattering $(n_{\mathrm{T}}, n_{\mathrm{S}}, n_{\mathrm{R}})$-MIMO channels, namely: i) the symbol error probability (SEP) [49], ii) the effective fading figure (EFF) [50]–[52], and iii) the capacity in a low-SNR regime [53], [54].

Diversity in communication can ameliorate system performances on behalf of error probability, information rate, and signal fluctuation due to fading. From an error probability viewpoint, the diversity attacks a high-SNR slope of the SEP curve, i.e., diversity order. In contrast, the diversity (from a capacity point of view) affects a low-SNR slope of the capacity curve rather than a high-SNR slope. For example, the high- and low-SNR slopes (bits/s/Hz per 3 dB) of the capacity for i.i.d. Rayleigh-fading MIMO channels are given by

$$S_{\infty} = \min(n_{\mathrm{T}}, n_{\mathrm{R}})$$
$$S_0 = \frac{2n_{\mathrm{T}}n_{\mathrm{R}}}{n_{\mathrm{T}} + n_{\mathrm{R}}}$$

respectively [53]. While the high-SNR capacity slope S_{∞} is limited by the spatial multiplexing gain $\min(n_{\mathrm{T}}, n_{\mathrm{R}})$, the low-SNR capacity slope S_0 is limited by the diversity gain amounting to the harmonic mean of n_{T} and n_{R}. Therefore, the capacity is multiplexing-limited in the high-SNR regime, but is diversity-limited in the low-SNR regime. At high SNR, the diversity advantage serves only to provide the power offset (i.e., the parallel shift of the capacity curve) [21]. These lessons stimulate a shift of focus to the low-SNR regime in analyzing the diversity effect on the capacity behavior. More inherently, diversity systems aim to reduce signal fluctuations due to the nature of fading. The EFF measure is defined as a variance-to-mean-square ratio (VMSR) of the instantaneous SNR (see Definition 1). This quantity can be used to assess the severity of fading and the effectiveness of diversity systems on reducing signal fluctuations. The main results of this paper can be summarized as follows.

- We show that the achievable diversity is of order
$$\frac{n_{\mathrm{T}}n_{\mathrm{S}}n_{\mathrm{R}}}{\max(n_{\mathrm{T}}, n_{\mathrm{S}}, n_{\mathrm{R}})}.$$
Hence, if the channel is "rich-enough," that is, the number of effective scatterers is greater than or equal to the numbers of transmit and receive antennas, the full spatial diversity order of $n_{\mathrm{T}}n_{\mathrm{R}}$ can be achieved even in the presence of double scattering.
- We derive exact analytical expressions for the SEP in three cases of particular interest:
  1) spatially uncorrelated double scattering (includes i.i.d. and keyhole channels as special cases);
  2) doubly correlated double scattering (includes a spatially correlated MIMO channel where spatial correlation is present at both the transmitter and the receiver);
  3) multiple-input single-output (MISO) double scattering (corresponds to a pure transmit diversity system wherein a burden of diversity reception at the receive terminal is moved to the transmitter—original motivation of space–time coding [6]–[8].
- We derive the EFF and the low-SNR capacity of double-scattering $(n_{\mathrm{T}}, n_{\mathrm{S}}, n_{\mathrm{R}})$-MIMO channels. The results show that these performance measures are completely characterized by the *correlation figures* of transmit, receive, and scatterer correlation matrices.[5]
- The EFF as a functional of the eigenvalues of correlation matrices is monotonically increasing in a sense of Schur (MIS).[6] We show that the maximum possible increase in the EFF due to double scattering is a sum of correlation figures of the transmit and receive correlation matrices, which eventuates when the scatterers tend to be fully correlated or the keyhole propagation takes place, that is, when only a single degree of freedom is available in the channel for communications.
- The low-SNR capacity slope as a functional of the eigenvalues of correlation matrices is monotonically decreasing in a sense of Schur (MDS). We also obtain the low-SNR capacity of a double-scattering MIMO channel without the constraint of orthogonal input signaling. This enables us to assess the penalty of the use of OSTBCs (for achieving full diversity with simple decoding) on spectral efficiency in the low-SNR regime.

We note in passing that all the mathematical and statistical results (on the monotonicity in a sense of *Schur* and random matrices) obtained in the appendices are applicable to many other problems related to multiple-antenna communications—for example, capacity analysis of MIMO relay channels [5] and spatially correlated MIMO channels [21]–[23], and error probability analysis of multiple-antenna systems with cochannel interference [55], [56].

This paper is organized as follows. In Section II, the system model considered in the paper is presented. Section III analyzes the achievable diversity and the SEP in the presence of double scattering. Section IV analyzes the EFF and the low-SNR capacity (with and without the use of OSTBCs) of double-scattering $(n_{\mathrm{T}}, n_{\mathrm{S}}, n_{\mathrm{R}})$-MIMO channels. Section V concludes the paper. In relation to our study, the notions of majorization and Schur monotonicity are briefly discussed in Appendix I. In Appendix II, we provide supplementary useful results on some statistics derived from complex Gaussian matrices.

[3]The Frobenius norm of an $m \times n$ matrix $\boldsymbol{A} = (A_{ij})$ is defined as

$$\|\boldsymbol{A}\|_{\mathrm{F}} \triangleq \sqrt{\mathrm{tr}(\boldsymbol{A}\boldsymbol{A}^{\dagger})} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|A_{ij}|^2\right)^{1/2}$$

where $\mathrm{tr}(\cdot)$ and $\dagger$ denote the trace operator and the transpose conjugate of a matrix, respectively.

[4]If the transmitter has channel knowledge, the maximum MIMO diversity can be achieved by *transmit beamforming* (often called maximum ratio transmission (MRT) or MIMO maximal-ratio combining) in the eigenspace of the largest eigenvalue of the Gramian matrix $\boldsymbol{H}^{\dagger}\boldsymbol{H}$ [45]–[48].

[5]The correlation figure is defined as a ratio of the second-order statistic of the spectra of correlation matrices to that of the fully correlated matrix (see Definition 2).

[6]See Appendix I for the notions of *Schur monotonicity* and *majorization*.

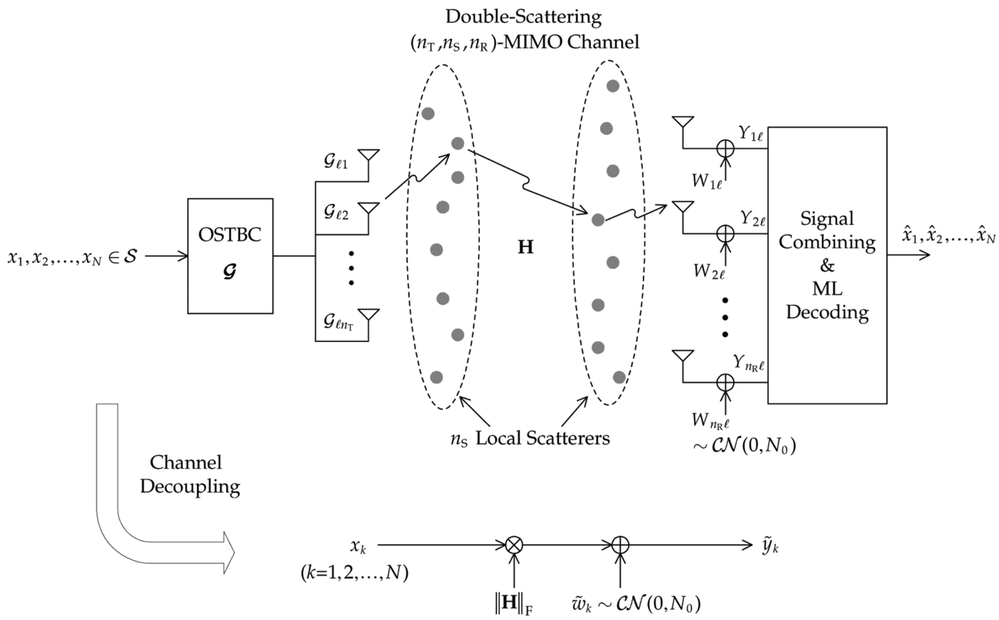

Fig. 1. Block diagram of a space–time block coded system in double-scattering (n_T, n_S, n_R)-MIMO channels and induced SISO subchannels.

Notation: Throughout the paper, we shall use the following notation. $\mathbb{N}$, $\mathbb{R}$, and $\mathbb{C}$ denote the natural numbers and the fields of real and complex numbers, respectively. The superscripts $*$, T, and $\dagger$ stand for the complex conjugate, transpose, and transpose conjugate, respectively. $\boldsymbol{I}_n$ and $\mathbf{0}_{m\times n}$ represent the $n \times n$ identity matrix and the $m \times n$ all-zero matrix, respectively. (A_{ij}) denotes the matrix with the (i,j)th entry A_{ij} and $\det_{1\le i,j\le n}(A_{ij})$ is the determinant of the $n \times n$ matrix (A_{ij}). $\mathrm{tr}(\boldsymbol{A})$, $\mathrm{etr}(\boldsymbol{A}) = e^{\mathrm{tr}(\boldsymbol{A})}$, and $\|\boldsymbol{A}\|_F$ denote the trace, exponential of the trace, and Frobenius norm of the matrix $\boldsymbol{A}$, respectively. $\otimes$ and $\oplus$ denote the Kronecker (direct) product and direct sum of matrices and $\mathrm{vec}(\boldsymbol{A})$ denotes the vector formed by stacking all the columns of $\boldsymbol{A}$ into a column vector. Also, we denote $\boldsymbol{A}_1 \otimes \boldsymbol{A}_2 \otimes \cdots \otimes \boldsymbol{A}_n$ by $\bigotimes_{i=1}^n \boldsymbol{A}_i$ and $\boldsymbol{A}_1 \oplus \boldsymbol{A}_2 \oplus \cdots \oplus \boldsymbol{A}_n$ by $\bigoplus_{i=1}^n \boldsymbol{A}_i$. With a slight abuse of notation, a positive-semidefinite matrix $\boldsymbol{A}$ is denoted by $\boldsymbol{A} \ge 0$ and a positive-definite matrix $\boldsymbol{A}$ is denoted by $\boldsymbol{A} > 0$. Finally, for a Hermitian matrix $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ with the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ in any order, $\varrho(\boldsymbol{A})$ denotes the number of distinct eigenvalues of $\boldsymbol{A}$. Also, $\lambda_{\langle k\rangle}$ and $\tau_k(\boldsymbol{A})$, $k = 1, 2, \ldots, \varrho(\boldsymbol{A})$, denote the distinct eigenvalues of $\boldsymbol{A}$ in decreasing order and its multiplicity, respectively, that is, $\lambda_{\langle 1\rangle} > \lambda_{\langle 2\rangle} > \cdots > \lambda_{\langle \varrho(\boldsymbol{A})\rangle}$ and $\sum_{k=1}^{\varrho(\boldsymbol{A})} \tau_k(\boldsymbol{A}) = n$.

II. System Model

We consider a MIMO wireless communication system with n_T transmit and n_R receive antennas, where the channel remains constant for an integer multiple of N_c $(\ge n_T)$ symbol periods and changes independently to a new value for each coherence time. We assume that the channel is perfectly known at the receiver but unknown at the transmitter.

A. Orthogonal Space–Time Block Codes

A space–time block-coded MIMO system in double-scattering channels is illustrated in Fig. 1. During an N_c-symbol interval, symbols $x_i \in \mathcal{S}$, $i = 1, 2, \ldots, N$, are encoded by an OSTBC defined by an $N_c \times n_T$ transmission matrix $\boldsymbol{\mathcal{G}}$, where $\mathcal{S}$ is two-dimensional signaling constellation [8], [9]. A general construction of complex OSTBCs with the minimal delay and maximal achievable rate was presented in [10, Proposition 2]. This construction of the OSTBC for n_T transmit antennas gives the maximal achievable rate [10, Theorem 1]

$$\mathcal{R} = \frac{\lceil \log_2 n_T \rceil + 1}{2^{\lceil \log_2 n_T \rceil}} \tag{1}$$

where $\lceil x \rceil$ denotes the smallest integer greater than or equal to x. For example, Alamouti's code $\begin{bmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{bmatrix}$ is a one-rate OSTBC employing two transmit antennas [7] and

$$\boldsymbol{\mathcal{G}}_4 = \begin{bmatrix} x_1 & x_2 & x_3 & 0 \\ -x_2^* & x_1^* & 0 & -x_3 \\ -x_3^* & 0 & x_1^* & x_2 \\ 0 & x_3^* & -x_2^* & x_1 \end{bmatrix} \tag{2}$$

is a $3/4$-rate OSTBC for four transmit antennas [10].

B. Signal and Channel Models

For a frequency-flat block-fading channel, the $n_R \times N_c$ received signal can be expressed in matrix notation as

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{\mathcal{G}}^T + \boldsymbol{W} \tag{3}$$

where $\boldsymbol{H} \in \mathbb{C}^{n_R \times n_T}$ is the random channel matrix whose (i,j)th entries H_{ij}, $i = 1, 2, \ldots, n_R$, $j = 1, 2, \ldots, n_T$, are complex propagation coefficients between the jth transmit antenna and the ith receive antenna with $\mathbb{E}\{|H_{ij}|^2\} = 1$, and $\boldsymbol{W} \sim \tilde{\mathcal{N}}_{n_R, N_c}(\mathbf{0}_{n_R\times N_c}, N_0 \boldsymbol{I}_{n_R}, \boldsymbol{I}_{N_c})$ is the complex additive white Gaussian noise (AWGN) matrix (see [21, Definition II.1] and [21, eq. (1)] for the definition and distribution of complex Gaussian matrices).[7] The total power transmitted through n_T antennas is assumed to be $\mathcal{P}$ and hence, the average SNR per receive antenna is equal to $\bar{\gamma} \triangleq \mathcal{P}/N_0$.

[7]There exist minor typos in [21, Definition II.1]; the covariance matrix $\boldsymbol{\Sigma} \otimes \boldsymbol{\Psi}$ should be read as $\boldsymbol{\Sigma}^T \otimes \boldsymbol{\Psi}$.

For double-scattering $(n_{\rm T}, n_{\rm S}, n_{\rm R})$-MIMO channels (see Fig. 1), the channel matrix $\boldsymbol{H}$ can be written as [21], [24]

$$\boldsymbol{H} = \frac{1}{\sqrt{n_{\rm S}}}\boldsymbol{\Phi}_{\rm R}^{1/2}\boldsymbol{H}_1\boldsymbol{\Phi}_{\rm S}^{1/2}\boldsymbol{H}_2\boldsymbol{\Phi}_{\rm T}^{1/2} \tag{4}$$

where $n_{\rm S}$ is the number of effective scatterers on each of the transmit and receive sides, $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ are statistically independent, $\boldsymbol{H}_1 \sim \tilde{\mathcal{N}}_{n_{\rm R},n_{\rm S}}(\boldsymbol{0}_{n_{\rm R}\times n_{\rm S}}, \boldsymbol{I}_{n_{\rm R}}, \boldsymbol{I}_{n_{\rm S}})$, $\boldsymbol{H}_2 \sim \tilde{\mathcal{N}}_{n_{\rm S},n_{\rm T}}(\boldsymbol{0}_{n_{\rm S}\times n_{\rm T}}, \boldsymbol{I}_{n_{\rm S}}, \boldsymbol{I}_{n_{\rm T}})$, and Hermitian positive-definite matrices $\boldsymbol{\Phi}_{\rm T}$, $\boldsymbol{\Phi}_{\rm S}$, and $\boldsymbol{\Phi}_{\rm R}$ are $n_{\rm T}\times n_{\rm T}$ transmit, $n_{\rm S}\times n_{\rm S}$ scatterer, and $n_{\rm R}\times n_{\rm R}$ receive correlation matrices with all-diagonal entries 1, respectively.[8] This model can include the rank-deficient effect of MIMO channels as well as spatial fading correlation by controlling $n_{\rm S}$ and the correlation matrices $\boldsymbol{\Phi}_{\rm T}$, $\boldsymbol{\Phi}_{\rm S}$, and $\boldsymbol{\Phi}_{\rm R}$. Therefore, (4) is a general family of MIMO channels spanning from the i.i.d. Rayleigh case ($n_{\rm S}\to\infty$ with $\boldsymbol{\Phi}_{\rm T} = \boldsymbol{I}_{n_{\rm T}}$, $\boldsymbol{\Phi}_{\rm S} = \boldsymbol{I}_{n_{\rm S}}$, $\boldsymbol{\Phi}_{\rm R} = \boldsymbol{I}_{n_{\rm R}}$) to the degenerate keyhole or pinhole case ($n_{\rm S} = 1$ with $\boldsymbol{\Phi}_{\rm T} = \boldsymbol{I}_{n_{\rm T}}$, $\boldsymbol{\Phi}_{\rm R} = \boldsymbol{I}_{n_{\rm R}}$) [24]. Note that the separability of correlation in (4) is a generalization of the well-known "Kronecker model" [17], [18]. Although there have been some attempts to report discrepancy between this separable correlation model and physical measurements (see, e.g., [57], [58]), the Kronecker correlation model has been accepted widely due to its analytical tractability and experimental validation by the European Project [19].

In [20], the so-called *stochastic* rank deficiency—meaning that the channel is rank deficient due to fading correlation, i.e., the correlation matrices have zero eigenvalues—was deemed as an important feature when dealing with fading correlation. However, this form of channel degeneracy cannot cover the case where the channel exhibits rank deficiency even when fading is uncorrelated. In contrast, we shall restrict $\boldsymbol{\Phi}_{\rm T}$, $\boldsymbol{\Phi}_{\rm S}$, and $\boldsymbol{\Phi}_{\rm R}$ to positive-definite (i.e., full rank) matrices in the paper. This implies that the rank of $\boldsymbol{H}$ is equal to $\min(n_{\rm T}, n_{\rm S}, n_{\rm R})$ with probability one. Therefore, rank deficiency can be distinguished from the fading correlation effect and may occur only due to the lack of scattering richness with $n_{\rm S}$ less than $\min(n_{\rm T}, n_{\rm R})$. This also enables us to discriminate a one-rank *fully* correlated scenario from a degenerate keyhole MIMO channel [29], and grants the channel to exhibit rank deficiency with uncorrelated fading (e.g., $n_{\rm S} < \min(n_{\rm T}, n_{\rm R})$ with $\boldsymbol{\Phi}_{\rm T} = \boldsymbol{I}_{n_{\rm T}}$, $\boldsymbol{\Phi}_{\rm S} = \boldsymbol{I}_{n_{\rm S}}$, $\boldsymbol{\Phi}_{\rm R} = \boldsymbol{I}_{n_{\rm R}}$).

Let $\boldsymbol{\Xi}_1 = \boldsymbol{\Phi}_{\rm R}^{1/2}\boldsymbol{H}_1$ and $\boldsymbol{\Xi}_2 = \boldsymbol{\Phi}_{\rm S}^{1/2}\boldsymbol{H}_2\boldsymbol{\Phi}_{\rm T}^{1/2}$, then we have

$$\boldsymbol{H} = \frac{1}{\sqrt{n_{\rm S}}}\boldsymbol{\Xi}_1\boldsymbol{\Xi}_2 \tag{5}$$

where $\boldsymbol{\Xi}_1 \sim \tilde{\mathcal{N}}_{n_{\rm R},n_{\rm S}}(\boldsymbol{0}_{n_{\rm R}\times n_{\rm S}}, \boldsymbol{\Phi}_{\rm R}, \boldsymbol{I}_{n_{\rm S}})$ and $\boldsymbol{\Xi}_2 \sim \tilde{\mathcal{N}}_{n_{\rm S},n_{\rm T}}(\boldsymbol{0}_{n_{\rm S}\times n_{\rm T}}, \boldsymbol{\Phi}_{\rm S}, \boldsymbol{\Phi}_{\rm T})$ are statistically independent complex Gaussian matrices.

III. Symbol Error Probability

With perfect channel knowledge at the receiver, orthogonal space-time block encoding and decoding convert a MIMO fading channel into N equivalent SISO subchannels, each for a different symbol, with a path gain $\|\boldsymbol{H}\|_{\rm F}$ [38]–[42] (as shown in Fig. 1). Consequently, the performance of OSTBCs is completely characterized by the statistical behavior of $\|\boldsymbol{H}\|_{\rm F}$ and the instantaneous SNR for each of the SISO subchannels, denoted by $\gamma_{\rm STBC}$, is given by [41], [42]

$$\gamma_{\rm STBC} = \frac{\bar{\gamma}\,\|\boldsymbol{H}\|_{\rm F}^2}{n_{\rm T}\mathcal{R}}. \tag{6}$$

To evaluate the SEP, we need the probability density function (pdf) or the moment generating function (MGF) of $\gamma_{\rm STBC}$. For double-scattering $(n_{\rm T}, n_{\rm S}, n_{\rm R})$-MIMO channels, the MGF of $\gamma_{\rm STBC}$ can be written as

$$\begin{aligned}
&\phi_{\gamma_{\rm STBC}}(s;\bar{\gamma}) \\
&\quad\triangleq \mathbb{E}\left\{\mathrm{etr}\left(-\frac{s\bar{\gamma}}{n_{\rm T}\mathcal{R}}\boldsymbol{H}\boldsymbol{H}^\dagger\right)\right\} \\
&\quad= \mathbb{E}_{\boldsymbol{\Xi}_1,\boldsymbol{\Xi}_2}\left\{\mathrm{etr}\left(-\frac{s\bar{\gamma}}{n_{\rm S}n_{\rm T}\mathcal{R}}\boldsymbol{\Xi}_1\boldsymbol{\Xi}_2\boldsymbol{\Xi}_2^\dagger\boldsymbol{\Xi}_1^\dagger\right)\right\} \\
&\quad= \mathbb{E}_{\boldsymbol{\Xi}_1}\left\{\det\left(\boldsymbol{I}_{n_{\rm S}n_{\rm T}} + \frac{s\bar{\gamma}}{n_{\rm S}n_{\rm T}\mathcal{R}}\boldsymbol{\Xi}_1^\dagger\boldsymbol{\Xi}_1\boldsymbol{\Phi}_{\rm S}\otimes\boldsymbol{\Phi}_{\rm T}\right)^{-1}\right\} \qquad (7)\\
&\quad= \mathbb{E}_{\boldsymbol{\Xi}_2}\left\{\det\left(\boldsymbol{I}_{n_{\rm R}n_{\rm S}} + \frac{s\bar{\gamma}}{n_{\rm S}n_{\rm T}\mathcal{R}}\boldsymbol{\Phi}_{\rm R}\otimes\boldsymbol{\Xi}_2\boldsymbol{\Xi}_2^\dagger\right)^{-1}\right\} \qquad (8)
\end{aligned}$$

where (7) and (8) follow from Lemma 1 in Appendix II.

A. Achievable Diversity

Before devoting to deriving the SEP expressions, we discuss the diversity order achieved by the OSTBC. In general, the achievable diversity order can be defined as

$$d \triangleq \lim_{\bar{\gamma}\to\infty}\frac{-\log P_{\rm e}}{\log\bar{\gamma}} \tag{9}$$

where $P_{\rm e}$ denotes the SEP for two-dimensional signaling constellation with polygonal decision boundaries. In the absence of double scattering, the OSTBC provides the maximum achievable diversity order of $n_{\rm T}n_{\rm R}$. The corresponding diversity order in double-scattering $(n_{\rm T}, n_{\rm S}, n_{\rm R})$-MIMO channels is given by the following result.

Theorem 1: The diversity order achieved by the OSTBC over double-scattering $(n_{\rm T}, n_{\rm S}, n_{\rm R})$-MIMO channels is

$$d_{\rm STBC} = \frac{n_{\rm T}n_{\rm S}n_{\rm R}}{\max(n_{\rm T}, n_{\rm S}, n_{\rm R})}. \tag{10}$$

Proof: See Appendix III-A.

Theorem 1 states that if the number of effective scatterers is greater than or equal to the numbers of transmit and receive antennas, the OSTBC provides the full diversity order of $n_{\rm T}n_{\rm R}$ even in the presence of double scattering.

We now present analytical expressions for the SEP of the OSTBC for three cases of particular interest—spatially uncorrelated double scattering, doubly correlated double scattering, and MISO double scattering. In what follows, a spatial correlation environment of double-scattering channels is denoted by $\mathbb{T} = (\boldsymbol{\Phi}_{\rm T}, \boldsymbol{\Phi}_{\rm S}, \boldsymbol{\Phi}_{\rm R})$ for given $n_{\rm T}$, $n_{\rm S}$, and $n_{\rm R}$.

[8] In general, a correlation matrix is positive semidefinite with all-diagonal entries 1.

B. Spatially Uncorrelated Double Scattering

Consider a spatial correlation environment $\mathbb{T}_{\rm uc} = (\boldsymbol{I}_{n_{\rm T}}, \boldsymbol{I}_{n_{\rm S}}, \boldsymbol{I}_{n_{\rm R}})$. This spatially uncorrelated double-scattering scenario includes i.i.d. and keyhole MIMO channels as special cases.

Let $n_1 = \min(n_{\rm T}, n_{\rm S})$, $n_2 = \max(n_{\rm T}, n_{\rm S})$, and the $n_1 \times n_1$ random matrix $\boldsymbol{\Upsilon}$ be

$$\boldsymbol{\Upsilon} = \begin{cases} \boldsymbol{\Xi}_2 \boldsymbol{\Xi}_2^\dagger, & \text{if } n_{\rm S} \le n_{\rm T} \\ \boldsymbol{\Xi}_2^\dagger \boldsymbol{\Xi}_2, & \text{if } n_{\rm S} > n_{\rm T} \end{cases} \tag{11}$$

which is a matrix quadratic form in complex Gaussian matrices [21, Definition II.3]. Then, from (8) and (147) in Appendix III, the SEP of the OSTBC with M-PSK signaling in double-scattering $(n_{\rm T}, n_{\rm S}, n_{\rm R})$-MIMO channels can be readily written as

$$P_{\rm e,\ MPSK} = \frac{1}{\pi}\int_0^{\Theta} \mathbb{E}\left\{ \det\left(\boldsymbol{I}_{n_1 n_{\rm R}} + \frac{g\bar{\gamma}\boldsymbol{\Phi}_{\rm R} \otimes \boldsymbol{\Upsilon}}{n_{\rm S} n_{\rm T} \mathcal{R} \sin^2\theta} \right)^{-1} \right\} d\theta \tag{12}$$

where we have used the fact that $\boldsymbol{\Xi}_2 \boldsymbol{\Xi}_2^\dagger$ and $\boldsymbol{\Xi}_2^\dagger \boldsymbol{\Xi}_2$ have the same nonzero eigenvalues.[9]

In the absence of spatial correlation, the random matrix $\boldsymbol{\Upsilon}$ has the Wishart distribution $\tilde{\mathcal{W}}_{n_1}(n_2, \boldsymbol{I}_{n_1})$ [21, Definition II.2]. Applying Corollary 4 in Appendix II to (12), we obtain the SEP for this spatially uncorrelated environment $\mathbb{T}_{\rm uc}$ as

$$P_{\rm e,\ MPSK}^{\rm uc\text{-}ds} = \frac{1}{\pi \mathcal{A}^{\rm uc\text{-}ds}} \int_0^{\Theta} \det\left\{ \mathsf{G}^{\rm uc\text{-}ds}(\theta) \right\} d\theta \tag{13}$$

where

$$\mathcal{A}^{\rm uc\text{-}ds} = \prod_{k=1}^{n_1} (n_2 - k)!\,(k-1)! \tag{14}$$

and $\mathsf{G}^{\rm uc\text{-}ds}(\theta) = \left(\mathsf{G}_{ij}^{\rm uc\text{-}ds}(\theta) \right)$ is the $n_1 \times n_1$ Hankel matrix whose (i,j)th entry is given by

$$\mathsf{G}_{ij}^{\rm uc\text{-}ds}(\theta) = (n_2 - n_1 + i + j - 2)! \times {}_2F_0\left(n_2 - n_1 + i + j - 1, n_{\rm R}; -\frac{g\bar{\gamma}}{n_{\rm S} n_{\rm T} \mathcal{R} \sin^2\theta} \right). \tag{15}$$

Example 1 (Uncorrelated Extremes—Keyhole and i.i.d.): The i.i.d. and keyhole MIMO channels are two extreme cases of spatially uncorrelated double scattering (i.e., $n_{\rm S} = \infty$ and $n_{\rm S} = 1$, respectively). If $n_{\rm S} = 1$, then $n_1 = 1$ and $n_2 = n_{\rm T}$. Hence, (13) reduces to [41, eq. (11)] for keyhole MIMO channels. As $n_{\rm S} \to \infty$, (13) becomes [42, eq. (26)] (with a Nakagami parameter $m = 1$) for i.i.d. Rayleigh-fading MIMO channels.

Fig. 2 shows the SEP of 8-PSK $\boldsymbol{\mathcal{G}}_4$ (2.25 bits/s/Hz) versus the SNR $\bar{\gamma}$ in spatially uncorrelated double-scattering $(4, n_{\rm S}, 2)$-MIMO channels when $n_{\rm S}$ varies from 1 (keyhole) to infinity (i.i.d. Rayleigh). We can see that as $n_{\rm S}$ increases, the SEP approaches that of i.i.d. Rayleigh-fading MIMO channels in the absence of double scattering. This resembles the behavior in Rayleigh-fading channels with diversity reception, that is, the channel behaves like an AWGN channel (diversity order of ∞) as the number of receive antennas increases. Observe that when $n_{\rm S} \ge 4$, the slope of the SEP curve at high SNR is identical to that of the i.i.d. case. This example confirms the result of Theorem 1: the diversity orders are equal to $d_{\rm STBC} = 2, 4$, and 6 for $n_{\rm S} = 1, 2$, and 3, respectively, whereas $d_{\rm STBC} = 8$ for $n_{\rm S} = 5, 10, 20, 100$, and ∞ (i.i.d.). A clearer understanding about the diversity behavior is obtained by referring to Fig. 3, where the SEPs of 16-PSK Alamouti (4 bits/s/Hz) and $\boldsymbol{\mathcal{G}}_4$ (3 bits/s/Hz) OSTBCs versus the SNR $\bar{\gamma}$ in spatially uncorrelated double-scattering $(n_{\rm T}, n_{\rm S}, n_{\rm R})$-MIMO channels are shown. Using (10), we can easily show that the Alamouti and $\boldsymbol{\mathcal{G}}_4$ codes achieve the diversity order of $d_{\rm STBC} = 2$ for $(2,3,1)$ and $(4,2,1)$ channels; $d_{\rm STBC} = 6$ for $(2,5,3)$ and $(4,3,2)$ channels; and $d_{\rm STBC} = 20$ for $(2,10,11)$ and $(4,5,5)$ channels. As can be seen, we obtain a close agreement in the slopes of the SEP curves, corresponding to the same value of $d_{\rm STBC}$, at high SNR.

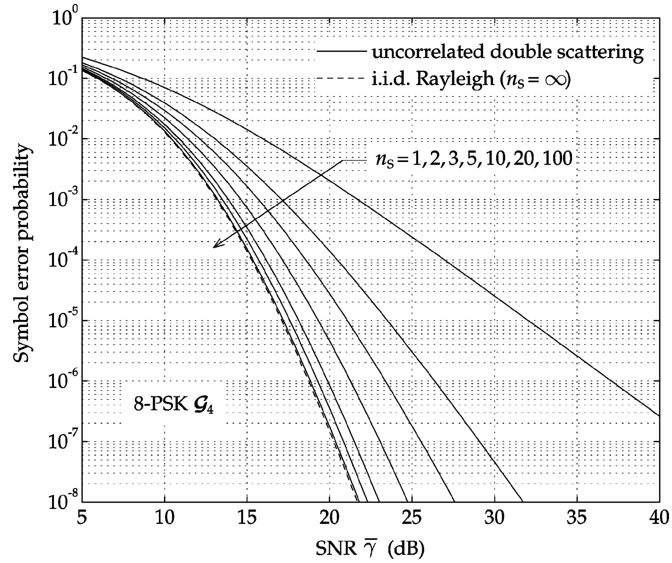

Fig. 2. SEP of 8-PSK $\boldsymbol{\mathcal{G}}_4$ (2.25 bits/s/Hz) versus $\bar{\gamma}$ in spatially uncorrelated double-scattering $(4, n_{\rm S}, 2)$-MIMO channels. $n_{\rm S} = 1, 2, 3, 5, 10, 20, 50, 100, \infty$ (i.i.d. Rayleigh).

C. Doubly Correlated Double Scattering

Consider a spatial correlation environment $\mathbb{T}_{\rm dc} = (\boldsymbol{\Phi}_{\rm T}, \boldsymbol{I}_{n_{\rm S}}, \boldsymbol{\Phi}_{\rm R})$, where spatial correlation exists only on the transmit and receive ends. Note that this scenario includes a spatially correlated MIMO channel in the absence of double scattering ($n_{\rm S} = \infty$) as a special case. Let λ_i^T and $\lambda_j^{\rm R}$, $i = 1, 2, \ldots, n_{\rm T}$, $j = 1, 2, \ldots, n_{\rm R}$, be the eigenvalues of $\boldsymbol{\Phi}_{\rm T}$ and $\boldsymbol{\Phi}_{\rm R}$ in any order, respectively. Suppose that $n_{\rm S} \ge n_{\rm T}$. Then, $\boldsymbol{\Upsilon} \sim \tilde{\mathcal{W}}_{n_{\rm T}}(n_{\rm S}, \boldsymbol{\Phi}_{\rm T})$. Applying Theorem 10 in Appendix II to (12), we obtain the SEP in the environment $\mathbb{T}_{\rm dc}$ as

$$P_{\rm s,\ MPSK}^{\rm dc\text{-}ds} = \frac{1}{\pi \mathcal{A}^{\rm dc\text{-}ds}} \times \int_0^{\Theta} \det\left(\left[\mathsf{G}_1^{\rm dc\text{-}ds}(\theta) \ \ldots \ \mathsf{G}_{\varrho(\boldsymbol{\Phi}_{\rm T})}^{\rm dc\text{-}ds}(\theta) \right] \right) d\theta \tag{16}$$

[9] As mentioned in the proof of Theorem 1, The SEP for the general case of arbitrary two-dimensional signaling constellation with polygonal decision boundaries can be written as a convex combination of terms akin to (147). Thus, our results can be easily extended to any two-dimensional signaling constellation.

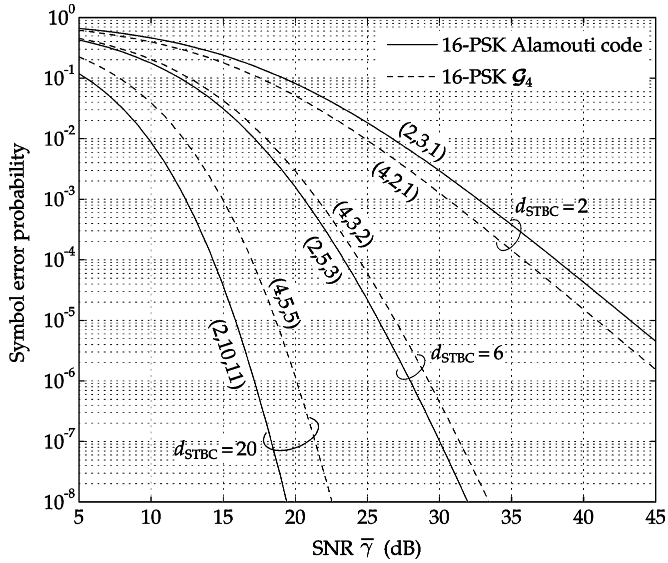

Fig. 3. SEP of 16-PSK Alamouti (4 bits/s/Hz) and $\mathcal{G}_4$ (3 bits/s/Hz) OSTBCs versus $\bar{\gamma}$ in spatially uncorrelated double-scattering $(n_{\rm T}, n_{\rm S}, n_{\rm R})$-MIMO channels. The Alamouti and $\mathcal{G}_4$ codes achieve the diversity order of $d_{\rm STBC} = 2$ in $(2,3,1)$ and $(4,2,1)$ links, respectively. The $d_{\rm STBC}$'s for $(2,5,3), (4,3,2)$ and $(2,10,11), (4,5,5)$ pairs are 6 and 20, respectively.

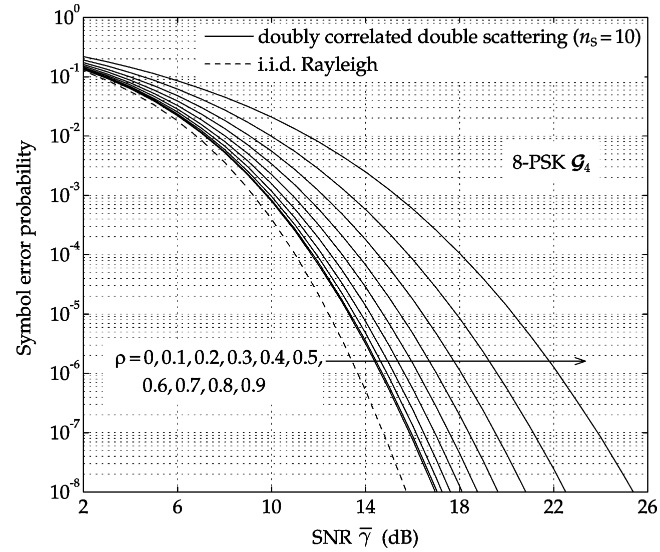

Fig. 4. SEP of 8-PSK $\mathcal{G}_4$ (2.25 bits/s/Hz) versus $\bar{\gamma}$ in doubly correlated double-scattering $(4,10,4)$-MIMO channels. The transmit and receive correlations follow the constant correlation $\mathbf{\Phi}_{\rm T} = \mathbf{\Phi}_{\rm R} = \mathbf{\Phi}_4^{(c)}(\rho)$ for $\rho = 0$ (spatially uncorrelated double-scattering),$0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8,$ and 0.9. For comparison, the SEP for i.i.d. Rayleigh-fading MIMO channels is also plotted.

with

$$\mathcal{A}^{\rm dc\text{-}ds} = \det\left(\left[\mathsf{B}_1^{\rm dc\text{-}ds} \quad \ldots \quad \mathsf{B}_{\varrho(\mathbf{\Phi}_{\rm T})}^{\rm dc\text{-}ds}\right]\right)\prod_{i=1}^{n_{\rm T}}(n_{\rm S}-i)! \quad (17)$$

where $\mathsf{B}_k^{\rm dc\text{-}ds} = \left(\mathsf{B}_{k,ij}^{\rm dc\text{-}ds}\right)$ and $\mathsf{G}_k^{\rm dc\text{-}ds}(\theta) = \left(\mathsf{G}_{k,ij}^{\rm dc\text{-}ds}(\theta)\right)$, $k = 1, 2, \ldots, \varrho(\mathbf{\Phi}_{\rm T})$, are $n_{\rm T} \times \tau_k(\mathbf{\Phi}_{\rm T})$ matrices whose (i,j)th entries are given, respectively, by

$$\mathsf{B}_{k,ij}^{\rm dc\text{-}ds} = (-1)^{i-j}(i-j+1)_{j-1}\,\lambda_{\langle k\rangle}^{T\ \ n_{\rm S}-i+j} \quad (18)$$

and by (19) (shown at the bottom of the page). In (19), $\mathcal{X}_{p,q}(\mathbf{\Phi}_{\rm R})$ is the (p,q)th characteristic coefficient of $\mathbf{\Phi}_{\rm R}$ (see Definition 4 in Appendix II).

Fig. 4 shows the SEP of 8-PSK $\mathcal{G}_4$ versus the SNR $\bar{\gamma}$ in doubly correlated double-scattering $(4,10,4)$-MIMO channels. In this figure, the transmit and receive correlations follow the constant correlation $\mathbf{\Phi}_{\rm T} = \mathbf{\Phi}_{\rm R} = \mathbf{\Phi}_4^{(c)}(\rho)$, defined by (53) in Appendix I, and the correlation coefficient ρ ranges from 0 (spatially uncorrelated double scattering) to 0.9. The characteristic coefficients of the constant correlation matrix are given by (131) and (132) (see Example 6 in Appendix II). For comparison, we also plot the SEP of i.i.d. Rayleigh-fading MIMO channels. In Fig. 4, we can see that the SNR penalty due to double scattering with $n_{\rm S} = 10$ (in the absence of spatial correlation) is about 1 dB at the SEP of 10^{-6} and it becomes larger than 2.5 dB for $\rho \geq 0.5$. In Fig. 5, the SEP of 8-PSK $\mathcal{G}_4$ at $\bar{\gamma} = 15$ dB is depicted as a function of a correlation coefficient ρ for doubly correlated double-scattering $(4, n_{\rm S}, 4)$-MIMO channels with constant correlation $\mathbf{\Phi}_{\rm T} = \mathbf{\Phi}_{\rm R} = \mathbf{\Phi}_4^{(c)}(\rho)$ when $n_{\rm S} = 5, 10, 20, 50, 100,$ and ∞ (doubly correlated Rayleigh). This figure demonstrates that double scattering and spatial correlation degrade the SEP performance considerably.

D. MISO Double Scattering

Finally, we consider a double-scattering MISO channel. This is a pure transmit diversity system wherein the burden of diversity reception at the receive terminal is moved to the transmitter. The SEP in double-scattering MISO channels can be obtained from (8) with $n_{\rm R} = 1$ as

$$P_{\rm e,\ MPSK}^{\rm miso\text{-}ds} = \frac{1}{\pi}\int_0^{\Theta}\mathbb{E}\left\{\det\left(\boldsymbol{I}_{n_{\rm S}} + \frac{g\bar{\gamma}\mathbf{\Xi}_2\mathbf{\Xi}_2^{\dagger}}{n_{\rm S}n_{\rm T}\mathcal{R}\sin^2\theta}\right)^{-1}\right\}d\theta. \quad (20)$$

Let $\lambda_i^{\rm S}$, $i = 1, 2, \ldots, n_{\rm S}$, be the eigenvalues of $\mathbf{\Phi}_{\rm S}$ in any order. Then, applying Theorem 11 in Appendix II to (20), we obtain

$$P_{\rm e,\ MPSK}^{\rm miso\text{-}ds} = \frac{1}{\pi}\sum_{p=1}^{\varrho(\mathbf{\Phi}_{\rm S})}\sum_{q=1}^{\varrho(\mathbf{\Phi}_{\rm T})}\sum_{i=1}^{\tau_p(\mathbf{\Phi}_{\rm S})}\sum_{j=1}^{\tau_q(\mathbf{\Phi}_{\rm T})}\left\{\mathcal{X}_{p,i}(\mathbf{\Phi}_{\rm S})\,\mathcal{X}_{q,j}(\mathbf{\Phi}_{\rm T}) \times \int_0^{\Theta}{}_2F_0\left(i, j; -\frac{g\bar{\gamma}\lambda_{\langle p\rangle}^{\rm S}\lambda_{\langle q\rangle}^{T}}{n_{\rm S}n_{\rm T}\mathcal{R}\sin^2\theta}\right)d\theta\right\} \quad (21)$$

$$\mathsf{G}_{k,ij}^{\rm dc\text{-}ds}(\theta) = \sum_{p=1}^{\varrho(\mathbf{\Phi}_{\rm R})}\sum_{q=1}^{\tau_p(\mathbf{\Phi}_{\rm R})}\left\{\mathcal{X}_{p,q}(\mathbf{\Phi}_{\rm R})\,\lambda_{\langle k\rangle}^{T\ \ n_{\rm S}-n_{\rm T}+i+j-1}(n_{\rm S}-n_{\rm T}+i+j-2)! \times {}_2F_0\left(n_{\rm S}-n_{\rm T}+i+j-1, q; -\frac{g\bar{\gamma}\lambda_{\langle p\rangle}^{\rm R}\lambda_{\langle k\rangle}^{T}}{n_{\rm S}n_{\rm T}\mathcal{R}\sin^2\theta}\right)\right\}. \quad (19)$$

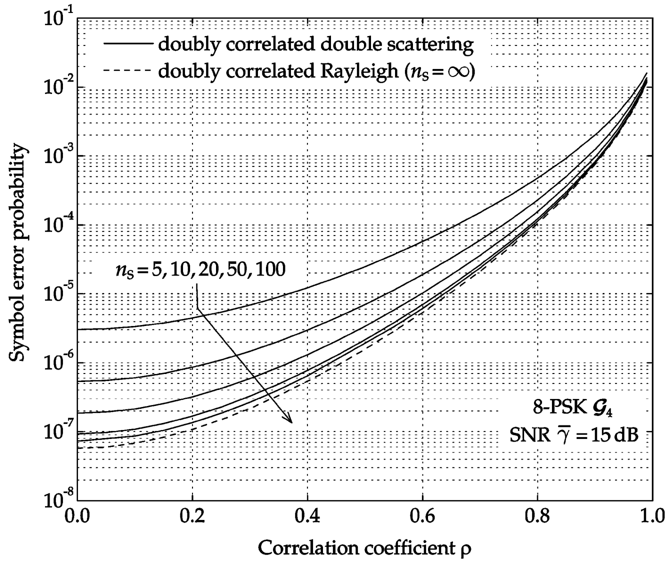

Fig. 5. SEP of 8-PSK $\mathcal{G}_4$ (2.25 bits/s/Hz) as a function of correlation coefficient ρ in doubly correlated double-scattering (4, n_S, 4)-MIMO channels with constant correlation $\mathbf{\Phi}_\mathrm{T} = \mathbf{\Phi}_\mathrm{R} = \mathbf{\Phi}_4^{(\mathrm{c})}(\rho)$. $n_\mathrm{S} = 5$, 10, 20, 50, 100, ∞ (doubly correlated Rayleigh) and $\bar{\gamma} = 15$ dB.

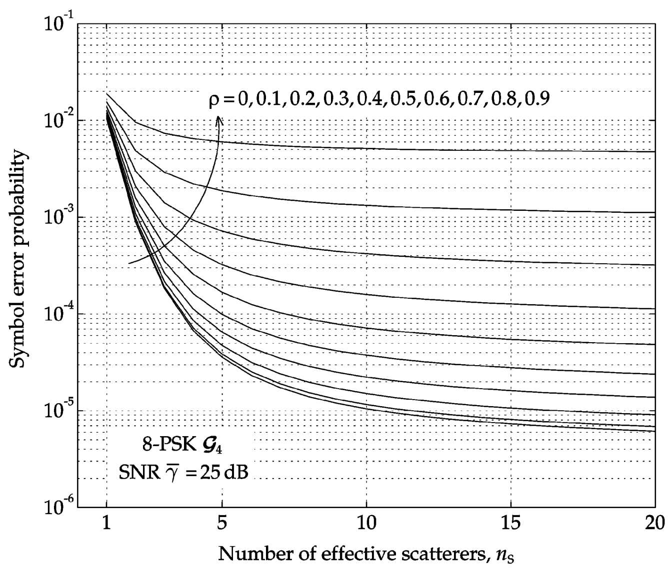

Fig. 6. SEP of 8-PSK $\mathcal{G}_4$ (2.25 bits/s/Hz) versus n_S in double-scattering $(4, n_\mathrm{S}, 1)$-MIMO channels. The transmit and scatterer correlations follow the constant correlation $\mathbf{\Phi}_\mathrm{T} = \mathbf{\Phi}_4^{(\mathrm{c})}(\rho)$ and $\mathbf{\Phi}_\mathrm{S} = \mathbf{\Phi}_{n_\mathrm{S}}^{(\mathrm{c})}(\rho)$ for $\rho = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8,$ and 0.9. $\bar{\gamma} = 25$ dB.

where $\mathcal{X}_{p,i}(\mathbf{\Phi}_\mathrm{S})$ and $\mathcal{X}_{q,j}(\mathbf{\Phi}_\mathrm{T})$ are the characteristic coefficients of $\mathbf{\Phi}_\mathrm{S}$ and $\mathbf{\Phi}_\mathrm{T}$, respectively.

The effects of the spatial correlation and the number of effective scatterers on the SEP performance in MISO channels can be ascertained by referring to Fig. 6, where the SEP of 8-PSK $\mathcal{G}_4$ at $\bar{\gamma} = 25$ dB versus n_S is depicted for double-scattering $(4, n_\mathrm{S}, 1)$-MIMO channels. The transmit and scatterer correlations follow the constant correlation $\mathbf{\Phi}_\mathrm{T} = \mathbf{\Phi}_4^{(\mathrm{c})}(\rho)$ and $\mathbf{\Phi}_\mathrm{S} = \mathbf{\Phi}_{n_\mathrm{S}}^{(\mathrm{c})}(\rho)$ where ρ varies from 0 to 0.9. Note that the maximum achievable diversity order is equal to $d_\mathrm{STBC} = 4$ for $n_\mathrm{S} \geq 4$. Hence, the SEP performance improves rapidly as n_S increases, and approaches the corresponding SEP in the absence of double scattering.

IV. Effective Fading Figure and Low-SNR Capacity

In this section, we access the combined effect of rank deficiency and spatial correlation on the performance of OSTBCs in terms of the EFF and the capacity in a low-SNR regime. It will be apparent that these performance measures are completely characterized by the *kurtosis* of $\|\boldsymbol{H}\|_\mathrm{F}$.

A. Effective Fading Figure

One of the goals of diversity systems is to reduce the signal fluctuation due to the stochastic nature of multipath fading. Therefore, it is of interest to characterize the variation of the instantaneous SNR at the output where the amount of signal fluctuations is measured. The following measure can be used to assess the severity of fading and the effectiveness of diversity systems on reducing signal fluctuations.

Definition 1 (Effective Fading Figure): For the instantaneous SNR γ at the output of interest in a communication system subject to fading, the EFF in decibels for the output SNR γ is defined as the VMSR of γ, i.e.,

$$\mathsf{EFF}_\gamma\,(\mathrm{dB}) \triangleq 10\log_{10}\left\{\frac{\mathbb{V}\mathrm{ar}\{\gamma\}}{(\mathbb{E}\{\gamma\})^2}\right\}. \tag{22}$$

It should be noted that the EFF is akin to the notions of the normalized standard deviation (NSD) of the instantaneous combiner output SNR [50]–[52] and the amount of fading (AF) [59], [60]. The AF, as defined in [59, eq. (2)], is purely to characterize the amount of random fluctuations in the channel itself and conveys no information about diversity systems. In contrast, the NSD is a measure of the signal fluctuations at the diversity combiner output, enabling us to compare the effectiveness of diversity combining techniques such as maximal-ratio combining (MRC), equal-gain combining (EGC), selection combining (SC), and hybrid section/maximal-ratio combining (H-S/MRC). If the signal fluctuation is measured at each branch output, the EFF is synonymous with the AF. In contrast, when the signal fluctuation is measured at the diversity combiner output, the EFF is equal to the square of the NSD of the instantaneous SNR at the combiner output. The term "AF" was also confusingly used for diversity systems in some literature with a view to bridging the philosophy between characterizing physical channel fading and quantifying the degree of diversity effectiveness [42], [61], [62].

By definition, the efficiency of OSTBCs on reducing the severity of fading can be assessed by

$$\begin{aligned}\mathsf{EFF}_\mathrm{STBC}\ (\mathrm{dB}) &\triangleq 10\log_{10}\left\{\frac{\mathbb{V}\mathrm{ar}\{\gamma_\mathrm{STBC}\}}{(\mathbb{E}\{\gamma_\mathrm{STBC}\})^2}\right\}\\ &= 10\log_{10}\{\kappa(\|\boldsymbol{H}\|_\mathrm{F}) - 1\}\end{aligned} \tag{23}$$

where $\kappa(\|\boldsymbol{H}\|_\mathrm{F})$ is the kurtosis of $\|\boldsymbol{H}\|_\mathrm{F}$ defined by

$$\begin{aligned}\kappa(\|\boldsymbol{H}\|_\mathrm{F}) &\triangleq \frac{\mathbb{E}\left\{[\|\boldsymbol{H}\|_\mathrm{F} - \mathbb{E}\{\|\boldsymbol{H}\|_\mathrm{F}\}]^4\right\}}{\left(\mathbb{E}\left\{[\|\boldsymbol{H}\|_\mathrm{F} - \mathbb{E}\{\|\boldsymbol{H}\|_\mathrm{F}\}]^2\right\}\right)^2}\\ &= \frac{\mathbb{E}\left\{\|\boldsymbol{H}\|_\mathrm{F}^4\right\}}{\left(\mathbb{E}\left\{\|\boldsymbol{H}\|_\mathrm{F}^2\right\}\right)^2}.\end{aligned} \tag{24}$$

In (24), the second equality follows from the fact that the kurtosis is invariant with respect to translations of a random variable. Note that the minimum EFF is equal to $-\infty$ decibels if there is no random fluctuation in the received signal. Also, the EFF is equal to 0 dB for Rayleigh fading without diversity and hence, $\mathsf{EFF}_{\mathrm{STBC}} > 0$ dB means that the variation of the instantaneous SNR in each SISO subchannel is more severe than that in Rayleigh fading.

1) Note on the Kurtosis of $\|\boldsymbol{H}\|_{\mathrm{F}}$*:* The kurtosis measures the peakedness or flatness of a distribution [63]. It has been revealed that this normalized form of the fourth statistic of fading distributions plays a key role in the low-SNR behavior of the spectral efficiency in fading channels [53], [64]. To proceed with deriving $\kappa(\|\boldsymbol{H}\|_{\mathrm{F}})$ for double-scattering $(n_{\mathrm{T}}, n_{\mathrm{S}}, n_{\mathrm{R}})$-MIMO channels, we first define the following scalar quantity related to a correlation matrix.

Definition 2 (Correlation Figure): For an arbitrary $n \times n$ correlation matrix $\mathbf{\Phi}$, the *correlation figure* of $\mathbf{\Phi}$ is defined by

$$\zeta(\mathbf{\Phi}) \triangleq \frac{\operatorname{tr}\left(\mathbf{\Phi}^2\right)}{\operatorname{tr}\left(\mathbf{1}_n^2\right)} = \frac{1}{n^2}\operatorname{tr}\left(\mathbf{\Phi}^2\right) \tag{25}$$

where $\mathbf{1}_n$ denotes the $n \times n$ all-one matrix.

Note that $\frac{1}{n} \le \zeta(\mathbf{\Phi}) \le 1$, where the lower and upper bounds correspond to uncorrelated and fully correlated cases, respectively.[10] The following Schur monotonicity properties hold for the correlation figure (the proofs are given in Appendix III-B).

Property 1: Let $\mathbf{\Phi}$ be an $n \times n$ correlation matrix. Then, the correlation figure $\zeta(\mathbf{\Phi})$ as a functional of the eigenvalues of $\mathbf{\Phi}$ is MIS, that is, if $\mathbf{\Phi} \preceq \grave{\mathbf{\Phi}}$, then

$$\zeta(\mathbf{\Phi}) \le \zeta(\grave{\mathbf{\Phi}}). \tag{26}$$

Property 2: Let $\mathbf{\Phi}_i$, $i = 1, 2, \ldots, m$, be $n_i \times n_i$ correlation matrices. Then, the product of correlation figures, $\prod_{i=1}^{m} \zeta(\mathbf{\Phi}_i)$, as a functional of the eigenvalues of $\bigotimes_{i=1}^{m} \mathbf{\Phi}_i$, is MIS, that is, if

$$\bigotimes_{i=1}^{m} \mathbf{\Phi}_i \preceq \bigotimes_{i=1}^{m} \grave{\mathbf{\Phi}}_i, \tag{27}$$

then

$$\prod_{i=1}^{m} \zeta(\mathbf{\Phi}_i) \le \prod_{i=1}^{m} \zeta(\grave{\mathbf{\Phi}}_i). \tag{28}$$

Property 3: Let $\mathbf{\Phi}_i$, $i = 1, 2, \ldots, m$, be $n_i \times n_i$ correlation matrices. Then, the sum of correlation figures $\sum_{i=1}^{m} \zeta(\mathbf{\Phi}_i)$, as a functional of the eigenvalues of $\bigoplus_{i=1}^{m} \frac{1}{n_i}\mathbf{\Phi}_i$, is MIS, that is, if

$$\bigoplus_{i=1}^{m} \frac{1}{n_i}\mathbf{\Phi}_i \preceq \bigoplus_{i=1}^{m} \frac{1}{n_i}\grave{\mathbf{\Phi}}_i \tag{29}$$

then

$$\sum_{i=1}^{m} \zeta(\mathbf{\Phi}_i) \le \sum_{i=1}^{m} \zeta(\grave{\mathbf{\Phi}}_i). \tag{30}$$

The next theorem shows that $\kappa(\|\boldsymbol{H}\|_{\mathrm{F}})$ depends exclusively on the spectra of spatial correlation matrices and is quantified solely by their correlation figures.

Theorem 2: For double-scattering $(n_{\mathrm{T}}, n_{\mathrm{S}}, n_{\mathrm{R}})$-MIMO channels, the kurtosis of $\|\boldsymbol{H}\|_{\mathrm{F}}$ is

$$\kappa(\|\boldsymbol{H}\|_{\mathrm{F}}) = \zeta(\mathbf{\Phi}_{\mathrm{T}})\,\zeta(\mathbf{\Phi}_{\mathrm{R}}) + \zeta(\mathbf{\Phi}_{\mathrm{T}})\,\zeta(\mathbf{\Phi}_{\mathrm{S}}) + \zeta(\mathbf{\Phi}_{\mathrm{R}})\,\zeta(\mathbf{\Phi}_{\mathrm{S}}) + 1. \tag{31}$$

Proof: See Appendix III-C. □

Example 2 (Spatially Uncorrelated Double Scattering): In the absence of spatial fading correlation ($\mathbb{T}_{\mathrm{uc}}$), we have

$$\kappa(\|\boldsymbol{H}\|_{\mathrm{F}}) = \frac{1}{n_{\mathrm{T}}n_{\mathrm{R}}} + \frac{1}{n_{\mathrm{T}}n_{\mathrm{S}}} + \frac{1}{n_{\mathrm{R}}n_{\mathrm{S}}} + 1. \tag{32}$$

As compared with the i.i.d. case, the keyhole increases the kurtosis of the fading distribution in SISO subchannels by twice the reciprocal of the harmonic mean between the numbers of transmit and receive antennas, that is, $\frac{1}{n_{\mathrm{T}}} + \frac{1}{n_{\mathrm{R}}}$.

Next, we show the Schur monotonicity property of $\kappa(\|\boldsymbol{H}\|_{\mathrm{F}})$.

Corollary 1: Let

$$\boldsymbol{\mathcal{J}}(\mathbb{T}) \triangleq \frac{\mathbf{\Phi}_{\mathrm{T}} \otimes \mathbf{\Phi}_{\mathrm{R}}}{n_{\mathrm{T}}n_{\mathrm{R}}} \oplus \frac{\mathbf{\Phi}_{\mathrm{T}} \otimes \mathbf{\Phi}_{\mathrm{S}}}{n_{\mathrm{T}}n_{\mathrm{S}}} \oplus \frac{\mathbf{\Phi}_{\mathrm{S}} \otimes \mathbf{\Phi}_{\mathrm{R}}}{n_{\mathrm{S}}n_{\mathrm{R}}} \tag{33}$$

for a spatial correlation environment $\mathbb{T} = (\mathbf{\Phi}_{\mathrm{T}}, \mathbf{\Phi}_{\mathrm{S}}, \mathbf{\Phi}_{\mathrm{R}})$. Then, the kurtosis of $\|\boldsymbol{H}\|_{\mathrm{F}}$, as a functional of the eigenvalues of $\boldsymbol{\mathcal{J}}(\mathbb{T})$, is a MIS (or isotone) function, that is, if $\boldsymbol{\mathcal{J}}(\mathbb{T}_1) \preceq \boldsymbol{\mathcal{J}}(\mathbb{T}_2)$, then

$$\kappa(\|\boldsymbol{H}\|_{\mathrm{F}}; \mathbb{T}_1) \le \kappa(\|\boldsymbol{H}\|_{\mathrm{F}}; \mathbb{T}_2). \tag{34}$$

Proof: It follows immediately from Theorem 2 and Properties 2 and 3 stating the fact that the product and sum of correlation figures preserve the monotonicity property. □

Corollary 1 implies that the less spatially correlated fading results in the less peaky fading distribution of each SISO subchannel.

2) Note on the EFF of γ_{STBC}*:* From Theorem 2 and (23), it is straightforward to see that the $\mathsf{EFF}_{\mathrm{STBC}}$ in double-scattering $(n_{\mathrm{T}}, n_{\mathrm{S}}, n_{\mathrm{R}})$-MIMO channels is given by

$$\mathsf{EFF}_{\mathrm{STBC}}(\mathrm{dB}) = 10\log_{10}\Big\{\zeta(\mathbf{\Phi}_{\mathrm{T}})\,\zeta(\mathbf{\Phi}_{\mathrm{R}}) + \zeta(\mathbf{\Phi}_{\mathrm{T}})\,\zeta(\mathbf{\Phi}_{\mathrm{S}}) + \zeta(\mathbf{\Phi}_{\mathrm{R}})\,\zeta(\mathbf{\Phi}_{\mathrm{S}})\Big\} \tag{35}$$

from which we can make the following observations on the $\mathsf{EFF}_{\mathrm{STBC}}$.

- The $\mathsf{EFF}_{\mathrm{STBC}}$ as a functional of the eigenvalues of $\boldsymbol{\mathcal{J}}(\mathbb{T})$ is MIS, that is,

$$\mathsf{EFF}_{\mathrm{STBC}}(\mathbb{T}_1) \le \mathsf{EFF}_{\mathrm{STBC}}(\mathbb{T}_2) \tag{36}$$

  whenever $\boldsymbol{\mathcal{J}}(\mathbb{T}_1) \preceq \boldsymbol{\mathcal{J}}(\mathbb{T}_2)$. This reveals that the less spatially correlated fading results in the less severe random

[10] Similar to (25), the *correlation number* was defined as $\frac{1}{n}\operatorname{tr}\left(\mathbf{\Phi}^2\right)$[54]. While the correlation figure and number are the second-order statistics of the spectra of a correlation matrix, normalized by those of fully correlated and uncorrelated matrices, respectively, the correlation figure is bounded by $0 \le \zeta(\mathbf{\Phi}) \le 1$ for any correlation structure, as $n \to \infty$.

fluctuations in equivalent SISO subchannels induced by OSTBCs.

- In the absence of double scattering, $\zeta(\mathbf{\Phi}_{\mathrm{S}})$ is equal to zero and thus, the double scattering together with spatial correlation causes the $\mathsf{EFF}_{\mathrm{STBC}}$ to increase by the amount of $\zeta(\mathbf{\Phi}_{\mathrm{T}})\zeta(\mathbf{\Phi}_{\mathrm{S}}) + \zeta(\mathbf{\Phi}_{\mathrm{R}})\zeta(\mathbf{\Phi}_{\mathrm{S}})$. In particular, the maximum increase in the $\mathsf{EFF}_{\mathrm{STBC}}$ is a sum of correlation figures of the transmit and receive correlation matrices, that is, $\zeta(\mathbf{\Phi}_{\mathrm{T}})+\zeta(\mathbf{\Phi}_{\mathrm{R}})$, which eventuates when $\mathbf{\Phi}_{\mathrm{S}}$ goes to be fully correlated or when the keyhole effect takes place.

B. Low-SNR Capacity

Recent information-theoretic studies show that the first-order analysis of the capacity versus the SNR fails to reveal the impact of the channel and that second-order analysis is required to assess the wideband or low-SNR performance of communication systems [53], [54]. In particular, it was demonstrated that the tradeoff between the capacity in bits per second per hertz (bits/s/Hz) and energy per bit required for reliable communication is the key measure of channel capacity in a low-SNR regime. In this regime, the capacity can be characterized by two parameters, namely, i) $\frac{E_{\mathrm{b}}}{N_0}_{\min}$, the minimum bit energy per noise level required to reliably communicate at any positive data rate (where E_{b} denotes the total transmitted energy per bit), and ii) S_0, the *low-SNR slope* (bits/s/Hz per 3 dB) of the capacity at the point $\frac{E_{\mathrm{b}}}{N_0}_{\min}$.

1) General Input Signaling: Before proceeding to study the low-SNR capacity achieved by OSTBCs, we first deal with the more general case of input signaling, assuming that the fading process is ergodic and coding is across many independent fading blocks without a delay constraint.

Theorem 3: Consider a general $(n_{\mathrm{T}}, n_{\mathrm{S}}, n_{\mathrm{R}})$-MIMO double scattering channel

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{X} + \boldsymbol{W} \tag{37}$$

where the channel matrix $\boldsymbol{H}$ is given by (4) at each coherence interval and the input signal $\boldsymbol{X} \in \mathbb{C}^{n_{\mathrm{T}}\times N_{\mathrm{c}}}$ is subject to the power constraint $\mathbb{E}\{\|\boldsymbol{X}\|_{\mathrm{F}}^2\} = N_{\mathrm{c}}\mathcal{P}$. Suppose that the receiver knows the realization of $\boldsymbol{H}$, but the transmitter has no channel knowledge. Then, the minimum required $\frac{E_{\mathrm{b}}}{N_0}$ for reliable communication is

$$\frac{E_{\mathrm{b}}}{N_0}_{\min} = \frac{\log_e 2}{n_{\mathrm{R}}} \tag{38}$$

and the low-SNR slope (bits/s/Hz per 3 dB) of the capacity is

$$S_0 = \frac{2}{\zeta(\mathbf{\Phi}_{\mathrm{T}}) + \zeta(\mathbf{\Phi}_{\mathrm{S}}) + \zeta(\mathbf{\Phi}_{\mathrm{R}}) + \zeta(\mathbf{\Phi}_{\mathrm{T}})\zeta(\mathbf{\Phi}_{\mathrm{S}})\zeta(\mathbf{\Phi}_{\mathrm{R}})}. \tag{39}$$

Proof: See Appendix III-D. □

From Theorem 3, we can make the following observations.

- The $\frac{E_{\mathrm{b}}}{N_0}_{\min}$ is inversely proportional to n_{R}, whereas the double scattering and spatial fading correlation as well as the numbers of transmit antennas and effective scatterers do not affect this measure. Moreover, regardless of the number of antennas and propagation conditions, the minimum received bit energy per noise level required for reliable communication, $\frac{E_{\mathrm{b}}^{\mathrm{r}}}{N_0}_{\min}$, is equal to

$$\frac{E_{\mathrm{b}}^{\mathrm{r}}}{N_0}_{\min} = n_{\mathrm{R}} \cdot \frac{E_{\mathrm{b}}}{N_0}_{\min} = -1.59 \text{ dB} \tag{40}$$

which is a fundamental feature of the channels where the additive noise is Gaussian [53, Theorem 1].

- The low-SNR slope S_0 as a functional of the eigenvalues of $\grave{\boldsymbol{\mathcal{J}}}(\mathbb{T})$ is MDS, that is, if $\grave{\boldsymbol{\mathcal{J}}}(\mathbb{T}_1) \preceq \grave{\boldsymbol{\mathcal{J}}}(\mathbb{T}_2)$, then

$$S_0(\mathbb{T}_1) \geq S_0(\mathbb{T}_2) \tag{41}$$

where $\grave{\boldsymbol{\mathcal{J}}}(\mathbb{T})$ is defined for the environment $\mathbb{T} = (\mathbf{\Phi}_{\mathrm{T}}, \mathbf{\Phi}_{\mathrm{S}}, \mathbf{\Phi}_{\mathrm{R}})$ as follows:

$$\grave{\boldsymbol{\mathcal{J}}}(\mathbb{T}) \triangleq \frac{\mathbf{\Phi}_{\mathrm{T}}}{n_{\mathrm{T}}} \oplus \frac{\mathbf{\Phi}_{\mathrm{S}}}{n_{\mathrm{S}}} \oplus \frac{\mathbf{\Phi}_{\mathrm{R}}}{n_{\mathrm{R}}} \oplus \frac{\mathbf{\Phi}_{\mathrm{T}} \otimes \mathbf{\Phi}_{\mathrm{S}} \otimes \mathbf{\Phi}_{\mathrm{R}}}{n_{\mathrm{T}}n_{\mathrm{S}}n_{\mathrm{R}}}. \tag{42}$$

Note that (41) follows from (39) and Properties 2 and 3. This MDS property reveals that the low-SNR slope decreases with the amount of spatial correlation in contrast to the high-SNR capacity slope $\min(n_{\mathrm{T}}, n_{\mathrm{R}}, n_{\mathrm{S}})$, which is invariant with respect to spatial correlation [21].

Example 3 (Dual-Antenna System): Consider $n_{\mathrm{T}} = n_{\mathrm{R}} = 2$. In the presence of spatially uncorrelated double scattering, the low-SNR slope for general double-scattering $(2, n_{\mathrm{S}}, 2)$-MIMO channels is

$$S_0 = 2 \cdot \left(1 + \frac{1}{n_{\mathrm{S}}} \cdot \frac{5}{4}\right)^{-1} \text{ bits/s/Hz per 3 dB} \tag{43}$$

which is bounded by $8/9 \leq S_0 \leq 2$. The lowest and highest slopes are achieved when $n_{\mathrm{S}} = 1$ (keyhole) and $n_{\mathrm{S}} = \infty$ (i.i.d.), respectively.

2) OSTBC Input Signaling: We now turn our attention to the low-SNR behavior of the capacity for double-scattering $(n_{\mathrm{T}}, n_{\mathrm{S}}, n_{\mathrm{R}})$-MIMO channels employing OSTBCs.

Theorem 4: Consider an $(n_{\mathrm{T}}, n_{\mathrm{S}}, n_{\mathrm{R}})$-MIMO double scattering channel

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{\mathcal{G}}^T + \boldsymbol{W}$$

where the channel matrix $\boldsymbol{H}$ is given by (4) at each coherence interval and the OSTBC $\boldsymbol{\mathcal{G}}$ is subject to the power constraint $\mathbb{E}\{\|\boldsymbol{\mathcal{G}}\|_{\mathrm{F}}^2\} = N_{\mathrm{c}}\mathcal{P}$. Then, the OSTBC achieves the minimum required $\frac{E_{\mathrm{b}}}{N_0}_{\min}$ the same as that without the orthogonal signaling constraint

$$\frac{E_{\mathrm{b}}}{N_0}_{\min}^{\mathrm{STBC}} = \frac{\log_e 2}{n_{\mathrm{R}}} \tag{44}$$

and the low-SNR slope (bits/s/Hz per 3 dB) of the capacity

$$S_0^{\mathrm{STBC}} = \frac{2\mathcal{R}}{\zeta(\mathbf{\Phi}_{\mathrm{T}})\zeta(\mathbf{\Phi}_{\mathrm{R}}) + \zeta(\mathbf{\Phi}_{\mathrm{T}})\zeta(\mathbf{\Phi}_{\mathrm{S}}) + \zeta(\mathbf{\Phi}_{\mathrm{R}})\zeta(\mathbf{\Phi}_{\mathrm{S}}) + 1}. \tag{45}$$

Proof: See Appendix III-E. □

From Theorem 4, we can make the following observations in parallel to Section IV-B1.

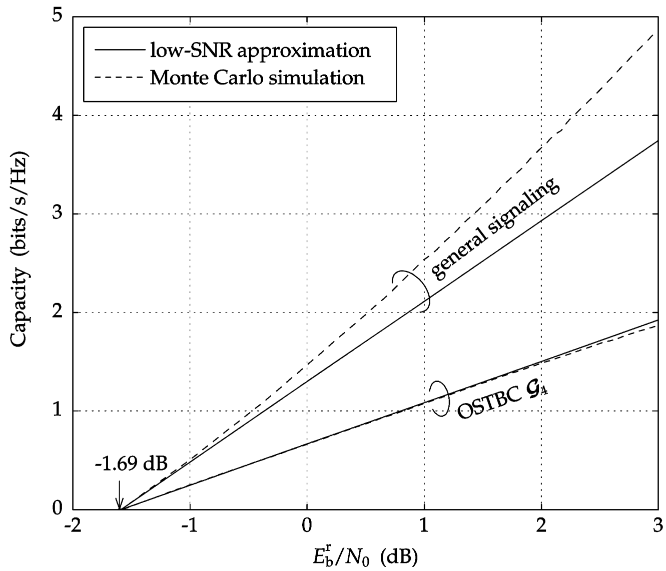

Fig. 7. Capacity in bits/s/Hz versus the received $\frac{E_\mathrm{b}}{N_0}$ for the general input signaling and OSTBC $\boldsymbol{\mathcal{G}}_4$ in double-scattering $(4, 20, 4)$-MIMO channels with exponential correlation $\boldsymbol{\Phi}_\mathrm{T} = \boldsymbol{\Phi}_\mathrm{R} = \boldsymbol{\Phi}_4^{(\mathrm{e})}(0.5)$ and $\boldsymbol{\Phi}_\mathrm{S} = \boldsymbol{\Phi}_{20}^{(\mathrm{e})}(0.5)$.

- As compared with the general case, the use of OSTBCs does not increase the minimum required $\frac{E_\mathrm{b}}{N_0}$ for reliable communication in MIMO channels.
- The low-SNR slope S_0^{STBC} as a functional of the eigenvalues of $\boldsymbol{\mathcal{J}}(\mathbb{T})$ is MDS, that is, if $\boldsymbol{\mathcal{J}}(\mathbb{T}_1) \preceq \boldsymbol{\mathcal{J}}(\mathbb{T}_2)$, then

$$S_0^{\mathrm{STBC}}(\mathbb{T}_1) \geq S_0^{\mathrm{STBC}}(\mathbb{T}_2). \tag{46}$$

  In contrast, we see from (159) that the high-SNR slope of the capacity is equal to $\mathcal{R}$, which does not depend on spatial correlation and double scattering.

Example 4 (Alamouti's Code): Consider $n_\mathrm{T} = n_\mathrm{R} = 2$. In the presence of spatially uncorrelated double scattering, the low-SNR slope for Alamouti's code with two receive antennas is

$$S_0^{\mathrm{STBC}} = \frac{8}{5} \cdot \left(1 + \frac{1}{n_\mathrm{S}} \cdot \frac{4}{5}\right)^{-1} \text{ bits/s/Hz per 3 dB} \tag{47}$$

which is bounded by $8/9 \leq S_0^{\mathrm{STBC}} \leq 8/5$.

In Fig. 7, the capacity (bits/s/Hz) versus $\frac{E_\mathrm{b}^\mathrm{r}}{N_0}_{\min}$ and its low-SNR approximation are depicted with and without the signaling constraint of the OSTBC $\boldsymbol{\mathcal{G}}_4$ in double-scattering $(4, 20, 4)$-MIMO channels with exponential correlation $\boldsymbol{\Phi}_\mathrm{T} = \boldsymbol{\Phi}_\mathrm{R} = \boldsymbol{\Phi}_4^{(\mathrm{e})}(0.5)$ and $\boldsymbol{\Phi}_\mathrm{S} = \boldsymbol{\Phi}_{20}^{(\mathrm{e})}(0.5)$. For the OSTBC $\boldsymbol{\mathcal{G}}_4$, the low-SNR approximation is remarkably accurate for a fairly wide range of $\frac{E_\mathrm{b}^\mathrm{r}}{N_0}_{\min}$, whereas there exists some discrepancy between the Monte Carlo simulation and the first-order approximation for the general input signaling—approximately 11% difference at $\frac{E_\mathrm{b}^\mathrm{r}}{N_0}_{\min} = 0$ dB, for example. In this scenario, the low-SNR slopes are 1.26 and 2.46 bits/s/Hz per 3 dB with and without the OSTBC input signaling constraint, respectively. Thus, the use of the OSTBC $\boldsymbol{\mathcal{G}}_4$ incurs about 49% reduction in the slope. This slope reduction is much smaller than that in a high-SNR regime: the high-SNR slope for the OSTBC $\boldsymbol{\mathcal{G}}_4$ is $\mathcal{R} = 3/4$ and the corresponding slope for the general signaling is equal to $\min(n_\mathrm{T}, n_\mathrm{R}, n_\mathrm{S}) = 4$ bits/s/Hz per 3 dB [21].

V. Conclusion

We investigated the combined effect of rank deficiency and spatial fading correlation on the diversity performance of MIMO systems. In particular, we considered double-scattering MIMO channels employing OSTBCs which use up all antennas to realize full diversity advantage. We characterized the effects of double scattering on the severity of fading and the low-SNR capacity by quantifying the EFF and the capacity slope in terms of the *correlation figures* of spatial correlation matrices. The Schur monotonicity properties were shown for these performance measures as functionals of the eigenvalues of correlation matrices. We also determined the required scattering richness of the channel to achieve the full diversity order of $n_\mathrm{T} n_\mathrm{R}$. Finally, we derived the exact SEP expressions for some classes of double scattering, which consolidate the effects of rank efficiency and spatial correlation on the SEP performance. On account of the generality of channel modeling, the results of the paper are substantial enough to encompass those for well-accepted existing models (e.g., i.i.d./spatially correlated/keyhole MIMO channels) as special cases of our solutions.

Appendix I
Majorization, Schur Monotonicity, and Correlation Matrices

We use the concept of majorization (see, e.g., [65]–[70]) as a mathematical tool to characterize different spatial correlation environments. Using the majorization theory, the analytical framework was established in [52] to assess the performance of multiple-antenna diversity systems with different power dispersion profiles. In particular, monotonicity theorems were proved for various performance measures such as the NSD of the output SNR, the ergodic capacity, the matched-filter bound, the inverse SEP, and the symbol error outage. The notion of majorization has also been used in [18], [36], [71] as a measure of correlation. In this appendix, we briefly discuss the basic properties of majorization and Schur monotonicity.

A. Majorization and Correlation Matrices

Given a real vector $\boldsymbol{a} = (a_1, \ldots, a_n)^T \in \mathbb{R}^n$, we rearrange its components in decreasing order as $a_{[1]} \geq a_{[2]} \geq \cdots \geq a_{[n]}$.

Definition 3: For $\boldsymbol{a} = (a_1, \ldots, a_n)^T, \boldsymbol{b} = (b_1, \ldots, b_n)^T \in \mathbb{R}^n$, we denote $\boldsymbol{a} \prec \boldsymbol{b}$ and say that $\boldsymbol{a}$ is *weakly majorized* (or *submajorized*) by $\boldsymbol{b}$ if

$$\sum_{i=1}^{k} a_{[i]} \leq \sum_{i=1}^{k} b_{[i]}, \quad k = 1, 2, \ldots, n. \tag{48}$$

If $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i$ holds in addition to $\boldsymbol{a} \prec \boldsymbol{b}$, then we say that $\boldsymbol{a}$ is *majorized* by $\boldsymbol{b}$ and denote as $\boldsymbol{a} \preceq \boldsymbol{b}$.

For example, if each $a_i \geq 0$ and $\sum_{i=1}^n a_i = n$, then

$$(1, 1, \ldots, 1)^T \preceq (a_1, a_2, \ldots, a_n)^T \preceq (n, 0, \ldots, 0)^T. \tag{49}$$

The Hardy–Littlewood–Pólya theorem [69, Theorem 2.B.2] argues that $\boldsymbol{a} \preceq \boldsymbol{b}$ if and only if there exist a *doubly stochastic* matrix $\boldsymbol{P}$ such that $\boldsymbol{a} = \boldsymbol{P}\boldsymbol{b}$.[11] Of particular interest are the majorization relations among Hermitian matrices in terms of their eigenvalue vectors to compare different spatial correlation environments. A Hermitian matrix $\boldsymbol{A}$ is said to be *majorized* by a Hermitian matrix $\boldsymbol{B}$, simply denoted by $\boldsymbol{A} \preceq \boldsymbol{B}$, if $\boldsymbol{\lambda}(\boldsymbol{A}) \preceq \boldsymbol{\lambda}(\boldsymbol{B})$ where $\boldsymbol{\lambda}(\cdot)$ denotes the vector of eigenvalues of a Hermitian matrix. For example, the well-known Schur's theorem [70, eq. (5.5.8)] on the relationship between the eigenvalues and diagonal entries of Hermitian matrices can be written as

$$\boldsymbol{A} \circ \boldsymbol{I}_n \preceq \boldsymbol{A} \text{ for Hermitian } \boldsymbol{A} \in \mathbb{C}^{n\times n} \tag{50}$$

where $\circ$ denotes a Hadamard (i.e., entrywise) product. One of the most useful results on the eigenvalue majorization is the following theorem.

Theorem 5 [67, Theorem 7.1]: A linear map $\mathcal{L} : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ is called *positive* if $\mathcal{L}(\boldsymbol{A}) \geq 0$ for $\boldsymbol{A} \in \mathbb{C}^{n\times n} \geq 0$ and *unital* if $\mathcal{L}(\boldsymbol{I}_n) = \boldsymbol{I}_n$. It is said to be *doubly stochastic* if $\mathcal{L}$ is a unital positive linear map with the trace-preserving property, i.e., $\mathrm{tr}\mathcal{L}(\boldsymbol{A}) = \mathrm{tr}(\boldsymbol{A}), \forall \boldsymbol{A} \in \mathbb{C}^{n\times n}$. Let $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ be Hermitian and $\mathcal{L}$ be a doubly stochastic map. Then

$$\mathcal{L}(\boldsymbol{A}) \preceq \boldsymbol{A}. \tag{51}$$

Recall that the Schur product theorem [70, Theorem 5.2.1] says that the Hadamard product of two positive semidefinite matrices is positive semidefinite. Therefore, if $\boldsymbol{\Phi} \in \mathbb{C}^{n\times n}$ is an arbitrary correlation matrix and define $\mathcal{L}(\boldsymbol{A}) = \boldsymbol{A} \circ \boldsymbol{\Phi}$, then $\mathcal{L}$ is obviously a doubly stochastic map on $\mathbb{C}^{n\times n}$.

Corollary 2: Let $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ be Hermitian and $\boldsymbol{\Phi} \in \mathbb{C}^{n\times n}$ be a correlation matrix. Then

$$\boldsymbol{A} \circ \boldsymbol{\Phi} \preceq \boldsymbol{A}. \tag{52}$$

[11]A square matrix is said to be doubly stochastic if all entries of the matrix are nonnegative and the sum of the entries in each row and column is equal to one.

In fact, this result was first given in [68, Corollary 2] without using the notion of doubly stochastic maps. From Corollary 2, we can obtain the eigenvalue majorization relations for the well-known correlation models—*constant*, *exponential*, and *tridiagonal correlation*—which have been widely used for many communication problems of multiple-antenna systems (see, e.g., [21]–[23], [49], [54], [72]).

Example 5 (Constant, Exponential, Tridiagonal Matrices): The nth-order constant, exponential, and tridiagonal matrices with a coefficient ρ, denoted by $\boldsymbol{\Phi}_n^{(\mathrm{c})}(\rho)$, $\boldsymbol{\Phi}_n^{(\mathrm{e})}(\rho)$, and $\boldsymbol{\Phi}_n^{(\mathrm{t})}(\rho)$, respectively, are $n \times n$ symmetric Toeplitz matrices of the structures shown in (53)–(55) at the bottom of the page. Note that $\boldsymbol{\Phi}_n^{(\mathrm{c})}(\rho)$, $\boldsymbol{\Phi}_n^{(\mathrm{e})}(\rho)$ with $\rho \in [0,1]$, and $\boldsymbol{\Phi}_n^{(\mathrm{t})}(\rho)$ with $\rho \in \left[0, 0.5/\cos\frac{\pi}{n+1}\right]$ are correlation matrices, since they are positive semidefinite for such values of ρ. Let $0 \leq \rho_1 \leq \rho_2$. Then, since

$$\boldsymbol{\Phi}_n^{(\mathrm{c})}(\rho_1) = \boldsymbol{\Phi}_n^{(\mathrm{c})}(\rho_2) \circ \boldsymbol{\Phi}_n^{(\mathrm{c})}\left(\frac{\rho_1}{\rho_2}\right)$$
$$\boldsymbol{\Phi}_n^{(\mathrm{e})}(\rho_1) = \boldsymbol{\Phi}_n^{(\mathrm{e})}(\rho_2) \circ \boldsymbol{\Phi}_n^{(\mathrm{e})}\left(\frac{\rho_1}{\rho_2}\right)$$
$$\boldsymbol{\Phi}_n^{(\mathrm{t})}(\rho_1) = \boldsymbol{\Phi}_n^{(\mathrm{t})}(\rho_2) \circ \boldsymbol{\Phi}_n^{(\mathrm{e})}\left(\frac{\rho_1}{\rho_2}\right)$$

it follows from Corollary 2 that

$$\boldsymbol{\Phi}_n^{(\mathrm{c})}(\rho_1) \preceq \boldsymbol{\Phi}_n^{(\mathrm{c})}(\rho_2) \tag{56}$$
$$\boldsymbol{\Phi}_n^{(\mathrm{e})}(\rho_1) \preceq \boldsymbol{\Phi}_n^{(\mathrm{e})}(\rho_2) \tag{57}$$
$$\boldsymbol{\Phi}_n^{(\mathrm{t})}(\rho_1) \preceq \boldsymbol{\Phi}_n^{(\mathrm{t})}(\rho_2). \tag{58}$$

Remark: If $0 \leq \rho_1 \leq \rho_2$, then $\boldsymbol{\Phi}_n^{(\mathrm{c})}\left(\frac{\rho_1}{\rho_2}\right)$ and $\boldsymbol{\Phi}_n^{(\mathrm{e})}\left(\frac{\rho_1}{\rho_2}\right)$ are positive semidefinite. Hence, the majorization relations (56)–(58) hold, although each matrix itself is only Hermitian but may not be positive semidefinite.

B. Schur Monotonicity

The concept of majorization is closely related to a MIS (or MDS) function. If a function f : (a subset of) $\mathbb{R}^n \to \mathbb{R}$ satisfies $f(a_1, \ldots, a_n) \leq f(b_1, \ldots, b_n)$ whenever $\boldsymbol{a} \preceq \boldsymbol{b}$, then f is

$$\boldsymbol{\Phi}_n^{(\mathrm{c})}(\rho) = \begin{bmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \cdots & \rho \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \cdots & 1 \end{bmatrix}_{n\times n} \tag{53}$$

$$\boldsymbol{\Phi}_n^{(\mathrm{e})}(\rho) = \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{(n-1)} \\ \rho & 1 & \rho & \cdots & \rho^{(n-2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{bmatrix}_{n\times n} \tag{54}$$

$$\boldsymbol{\Phi}_n^{(\mathrm{t})}(\rho) = \begin{bmatrix} 1 & \rho & & & & 0 \\ \rho & 1 & \rho & & & \\ & \rho & 1 & \rho & & \\ & & \ddots & \ddots & \ddots & \\ & & & \rho & 1 & \rho \\ 0 & & & & \rho & 1 \end{bmatrix}_{n\times n}. \tag{55}$$

called a MIS (or isotone) function on (a subset of) $\mathbb{R}^n$. The following theorem gives a necessary and sufficient condition for f to be MIS.

Theorem 6 (Schur 1923): Let $\mathbb{I} \subset \mathbb{R}$ and $f : \mathbb{I}^n \to \mathbb{R}$ be continuously differentiable. Then, the function f is MIS on $\mathbb{I}^n$ if and only if

$$f \text{ is symmetric on } \mathbb{I}^n \tag{59}$$

and for all $i \neq j$

$$(a_i - a_j)\left[\frac{\partial f}{\partial a_i} - \frac{\partial f}{\partial a_j}\right] \geq 0, \quad \forall\, \boldsymbol{a} \in \mathbb{I}^n. \tag{60}$$

Note that Schur's condition (60) can be replaced by

$$(a_1 - a_2)\left[\frac{\partial f}{\partial a_1} - \frac{\partial f}{\partial a_2}\right] \geq 0, \quad \forall\, \boldsymbol{a} \in \mathbb{I}^n \tag{61}$$

because of the symmetry. If f is MIS on $\mathbb{I}^n$, then $-f$ is a MDS function on $\mathbb{I}^n$.

APPENDIX II
SOME STATISTICS DERIVED FROM COMPLEX GAUSSIAN MATRICES

This appendix gives useful results on some statistics derived from complex Gaussian matrices.

A. Preliminary Results

Lemma 1: Let $\boldsymbol{X}_k \sim \tilde{\mathcal{N}}_{m,n}(\boldsymbol{0}_{m\times n}, \boldsymbol{\Sigma}, \boldsymbol{\Psi}_k)$, $k = 1, \ldots, p$, be statistically independent complex Gaussian matrices and

$$\boldsymbol{X} = [\,\boldsymbol{X}_1 \quad \ldots \quad \boldsymbol{X}_p\,] \sim \tilde{\mathcal{N}}_{m,np}\left(\boldsymbol{0}_{m\times np}, \boldsymbol{\Sigma}, \bigoplus_{k=1}^{p} \boldsymbol{\Psi}_k\right). \tag{62}$$

Then, for $\boldsymbol{A} \in \mathbb{C}^{m\times m} \geq 0$ and $\boldsymbol{B} = \bigoplus_{k=1}^{p} \boldsymbol{B}_k$, $\boldsymbol{B}_k \in \mathbb{C}^{n\times n} \geq 0$, we have

$$\mathbb{E}\left\{\operatorname{etr}(-\boldsymbol{A}\boldsymbol{X}\boldsymbol{B}\boldsymbol{X}^\dagger)\right\} = \prod_{k=1}^{p} \det(\boldsymbol{I}_{mn} + \boldsymbol{A}\boldsymbol{\Sigma} \otimes \boldsymbol{\Psi}_k \boldsymbol{B}_k)^{-1}. \tag{63}$$

Proof: Since $\boldsymbol{A}\boldsymbol{X}\boldsymbol{B}\boldsymbol{X}^\dagger = \sum_{k=1}^{p} \boldsymbol{A}\boldsymbol{X}_k\boldsymbol{B}_k\boldsymbol{X}_k^\dagger$, we have

$$\mathbb{E}\left\{\operatorname{etr}(-\boldsymbol{A}\boldsymbol{X}\boldsymbol{B}\boldsymbol{X}^\dagger)\right\} = \prod_{k=1}^{p} \mathbb{E}_{\boldsymbol{X}_k}\left\{\operatorname{etr}(-\boldsymbol{A}\boldsymbol{X}_k\boldsymbol{B}_k\boldsymbol{X}_k^\dagger)\right\}. \tag{64}$$

Therefore[12]

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{X}_k}\left\{\operatorname{etr}(-\boldsymbol{A}\boldsymbol{X}_k\boldsymbol{B}_k\boldsymbol{X}_k^\dagger)\right\} \\
&\quad = c_k \int_{\boldsymbol{X}_k} \operatorname{etr}(-\boldsymbol{A}\boldsymbol{X}_k\boldsymbol{B}_k\boldsymbol{X}_k^\dagger - \boldsymbol{\Sigma}^{-1}\boldsymbol{X}_k\boldsymbol{\Psi}_k^{-1}\boldsymbol{X}_k^\dagger)\, d\boldsymbol{X}_k \\
&\quad = c_k \int_{\boldsymbol{X}_k} \exp\left\{-(\operatorname{vec}(\boldsymbol{X}_k^\dagger))^\dagger \boldsymbol{T}_k \operatorname{vec}(\boldsymbol{X}_k^\dagger)\right\} d\boldsymbol{X}_k \\
&\quad = c_k \pi^{mn} \det(\boldsymbol{T}_k)^{-1} \\
&\quad = \det(\boldsymbol{I}_{mn} + \boldsymbol{A}\boldsymbol{\Sigma} \otimes \boldsymbol{\Psi}_k\boldsymbol{B}_k)^{-1}
\end{aligned} \tag{65}$$

[12]If $\boldsymbol{X} = (X_{ij})$ is an $m \times n$ matrix of functionally independent complex variables, then

$$d\boldsymbol{X} = \prod_{i=1}^{m}\prod_{j=1}^{n} d\,\mathfrak{Re}X_{ij}\, d\,\Im m X_{ij}.$$

where $c_k = \pi^{-mn} \det(\boldsymbol{\Sigma})^{-n} \det(\boldsymbol{\Psi}_k)^{-m}$ and

$$\boldsymbol{T}_k = (\boldsymbol{\Sigma}^T \otimes \boldsymbol{\Psi}_k)^{-1} + \boldsymbol{A}^T \otimes \boldsymbol{B}_k.$$

Combining (64) and (65) completes the proof. □

Lemma 2: Let $\boldsymbol{X} \sim \tilde{\mathcal{N}}_{m,n}(\boldsymbol{0}_{m\times n}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$. Then, for $m \times n$ matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, we have

$$\mathbb{E}\left\{\operatorname{etr}(\boldsymbol{X}^\dagger\boldsymbol{A} + \boldsymbol{B}^\dagger\boldsymbol{X})\right\} = \operatorname{etr}(\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Psi}\boldsymbol{B}^\dagger). \tag{66}$$

Proof: Let $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ be $m \times n$ matrices such that

$$\begin{aligned}
&\operatorname{tr}\left(\boldsymbol{X}^\dagger\boldsymbol{A} + \boldsymbol{B}^\dagger\boldsymbol{X} - \boldsymbol{\Sigma}^{-1}\boldsymbol{X}\boldsymbol{\Psi}^{-1}\boldsymbol{X}^\dagger\right) \\
&\quad = \operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{M}_1\boldsymbol{\Psi}^{-1}\boldsymbol{M}_2^\dagger\right) \\
&\qquad + \operatorname{tr}\left\{-\boldsymbol{\Sigma}^{-1}(\boldsymbol{X} - \boldsymbol{M}_1)\boldsymbol{\Psi}^{-1}(\boldsymbol{X} - \boldsymbol{M}_2)^\dagger\right\}.
\end{aligned} \tag{67}$$

Then, since

$$\begin{aligned}
&\int_{\boldsymbol{X}} \operatorname{etr}\left\{-\boldsymbol{\Sigma}^{-1}(\boldsymbol{X} - \boldsymbol{M}_1)\boldsymbol{\Psi}^{-1}(\boldsymbol{X} - \boldsymbol{M}_2)^\dagger\right\} d\boldsymbol{X} \\
&\qquad = \pi^{mn}\det(\boldsymbol{\Sigma})^n \det(\boldsymbol{\Psi})^m
\end{aligned} \tag{68}$$

we get

$$\mathbb{E}\left\{\operatorname{etr}(\boldsymbol{X}^\dagger\boldsymbol{A} + \boldsymbol{B}^\dagger\boldsymbol{X})\right\} = \operatorname{etr}(\boldsymbol{\Sigma}^{-1}\boldsymbol{M}_1\boldsymbol{\Psi}^{-1}\boldsymbol{M}_2^\dagger). \tag{69}$$

By comparing both the sides of (67), we have

$$\boldsymbol{M}_1 = \boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Psi} \tag{70}$$

$$\boldsymbol{M}_2 = \boldsymbol{\Sigma}\boldsymbol{B}\boldsymbol{\Psi}. \tag{71}$$

Finally, substituting (70) and (71) into (69) completes the proof. □

Lemma 3: Let $\boldsymbol{X} \sim \tilde{\mathcal{N}}_{m,n}(\boldsymbol{M}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$. Then, the characteristic function of $\boldsymbol{X}$ is

$$\begin{aligned}
\Phi_{\boldsymbol{X}}(\boldsymbol{Z}) &\triangleq \mathbb{E}\left\{\exp\left[\jmath\, \mathfrak{Re}\operatorname{tr}(\boldsymbol{X}\boldsymbol{Z}^\dagger)\right]\right\} \\
&= \exp\left[\jmath\, \mathfrak{Re}\operatorname{tr}(\boldsymbol{M}\boldsymbol{Z}^\dagger) - \frac{1}{4}\operatorname{tr}(\boldsymbol{\Sigma}\boldsymbol{Z}\boldsymbol{\Psi}\boldsymbol{Z}^\dagger)\right]
\end{aligned} \tag{72}$$

where $\jmath = \sqrt{-1}$ and $\boldsymbol{Z} \in \mathbb{C}^{m\times n}$ is an arbitrary matrix.

Proof: Let $\boldsymbol{X}_1 \sim \tilde{\mathcal{N}}_{m,n}(\boldsymbol{0}_{m\times n}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$. Then

$$\Phi_{\boldsymbol{X}}(\boldsymbol{Z}) = \exp\left[\jmath\, \mathfrak{Re}\operatorname{tr}(\boldsymbol{M}\boldsymbol{Z}^\dagger)\right] \mathbb{E}\left\{\exp\left[\jmath\, \mathfrak{Re}\operatorname{tr}(\boldsymbol{X}_1\boldsymbol{Z}^\dagger)\right]\right\}. \tag{73}$$

Since

$$\mathfrak{Re}\operatorname{tr}(\boldsymbol{X}_1\boldsymbol{Z}^\dagger) = \frac{1}{2}\operatorname{tr}(\boldsymbol{Z}^\dagger\boldsymbol{X}_1 + \boldsymbol{X}_1^\dagger\boldsymbol{Z}) \tag{74}$$

it follows from Lemma 2 that

$$\mathbb{E}\left\{\exp\left[\jmath\, \mathfrak{Re}\operatorname{tr}(\boldsymbol{X}_1\boldsymbol{Z}^\dagger)\right]\right\} = \operatorname{etr}\left(-\frac{1}{4}\boldsymbol{\Sigma}\boldsymbol{Z}\boldsymbol{\Psi}\boldsymbol{Z}^\dagger\right). \tag{75}$$

Combining (73) and (75) completes the proof. □

We remark that Lemma 3 is a counterpart result of the real case in [73, Theorem 2.3.2].

B. Hypergeometric Functions of Matrix Arguments

The hypergeometric functions of matrix arguments often appear in deriving the distributions and statistics of random matrices [73]–[77]. In parallel to the hypergeometric functions of a scalar argument, the hypergeometric functions of one or two matrix arguments can be expressed as an infinite series of zonal polynomials[13]

$$\begin{aligned}{}_p\tilde{F}_q\left(a_1,a_2,\ldots,a_p;b_1,b_2,\ldots,b_q;\boldsymbol{A}\right) &= \sum_{k=0}^{\infty}\sum_{\kappa}\frac{[a_1]_\kappa[a_2]_\kappa\cdots[a_p]_\kappa}{[b_1]_\kappa[b_2]_\kappa\cdots[b_q]_\kappa}\frac{\tilde{C}_\kappa(\boldsymbol{A})}{k!} \quad (76)\\ {}_p\tilde{F}_q^{(n)}\left(a_1,a_2,\ldots,a_p;b_1,b_2,\ldots,b_q;\boldsymbol{A},\boldsymbol{B}\right) &= \sum_{k=0}^{\infty}\sum_{\kappa}\frac{[a_1]_\kappa[a_2]_\kappa\cdots[a_p]_\kappa}{[b_1]_\kappa[b_2]_\kappa\cdots[b_q]_\kappa}\frac{\tilde{C}_\kappa(\boldsymbol{A})\tilde{C}_\kappa(\boldsymbol{B})}{k!\tilde{C}_\kappa(\boldsymbol{I}_n)} \quad (77)\end{aligned}$$

with Hermitian $\boldsymbol{A}\in\mathbb{C}^{n\times n}$ and $\boldsymbol{B}\in\mathbb{C}^{n\times n}$. In (76) and (77), $\kappa=(k_1,k_2,\ldots,k_n)$ denotes a partition of the nonnegative integer k such that $k_1\ge k_2\ge\cdots\ge k_n\ge 0$ and $\sum_{i=1}^n k_i=k$, $[a]_\kappa$ is the complex multivariate hypergeometric coefficient of the partition κ [75, eq. (84)], and $\tilde{C}_\kappa(\cdot)$ is the zonal polynomial of a Hermitian matrix [75, eq. (85)]. Although these functions are of great interest from an analytical point of view, the practical difficulty lies in their numerical aspects. The determinantal representation for the hypergeometric function of two Hermitian matrices [77, Lemma 3] settles this computational problem and has been widely used in the literature of multiple-antenna communication theory (see, e.g., [22], [23], [55], [56]). However, [77, Lemma 3] is valid only for the case of two matrix arguments with the same dimension and the distinct eigenvalues. In the following lemma, we generalize [77, Lemma 3] for the case that two matrix arguments have different matrix dimension and eigenvalues of arbitrary multiplicity.

Lemma 4: Let $\boldsymbol{\Lambda}\in\mathbb{C}^{m\times m}$ and $\boldsymbol{\Sigma}\in\mathbb{C}^{n\times n}$, $m\le n$, be Hermitian matrices with the ordered eigenvalues $\lambda_1\ge\lambda_2\ge\cdots\ge\lambda_m$ and $\sigma_1\ge\sigma_2\ge\cdots\ge\sigma_n$, respectively. Given $a_i,b_j\in\mathbb{C}$ where $i=1,2,\ldots,p$ and $j=1,2,\ldots,q$, define

$$\mathcal{H}_{p,q}^{n,\nu}(x)\triangleq{}_pF_q(a_1-n+\nu,\ldots,a_p-n+\nu;\; b_1-n+\nu,\ldots,b_q-n+\nu;x) \quad (78)$$

$$\chi_{p,q}^{n,\nu}\triangleq\frac{\prod_{j=1}^q(b_j-n+1)_\nu}{\prod_{i=1}^p(a_i-n+1)_\nu} \quad (79)$$

where ν is an arbitrary nonnegative integer, $(a)_n=a(a+1)\cdots(a+n-1)$, $(a)_0=1$ is the Pochhammer symbol, and ${}_pF_q(a_1,a_2,\ldots,a_p;b_1,b_2,\ldots,b_q;z)$ is the generalized hypergeometric function of scalar argument [79, eq. (9.14.1)]. Then, we have a generic determinantal expression for the hypergeometric function of two Hermitian matrices shown as (80) at the bottom of the page. In (80)

$$K_{p,q}^{m,n}=\prod_{i=1}^m\chi_{p,q}^{n,n-i}\cdot(n-i)! \quad (81)$$

and $\boldsymbol{\mathcal{Y}}_k=(\mathcal{Y}_{k,ij})$ and $\boldsymbol{\mathcal{Z}}_{(l),k}=(\mathcal{Z}_{(l),k,ij})$, $l\le n$, $k=1,2,\ldots,\varrho(\boldsymbol{\Sigma})$, are $m\times\tau_k(\boldsymbol{\Sigma})$ and $l\times\tau_k(\boldsymbol{\Sigma})$ matrices, whose (i,j)th entries are given, respectively, by

$$\mathcal{Y}_{k,ij}=\frac{\lambda_i^{j-1}}{\chi_{p,q}^{n,j-1}}\cdot\mathcal{H}_{p,q}^{n,j}\left(\lambda_i\sigma_{\langle k\rangle}\right) \quad (82)$$

$$\mathcal{Z}_{(l),k,ij}=(i-j+1)_{j-1}\,\sigma_{\langle k\rangle}^{i-j}. \quad (83)$$

In particular, for ${}_0\tilde{F}_0^{(n)}(\boldsymbol{\Lambda},\boldsymbol{\Sigma})$, $K_{p,q}^{m,n}$ in (81) and the (i,j)th entry of $\boldsymbol{\mathcal{Y}}_k$ in (82) reduce to

$$K_{0,0}^{m,n}=\prod_{i=1}^m(n-i)! \quad (84)$$

$$\mathcal{Y}_{k,ij}=\lambda_i^{j-1}e^{\lambda_i\sigma_{\langle k\rangle}}. \quad (85)$$

Proof: Let us dilate the $m\times m$ matrix $\boldsymbol{\Lambda}$ to the $n\times n$ matrix $\boldsymbol{\Lambda}\oplus\boldsymbol{0}_{n-m}$ by affixing zero elements. Then, this augmented matrix $\boldsymbol{\Lambda}\oplus\boldsymbol{0}_{n-m}$ has the eigenvalues $\lambda_1,\lambda_2,\ldots,\lambda_m$ and $(n-m)$ additional zero eigenvalues. Note that zonal polynomials depend on its Hermitian matrix arguments through Schur functions in the eigenvalues of matrix arguments [75]–[78]. Since Schur functions are invariant to augmenting zero elements [80], it is easy to show that

$$\tilde{C}_\kappa(\boldsymbol{\Lambda}\oplus\boldsymbol{0}_{n-m})=\tilde{C}_\kappa(\boldsymbol{\Lambda}). \quad (86)$$

Let $\lambda_{m+1},\lambda_{m+2},\ldots,\lambda_n$ be $(n-m)$ additional zero eigenvalues and denote the left-hand side of (80) by $\mathrm{LHS}_{(80)}$ for convenience. Then, it follows from (86) and [7, Lemma 3] that

$$\mathrm{LHS}_{(80)}=K_{p,q}^{n,n}\frac{\det_{1\le i,j\le n}\left(\mathcal{H}_{p,q}^{n,1}(\lambda_i\sigma_j)\right)}{\prod_{i<j}^n(\lambda_j-\lambda_i)(\sigma_j-\sigma_i)}\cdot\prod_{i<j}^m(\lambda_j-\lambda_i). \quad (87)$$

From a computational point of view, (87) presents numerical difficulty since the Vandermonde determinant $\prod_{i<j}^n(\lambda_j-\lambda_i)$ or $\prod_{i<j}^n(\sigma_j-\sigma_i)$ becomes zero when some of the λ_i's or σ_i's are equal. This can be alleviated by using Cauchy's mean value theorem (or L'Hôspital's rule)

$$\mathrm{LHS}_{(80)}=K_{p,q}^{n,n}\lim_{\boldsymbol{\sigma}\to\tilde{\boldsymbol{\sigma}}}\lim_{\{\lambda_k\}_{k=m+1}^n\to 0}\left\{\prod_{i<j}^m(\lambda_j-\lambda_i)\times\frac{\det_{1\le i,j\le n}\left(\mathcal{H}_{p,q}^{n,1}(\lambda_i\sigma_j)\right)}{\prod_{i<j}^n(\lambda_j-\lambda_i)(\sigma_j-\sigma_i)}\right\} \quad (88)$$

[13]Zonal polynomials of a symmetric matrix were introduced in [74] using group representation theory. In parallel to a real matrix argument, zonal polynomials of a Hermitian matrix were defined in [75] as natural extension of the real case. Those polynomials are homogeneous symmetric functions in the eigenvalues of matrix argument and can be constructed in terms of homogeneous symmetric polynomials such as monomial symmetric functions, elementary symmetric functions, and Schur functions [78].

$${}_p\tilde{F}_q^{(n)}(a_1,\ldots,a_p;b_1,\ldots,b_q;\boldsymbol{\Lambda},\boldsymbol{\Sigma})\prod_{i<j}^m(\lambda_j-\lambda_i)=\frac{K_{p,q}^{m,n}}{\det(\boldsymbol{\Lambda})^{n-m}}\frac{\det\left(\begin{bmatrix}\boldsymbol{\mathcal{Z}}_{(n-m),1} & \boldsymbol{\mathcal{Z}}_{(n-m),2} & \cdots & \boldsymbol{\mathcal{Z}}_{(n-m),\varrho(\boldsymbol{\Sigma})}\\ \boldsymbol{\mathcal{Y}}_1 & \boldsymbol{\mathcal{Y}}_2 & \cdots & \boldsymbol{\mathcal{Y}}_{\varrho(\boldsymbol{\Sigma})}\end{bmatrix}\right)}{\det\left(\begin{bmatrix}\boldsymbol{\mathcal{Z}}_{(n),1} & \boldsymbol{\mathcal{Z}}_{(n),2} & \cdots & \boldsymbol{\mathcal{Z}}_{(n),\varrho(\boldsymbol{\Sigma})}\end{bmatrix}\right)}. \quad (80)$$

where $\boldsymbol{\sigma} \rightarrow \tilde{\boldsymbol{\sigma}}$ means that

$$\{\sigma_i\}_{i=1}^{\tau_1(\boldsymbol{\Sigma})} \rightarrow \sigma_{\langle 1\rangle},$$
$$\{\sigma_i\}_{i=\tau_1(\boldsymbol{\Sigma})+1}^{\tau_1(\boldsymbol{\Sigma})+\tau_2(\boldsymbol{\Sigma})} \rightarrow \sigma_{\langle 2\rangle},$$
$$\vdots$$
$$\{\sigma_i\}_{i=n-\tau_{\varrho(\boldsymbol{\Sigma})}(\boldsymbol{\Sigma})+1}^{n} \rightarrow \sigma_{\langle \varrho(\boldsymbol{\Sigma})\rangle}.$$

Let n-dimensional vectors $\boldsymbol{u}(z)$ and $\boldsymbol{v}(z)$ be

$$\boldsymbol{u}(z) = \left(\mathcal{H}_{p,q}^{n,1}(\sigma_1 z), \mathcal{H}_{p,q}^{n,1}(\sigma_2 z), \ldots, \mathcal{H}_{p,q}^{n,1}(\sigma_n z)\right) \tag{89}$$
$$\boldsymbol{v}(z) = \left(1, z, \ldots, z^{n-1}\right) \tag{90}$$

and let $\boldsymbol{u}^{(k)}(z)$ and $\boldsymbol{v}^{(k)}(z)$ be the kth derivatives of $\boldsymbol{u}(z)$ and $\boldsymbol{v}(z)$ with respect to z, respectively. Note that the jth components $u_j^{(k)}(z)$ and $v_j^{(k)}(z)$, $j = 1, 2, \ldots, n$, of $\boldsymbol{u}^{(k)}(z)$ and $\boldsymbol{v}^{(k)}(z)$ are given, respectively, by

$$u_j^{(k)}(z) = \frac{\sigma_j^k}{\chi_{p,q}^{n,k}} \cdot \mathcal{H}_{p,q}^{n,k+1}(\sigma_j z) \tag{91}$$
$$v_j^{(k)}(z) = (j-k)_k \, z^{j-k-1} \tag{92}$$

where (91) follows from the differentiation identity of [81, eq. (7.2.3.47)]. Then, taking the limits on λ_k's, we get

$$\lim_{\{\lambda_k\}_{k=m+1}^{n} \rightarrow 0} \frac{\det\limits_{1\le i,j\le n}\left(\mathcal{H}_{p,q}^{n,1}(\lambda_i \sigma_j)\right)}{\prod_{i<j}^{n}(\lambda_j - \lambda_i)} = \frac{\det\left(\begin{bmatrix}\boldsymbol{U}_{\mathrm{A}} \\ \boldsymbol{U}_{\mathrm{B}}\end{bmatrix}\right)}{\det\left(\begin{bmatrix}\boldsymbol{V}_{\mathrm{A}} \\ \boldsymbol{V}_{\mathrm{B}}\end{bmatrix}\right)} \tag{93}$$

with the $(n-m) \times n$ matrices

$$\boldsymbol{U}_{\mathrm{A}} = (U_{\mathrm{A},ij}) = \begin{bmatrix} \boldsymbol{u}^{(0)}(0) \\ \boldsymbol{u}^{(1)}(0) \\ \vdots \\ \boldsymbol{u}^{(n-m-1)}(0) \end{bmatrix} \tag{94}$$

$$\boldsymbol{V}_{\mathrm{A}} = (V_{\mathrm{A},ij}) = \begin{bmatrix} \boldsymbol{v}^{(0)}(0) \\ \boldsymbol{v}^{(1)}(0) \\ \vdots \\ \boldsymbol{v}^{(n-m-1)}(0) \end{bmatrix} \tag{95}$$

and the $m \times n$ matrices $\boldsymbol{U}_{\mathrm{B}} = \left(\mathcal{H}_{p,q}^{n,1}(\lambda_i \sigma_j)\right)$ and $\boldsymbol{V}_{\mathrm{B}} = \left(\lambda_i^{j-1}\right)$. From (91) and (92), it is easy to see that the (i,j)th entries of $\boldsymbol{U}_{\mathrm{A}}$ and $\boldsymbol{V}_{\mathrm{A}}$ are given, respectively, by

$$U_{\mathrm{A},ij} = u_j^{(i-1)}(0) = \frac{\sigma_j^{i-1}}{\chi_{p,q}^{n,i-1}} \tag{96}$$

$$V_{\mathrm{A},ij} = v_j^{(i-1)}(0) = \begin{cases} (i-1)!, & \text{if } i = j \\ 0, & \text{otherwise.} \end{cases} \tag{97}$$

Now, using the result on the determinant of a partitioned matrix for invertible $\boldsymbol{A}$

$$\det\left(\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D}\end{bmatrix}\right) = \det(\boldsymbol{A}) \det\left(\boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\right) \tag{98}$$

we have

$$\det\left(\begin{bmatrix}\boldsymbol{V}_{\mathrm{A}} \\ \boldsymbol{V}_{\mathrm{B}}\end{bmatrix}\right) = \det\left(\begin{bmatrix} \lambda_1^{n-m} & \cdots & \lambda_1^{n-1} \\ \lambda_2^{n-m} & \cdots & \lambda_2^{n-1} \\ \vdots & \ddots & \vdots \\ \lambda_m^{n-m} & \cdots & \lambda_m^{n-1} \end{bmatrix}\right) \prod_{l=1}^{n-m}(l-1)!$$
$$= \prod_{l=1}^{n-m}(l-1)! \prod_{k=1}^{m} \lambda_k^{n-m} \prod_{i<j}^{m}(\lambda_j - \lambda_i). \tag{99}$$

Hence, combining (88), (93), and (99) gives

$$\mathrm{LHS}_{(80)} = \frac{K_{p,q}^{m,n}}{\det(\boldsymbol{\Lambda})^{n-m}} \lim_{\boldsymbol{\sigma} \rightarrow \tilde{\boldsymbol{\sigma}}} \frac{\det\left(\begin{bmatrix}\tilde{\boldsymbol{U}}_{\mathrm{A}} \\ \boldsymbol{U}_{\mathrm{B}}\end{bmatrix}\right)}{\prod_{i<j}^{n}(\sigma_j - \sigma_i)} \tag{100}$$

where $\tilde{\boldsymbol{U}}_{\mathrm{A}} = \left(\sigma_j^{i-1}\right)$ is the $(n-m) \times n$ submatrix of the Vandermonde matrix of $\sigma_1, \sigma_2, \ldots, \sigma_n$.

Using similar steps leading to (93), we obtain (101) at the bottom of the page, where the (i,j)th entries of $m \times \tau_k(\boldsymbol{\Sigma})$ matrices $\boldsymbol{\mathcal{Y}}_k$ and $l \times \tau_k(\boldsymbol{\Sigma})$ matrices $\boldsymbol{\mathcal{Z}}_{(l),k}$, $l \le n$, $k = 1, 2, \ldots, \varrho(\boldsymbol{\Sigma})$, are given by (82) and (83), respectively. Finally, substituting (101) into (100) completes the proof of the lemma. □

As a by-product of Lemma 4, we obtain the following determinantal formula for the hypergeometric function of one matrix argument.

Corollary 3: If $\boldsymbol{\Sigma} = \boldsymbol{I}_n$ in Lemma 4, then we have

$${}_p\tilde{F}_q(a_1, \ldots, a_p; b_1, \ldots, b_q; \boldsymbol{\Lambda}) \cdot \prod_{i<j}^{m}(\lambda_j - \lambda_i)$$
$$= \det_{1\le i,j\le m}\left(\lambda_i^{j-1}\mathcal{H}_{p,q}^{n,n-m+j}(\lambda_i)\right). \tag{102}$$

Proof: The result follows immediately from (98) and Lemma 4 with $\varrho(\boldsymbol{\Sigma}) = 1$, $\tau_1(\boldsymbol{\Sigma}) = n$, and $\sigma_{\langle 1\rangle} = 1$. □

C. Some Statistics

Lemma 5: Let $\boldsymbol{X} \sim \tilde{\mathcal{N}}_{m,n}(\boldsymbol{0}_{m\times n}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$. Then, for $\boldsymbol{A} \in \mathbb{C}^{m\times m} \ge 0$ and $\boldsymbol{B} \in \mathbb{C}^{n\times n} \ge 0$, the kth-order cumulant of $\mathrm{tr}(\boldsymbol{AXBX}^\dagger)$ is

$$\mathbb{C}\mathrm{um}_k\left\{\mathrm{tr}(\boldsymbol{AXBX}^\dagger)\right\} \triangleq (-1)^k \frac{d^k}{ds^k} \ln \phi_{\mathrm{tr}(\boldsymbol{AXBX}^\dagger)}(s)\Big|_{s=0}$$
$$= (k-1)!\,\mathrm{tr}\left\{(\boldsymbol{A\Sigma})^k\right\}\mathrm{tr}\left\{(\boldsymbol{\Psi B})^k\right\} \tag{103}$$

where $\phi_{\mathrm{tr}(\boldsymbol{AXBX}^\dagger)}(s) \triangleq \mathbb{E}\left\{\mathrm{etr}\left(-s\boldsymbol{AXBX}^\dagger\right)\right\}$ is the MGF of $\mathrm{tr}(\boldsymbol{AXBX}^\dagger)$.

Proof: Since

$$\mathrm{tr}(\boldsymbol{AXBX}^\dagger) = \left(\mathrm{vec}(\boldsymbol{X}^\dagger)\right)^\dagger (\boldsymbol{A}^T \otimes \boldsymbol{B})\mathrm{vec}(\boldsymbol{X}^\dagger)$$

$$\lim_{\boldsymbol{\sigma} \rightarrow \tilde{\boldsymbol{\sigma}}} \frac{\det\left(\begin{bmatrix}\tilde{\boldsymbol{U}}_{\mathrm{A}} \\ \boldsymbol{U}_{\mathrm{B}}\end{bmatrix}\right)}{\prod_{i<j}^{n}(\sigma_j - \sigma_i)} = \frac{\det\left(\begin{bmatrix} \boldsymbol{\mathcal{Z}}_{(n-m),1} & \boldsymbol{\mathcal{Z}}_{(n-m),2} & \cdots & \boldsymbol{\mathcal{Z}}_{(n-m),\varrho(\boldsymbol{\Sigma})} \\ \boldsymbol{\mathcal{Y}}_1 & \boldsymbol{\mathcal{Y}}_2 & \cdots & \boldsymbol{\mathcal{Y}}_{\varrho(\boldsymbol{\Sigma})} \end{bmatrix}\right)}{\det\left(\begin{bmatrix} \boldsymbol{\mathcal{Z}}_{(n),1} & \boldsymbol{\mathcal{Z}}_{(n),2} & \cdots & \boldsymbol{\mathcal{Z}}_{(n),\varrho(\boldsymbol{\Sigma})} \end{bmatrix}\right)} \tag{101}$$

is a quadratic form in complex Gaussian variables, whose characteristic function has been reported in [82], it can be readily shown that

$$\phi_{\mathrm{tr}(\boldsymbol{AXBX}^\dagger)}(s) = \det\left\{\boldsymbol{I}_{mn} + s(\boldsymbol{\Sigma}^T \otimes \boldsymbol{\Psi})(\boldsymbol{A}^T \otimes \boldsymbol{B})\right\}^{-1} = \det\left(\boldsymbol{I}_{mn} + s\boldsymbol{A\Sigma} \otimes \boldsymbol{\Psi B}\right)^{-1}. \quad (104)$$

Therefore

$$\frac{d^k}{ds^k}\ln\phi_{\mathrm{tr}(\boldsymbol{AXBX}^\dagger)}(s) = (-1)^k (k-1)! \times \mathrm{tr}\left\{\left[\left(\boldsymbol{I}_{mn} + s\boldsymbol{A\Sigma}\otimes\boldsymbol{\Psi B}\right)^{-1}\left(\boldsymbol{A\Sigma}\otimes\boldsymbol{\Psi B}\right)\right]^k\right\}. \quad (105)$$

Hence, we obtain the result (103) from (105) with $s = 0$. □

We remark that the cumulants, except for the first-order cumulant, are invariant with respect to translations of a random variable. The first- and second-order cumulants are the mean and variance of the underlying random variable, respectively, and other higher order statistics can also be obtained from general relationships between the cumulants and moments. Lemma 5 reveals that all cumulants of $\mathrm{tr}\left(\boldsymbol{AXBX}^\dagger\right)$ as functionals of the eigenvalues of $\boldsymbol{A\Sigma}$ and $\boldsymbol{\Psi B}$ are MIS.

Lemma 6: Let $\boldsymbol{X} \sim \tilde{\mathcal{N}}_{m,n}(\boldsymbol{0}_{m\times n}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$. Then, for $\boldsymbol{A} \in \mathbb{C}^{m\times m} \geq 0$ and $\boldsymbol{B} \in \mathbb{C}^{n\times n} \geq 0$, we have

$$\mathbb{E}\left\{\mathrm{tr}\left[(\boldsymbol{AXBX}^\dagger)^2\right]\right\} = \mathrm{tr}^2(\boldsymbol{A\Sigma})\,\mathrm{tr}\left\{(\boldsymbol{\Psi B})^2\right\} + \mathrm{tr}^2(\boldsymbol{\Psi B})\,\mathrm{tr}\left\{(\boldsymbol{A\Sigma})^2\right\}. \quad (106)$$

Proof: We first start with the characteristic function of $\boldsymbol{S} = (S_{ij}) = \boldsymbol{A}^{1/2}\boldsymbol{XB}^{1/2}$. Let $\tilde{\boldsymbol{\Sigma}} = (\tilde{\Sigma}_{ij}) = \boldsymbol{A}^{1/2}\boldsymbol{\Sigma A}^{1/2}$ and $\tilde{\boldsymbol{\Psi}} = (\tilde{\Psi}_{ij}) = \boldsymbol{B}^{1/2}\boldsymbol{\Psi B}^{1/2}$. Then

$$\begin{aligned}\Phi_{\boldsymbol{S}}(\boldsymbol{Z}) &= \mathbb{E}\left\{\exp\left[\jmath\,\Re\mathfrak{e}\,\mathrm{tr}\left(\boldsymbol{A}^{1/2}\boldsymbol{XB}^{1/2}\boldsymbol{Z}^\dagger\right)\right]\right\}\\ &= \Phi_{\boldsymbol{X}}\left(\boldsymbol{A}^{1/2}\boldsymbol{ZB}^{1/2}\right)\\ &\overset{(a)}{=} \mathrm{etr}\left(-\frac{1}{4}\tilde{\boldsymbol{\Sigma}}\boldsymbol{Z}\tilde{\boldsymbol{\Psi}}\boldsymbol{Z}^\dagger\right)\\ &= e^{\varphi(\boldsymbol{Z})}\end{aligned} \quad (107)$$

where (a) follows from Lemma 3 and

$$\varphi(\boldsymbol{Z}) = -\frac{1}{4}\sum_{i=1}^{m}\sum_{p=1}^{m}\sum_{q=1}^{n}\sum_{j=1}^{n}\tilde{\Sigma}_{ip}Z_{pj}\tilde{\Psi}_{jq}Z_{iq}^*. \quad (108)$$

It follows from the characteristic function $\Phi_{\boldsymbol{S}}(\boldsymbol{Z})$ in (107) that

$$\begin{aligned}&\mathbb{E}\left\{S_{i_1j_1}S_{i_2j_2}^*S_{i_3j_3}S_{i_4j_4}^*\right\}\\ &= \frac{1}{\jmath^4}\left.\frac{\partial\Phi_{\boldsymbol{S}}(\boldsymbol{Z})}{\partial Z_{i_1j_1}\partial Z_{i_2j_2}^*\partial Z_{i_3j_3}\partial Z_{i_4j_4}^*}\right|_{\boldsymbol{Z}=\boldsymbol{0}}\\ &= \frac{1}{\jmath^4}\left.\left[\frac{\partial\varphi_3(\boldsymbol{Z})}{\partial\Re\mathfrak{e}Z_{i_4j_4}} - \jmath\frac{\partial\varphi_3(\boldsymbol{Z})}{\partial\Im\mathfrak{m}Z_{i_4j_4}}\right]\right|_{\boldsymbol{Z}=\boldsymbol{0}}\\ &= \tilde{\Sigma}_{i_1i_2}\tilde{\Psi}_{j_1j_2}^*\tilde{\Sigma}_{i_3i_4}\tilde{\Psi}_{j_3j_4}^* + \tilde{\Sigma}_{i_1i_4}\tilde{\Psi}_{j_1j_4}^*\tilde{\Sigma}_{i_3i_2}\tilde{\Psi}_{j_3j_2}^*\end{aligned} \quad (109)$$

with

$$\varphi_1(\boldsymbol{Z}) = e^{\varphi(\boldsymbol{Z})}\left[\frac{\partial\varphi(\boldsymbol{Z})}{\partial\Re\mathfrak{e}Z_{i_1j_1}} + \jmath\frac{\partial\varphi(\boldsymbol{Z})}{\partial\Im\mathfrak{m}Z_{i_1j_1}}\right] \quad (110)$$

$$\varphi_2(\boldsymbol{Z}) = \frac{\partial\varphi_1(\boldsymbol{Z})}{\partial\Re\mathfrak{e}Z_{i_2j_2}} - \jmath\frac{\partial\varphi_1(\boldsymbol{Z})}{\partial\Im\mathfrak{m}Z_{i_2j_2}} \quad (111)$$

$$\varphi_3(\boldsymbol{Z}) = \frac{\partial\varphi_2(\boldsymbol{Z})}{\partial\Re\mathfrak{e}Z_{i_3j_3}} + \jmath\frac{\partial\varphi_2(\boldsymbol{Z})}{\partial\Im\mathfrak{m}Z_{i_3j_3}}. \quad (112)$$

Using (109), we obtain

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{X}}\left\{\mathrm{tr}\left[(\boldsymbol{AXBX}^\dagger)^2\right]\right\}\\ &= \mathbb{E}_{\boldsymbol{S}}\left\{\mathrm{tr}\left[(\boldsymbol{SS}^\dagger)^2\right]\right\}\\ &= \sum_{i=1}^{m}\sum_{p=1}^{n}\sum_{q=1}^{n}\sum_{j=1}^{m}\mathbb{E}\left\{S_{ip}S_{jp}^*S_{jq}S_{iq}^*\right\}\\ &= \sum_{i=1}^{m}\sum_{p=1}^{n}\sum_{q=1}^{n}\sum_{j=1}^{m}\left(\tilde{\Sigma}_{ij}\tilde{\Psi}_{pp}\tilde{\Sigma}_{ji}\tilde{\Psi}_{qq} + \tilde{\Sigma}_{ii}\tilde{\Psi}_{pq}\tilde{\Sigma}_{jj}\tilde{\Psi}_{qp}\right)\\ &= \mathrm{tr}^2(\tilde{\boldsymbol{\Sigma}})\mathrm{tr}(\tilde{\boldsymbol{\Psi}}^2) + \mathrm{tr}^2(\tilde{\boldsymbol{\Psi}})\mathrm{tr}(\tilde{\boldsymbol{\Sigma}}^2)\end{aligned} \quad (113)$$

from which (106) follows readily. □

Theorem 7: Let $\boldsymbol{X}_1 \sim \tilde{\mathcal{N}}_{m,p}(\boldsymbol{0}_{m\times p}, \boldsymbol{\Sigma}_1, \boldsymbol{\Psi}_1)$ and $\boldsymbol{X}_2 \sim \tilde{\mathcal{N}}_{p,n}(\boldsymbol{0}_{p\times n}, \boldsymbol{\Sigma}_2, \boldsymbol{\Psi}_2)$ be statistically independent complex Gaussian matrices. Then

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{X}_1,\boldsymbol{X}_2}\left\{\mathrm{tr}^2(\boldsymbol{X}_1\boldsymbol{X}_2\boldsymbol{X}_2^\dagger\boldsymbol{X}_1^\dagger)\right\}\\ &= \mathrm{tr}\left(\boldsymbol{\Sigma}_1^2\right)\mathrm{tr}^2(\boldsymbol{\Psi}_1\boldsymbol{\Sigma}_2)\,\mathrm{tr}\left(\boldsymbol{\Psi}_2^2\right)\\ &\quad + \mathrm{tr}\left(\boldsymbol{\Sigma}_1^2\right)\mathrm{tr}^2(\boldsymbol{\Psi}_2)\,\mathrm{tr}\left\{(\boldsymbol{\Psi}_1\boldsymbol{\Sigma}_2)^2\right\}\\ &\quad + \mathrm{tr}^2(\boldsymbol{\Sigma}_1)\,\mathrm{tr}\left\{(\boldsymbol{\Psi}_1\boldsymbol{\Sigma}_2)^2\right\}\mathrm{tr}\left(\boldsymbol{\Psi}_2^2\right)\\ &\quad + \mathrm{tr}^2(\boldsymbol{\Sigma}_1)\,\mathrm{tr}^2(\boldsymbol{\Psi}_1\boldsymbol{\Sigma}_2)\,\mathrm{tr}^2(\boldsymbol{\Psi}_2)\end{aligned} \quad (114)$$

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{X}_1,\boldsymbol{X}_2}\left\{\mathrm{tr}\left[(\boldsymbol{X}_1\boldsymbol{X}_2\boldsymbol{X}_2^\dagger\boldsymbol{X}_1^\dagger)^2\right]\right\}\\ &= \mathrm{tr}^2(\boldsymbol{\Sigma}_1)\,\mathrm{tr}^2(\boldsymbol{\Psi}_1\boldsymbol{\Sigma}_2)\,\mathrm{tr}\left(\boldsymbol{\Psi}_2^2\right)\\ &\quad + \mathrm{tr}^2(\boldsymbol{\Sigma}_1)\,\mathrm{tr}^2(\boldsymbol{\Psi}_2)\,\mathrm{tr}\left\{(\boldsymbol{\Psi}_1\boldsymbol{\Sigma}_2)^2\right\}\\ &\quad + \mathrm{tr}\left\{(\boldsymbol{\Psi}_1\boldsymbol{\Sigma}_2)^2\right\}\mathrm{tr}\left(\boldsymbol{\Psi}_2^2\right)\mathrm{tr}\left(\boldsymbol{\Sigma}_1^2\right)\\ &\quad + \mathrm{tr}^2(\boldsymbol{\Psi}_1\boldsymbol{\Sigma}_2)\,\mathrm{tr}^2(\boldsymbol{\Psi}_2)\,\mathrm{tr}\left(\boldsymbol{\Sigma}_1^2\right).\end{aligned} \quad (115)$$

Proof: Using the first two cumulants from Lemma 5, we get

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{X}_1,\boldsymbol{X}_2}\left\{\mathrm{tr}^2(\boldsymbol{X}_1\boldsymbol{X}_2\boldsymbol{X}_2^\dagger\boldsymbol{X}_1^\dagger)\right\}\\ &= \mathbb{E}_{\boldsymbol{X}_2}\left\{\mathrm{tr}\left(\boldsymbol{\Sigma}_1^2\right)\mathrm{tr}\left[(\boldsymbol{X}_2\boldsymbol{X}_2^\dagger\boldsymbol{\Psi}_1)^2\right] + \mathrm{tr}^2(\boldsymbol{\Sigma}_1)\,\mathrm{tr}^2(\boldsymbol{X}_2\boldsymbol{X}_2^\dagger\boldsymbol{\Psi}_1)\right\}\end{aligned} \quad (116)$$

where it follows from Lemma 6 that

$$\mathbb{E}_{\boldsymbol{X}_2}\left\{\mathrm{tr}\left[(\boldsymbol{X}_2\boldsymbol{X}_2^\dagger\boldsymbol{\Psi}_1)^2\right]\right\} = \mathrm{tr}^2(\boldsymbol{\Psi}_1\boldsymbol{\Sigma}_2)\,\mathrm{tr}\left(\boldsymbol{\Psi}_2^2\right) + \mathrm{tr}^2(\boldsymbol{\Psi}_2)\,\mathrm{tr}\left\{(\boldsymbol{\Psi}_1\boldsymbol{\Sigma}_2)^2\right\} \tag{117}$$

and from Lemma 5 that

$$\mathbb{E}_{\boldsymbol{X}_2}\left\{\mathrm{tr}^2(\boldsymbol{X}_2\boldsymbol{X}_2^\dagger\boldsymbol{\Psi}_1)\right\} = \mathrm{tr}\left\{(\boldsymbol{\Psi}_1\boldsymbol{\Sigma}_2)^2\right\}\mathrm{tr}\left(\boldsymbol{\Psi}_2^2\right) + \mathrm{tr}^2(\boldsymbol{\Psi}_1\boldsymbol{\Sigma}_2)\,\mathrm{tr}^2(\boldsymbol{\Psi}_2). \tag{118}$$

Combining (116)–(118) yields the desired result (114).

Similar to (116), we have

$$\mathbb{E}_{\boldsymbol{X}_1,\boldsymbol{X}_2}\left\{\mathrm{tr}\left[(\boldsymbol{X}_1\boldsymbol{X}_2\boldsymbol{X}_2^\dagger\boldsymbol{X}_1^\dagger)^2\right]\right\} = \mathbb{E}_{\boldsymbol{X}_2}\left\{\mathrm{tr}^2(\boldsymbol{\Sigma}_1)\,\mathrm{tr}\left[(\boldsymbol{X}_2\boldsymbol{X}_2^\dagger\boldsymbol{\Psi}_1)^2\right] + \mathrm{tr}^2(\boldsymbol{X}_2\boldsymbol{X}_2^\dagger\boldsymbol{\Psi}_1)\mathrm{tr}\left(\boldsymbol{\Sigma}_1^2\right)\right\}. \tag{119}$$

From (117)–(119), we obtain the desired result (115). □

Theorem 8: Let $\boldsymbol{X} \sim \tilde{\mathcal{N}}_{m,n}(\boldsymbol{0}_{m\times n}, \boldsymbol{\Sigma}, \boldsymbol{I}_n)$, $m \le n$, and $\sigma_1, \sigma_2, \ldots, \sigma_m$ be the eigenvalues of $\boldsymbol{\Sigma}$ in any order. Then, the joint probability density function (pdf) of the ordered eigenvalues $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_m > 0$ of a central complex Wishart matrix $\boldsymbol{X}\boldsymbol{X}^\dagger \sim \tilde{\mathcal{W}}_m(n, \boldsymbol{\Sigma})$ is given by

$$p_{\boldsymbol{\lambda}}(\lambda_1, \lambda_2, \ldots, \lambda_m) = \mathcal{A}^{-1}\det\left(\left[\boldsymbol{G}_1 \quad \boldsymbol{G}_2 \quad \ldots \quad \boldsymbol{G}_{\varrho(\boldsymbol{\Sigma})}\right]\right) \times \det_{1\le i,j\le m}\left(\lambda_j^{i-1}\right)\prod_{k=1}^{m}\lambda_k^{n-m} \tag{120}$$

where

$$\mathcal{A} = K_{0,0}^{m,n}\cdot\det\left(\left[\boldsymbol{\mathcal{B}}_1 \quad \boldsymbol{\mathcal{B}}_2 \quad \ldots \quad \boldsymbol{\mathcal{B}}_{\varrho(\boldsymbol{\Sigma})}\right]\right) \tag{121}$$

and $\boldsymbol{G}_k = \left(G_{k,ij}\right)$ and $\boldsymbol{\mathcal{B}}_k = \left(\mathcal{B}_{k,ij}\right)$, $k = 1, 2, \ldots, \varrho(\boldsymbol{\Sigma})$ are $m \times \tau_k(\boldsymbol{\Sigma})$ matrices, whose (i,j)th entries are given, respectively, by

$$G_{k,ij} = \lambda_i^{j-1}e^{-\lambda_i/\sigma_{\langle k\rangle}} \tag{122}$$

$$\mathcal{B}_{k,ij} = (-1)^{i-j}(i-j+1)_{j-1}\,\sigma_{\langle k\rangle}^{n-i+j}. \tag{123}$$

Proof: The joint eigenvalue density $p_{\boldsymbol{\lambda}}(\lambda_1, \lambda_2, \ldots, \lambda_m)$ is given by [75, eq. (95)] in terms of the hypergeometric function of matrix arguments. To render this joint pdf more amenable to further analysis and computationally tractable, we apply Lemma 4 to [75, eq. (95)], which results in (120) after some algebra. □

Note that (120) is valid for any covariance matrix $\boldsymbol{\Sigma}$ with the eigenvalues of arbitrary multiplicity and hence, generalizes the previous determinantal representation for the joint eigenvalue pdf of Wishart matrices. If $\boldsymbol{\Sigma} = \boldsymbol{I}_m$ in Theorem 8, all of the eigenvalues are identically equal to one and hence, with $\varrho(\boldsymbol{\Sigma}) = 1$, $\tau_1(\boldsymbol{\Sigma}) = m$, and $\sigma_{\langle 1\rangle} = 1$, (120) reduces to [22, eq. (6)]. Furthermore, if all the eigenvalues of $\boldsymbol{\Sigma}$ are distinct, then, with $\varrho(\boldsymbol{\Sigma}) = m$ and $\tau_1(\boldsymbol{\Sigma}) = \tau_2(\boldsymbol{\Sigma}) = \cdots = \tau_m(\boldsymbol{\Sigma}) = 1$, (120) reduces to [22, eq.(18)].

Theorem 9: Let $\boldsymbol{X} \sim \tilde{\mathcal{N}}_{m,n}(\boldsymbol{0}_{m\times n}, \boldsymbol{I}_m, \boldsymbol{\Psi})$, $m \le n$, $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ be Hermitian positive definite, and $\beta_1, \beta_2, \ldots, \beta_n$ be the eigenvalues of $\boldsymbol{A}^{1/2}\boldsymbol{\Psi}\boldsymbol{A}^{1/2}$ in any order. Then, the joint pdf of the ordered eigenvalues $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_m > 0$ of a matrix quadratic form $\boldsymbol{X}\boldsymbol{A}\boldsymbol{X}^\dagger$ is given by (124) at the bottom of the page. In (124), $\boldsymbol{Q}_k = \left(Q_{k,ij}\right)$ and $\boldsymbol{\mathcal{V}}_{(l),k} = \left(\mathcal{V}_{(l),k,ij}\right)$, $l \le n$, $k = 1, 2, \ldots, \varrho\left(\boldsymbol{A}^{1/2}\boldsymbol{\Psi}\boldsymbol{A}^{1/2}\right)$, are $m \times \tau_k\left(\boldsymbol{A}^{1/2}\boldsymbol{\Psi}\boldsymbol{A}^{1/2}\right)$ and $l \times \tau_k\left(\boldsymbol{A}^{1/2}\boldsymbol{\Psi}\boldsymbol{A}^{1/2}\right)$ matrices, whose (i,j)th entries are given, respectively, by

$$Q_{k,ij} = \lambda_i^{j-1}e^{-\lambda_i/\beta_{\langle k\rangle}} \tag{125}$$

$$\mathcal{V}_{(l),k,ij} = (-1)^{i-j}(i-j+1)_{j-1}\,\beta_{\langle k\rangle}^{-i+j}. \tag{126}$$

Proof: Let $\boldsymbol{S} = \boldsymbol{X}\boldsymbol{A}\boldsymbol{X}^\dagger$, then $\boldsymbol{S} \sim \tilde{\mathcal{Q}}_{m,n}(\boldsymbol{A}, \boldsymbol{I}_m, \boldsymbol{\Psi})$ is a positive-definite quadratic form in the complex Gaussian matrix [21, Definition II.3]. Using the pdf [23, eq. (2)], we can write the joint eigenvalue pdf of $\boldsymbol{S}$ in the form

$$\begin{aligned} p_{\boldsymbol{\lambda}}(\lambda_1, \lambda_2, \ldots, \lambda_m) &= \frac{\pi^{m(m-1)}}{\tilde{\Gamma}_m(m)}\int_{\boldsymbol{U}\in\mathcal{U}(m)} p_{\boldsymbol{S}}\left(\boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^\dagger\right)\prod_{i<j}^{m}(\lambda_i - \lambda_j)^2\,[d\boldsymbol{U}] \\ &= \frac{\pi^{m(m-1)}\det(\boldsymbol{A}\boldsymbol{\Psi})^{-m}}{\tilde{\Gamma}_m(n)\,\tilde{\Gamma}_m(m)}\,{}_0\tilde{F}_0^{(n)}\left(\boldsymbol{D}, -\boldsymbol{\Psi}^{-1}\boldsymbol{A}^{-1}\right) \\ &\quad \times \prod_{k=1}^{m}\lambda_k^{n-m}\prod_{i<j}^{m}(\lambda_i - \lambda_j)^2 \end{aligned} \tag{127}$$

where

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

$\mathfrak{Re}(\alpha) > m-1$, is the complex multivariate gamma function, $\Gamma(\cdot)$ is the gamma function, and $\boldsymbol{D} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m)$. In (127), $\mathcal{U}(m) = \left\{\boldsymbol{U} : \boldsymbol{U}\boldsymbol{U}^\dagger = \boldsymbol{I}_m\right\}$ is the unitary group of order m and $[d\boldsymbol{U}]$ is the unitary invariant Haar measure on the unitary group $\mathcal{U}(m)$ normalized to make the total volume unity. Similar to Theorem 8, we obtain the desired result (124) applying Lemma 4 to (127). □

Definition 4 (Characteristic Coefficient): Let $\boldsymbol{A}$ be an $n \times n$ Hermitian matrix with the eigenvalues $\alpha_1, \alpha_2, \ldots, \alpha_n$ in any order. Then, the (i,j)th characteristic coefficient $\mathcal{X}_{i,j}(\boldsymbol{A})$,

---

$$p_{\boldsymbol{\lambda}}(\lambda_1, \lambda_2, \ldots, \lambda_m) = \frac{\det\left(\begin{bmatrix}\boldsymbol{\mathcal{V}}_{(n-m),1} & \boldsymbol{\mathcal{V}}_{(n-m),2} & \cdots & \boldsymbol{\mathcal{V}}_{(n-m),\varrho\left(\boldsymbol{A}^{1/2}\boldsymbol{\Psi}\boldsymbol{A}^{1/2}\right)} \\ \boldsymbol{Q}_1 & \boldsymbol{Q}_2 & \cdots & \boldsymbol{Q}_{\varrho\left(\boldsymbol{A}^{1/2}\boldsymbol{\Psi}\boldsymbol{A}^{1/2}\right)}\end{bmatrix}\right)}{K_{0,0}^{m,m}\det(\boldsymbol{A}\boldsymbol{\Psi})^m\det\left(\left[\boldsymbol{\mathcal{V}}_{(n),1} \quad \boldsymbol{\mathcal{V}}_{(n),2} \quad \cdots \quad \boldsymbol{\mathcal{V}}_{(n),\varrho\left(\boldsymbol{A}^{1/2}\boldsymbol{\Psi}\boldsymbol{A}^{1/2}\right)}\right]\right)}\det_{1\le i,j\le m}\left(\lambda_j^{i-1}\right) \tag{124}$$

$i = 1, 2, \ldots, \varrho(\boldsymbol{A})$, $j = 1, 2, \ldots, \tau_i(\boldsymbol{A})$, is defined as a partial fraction expansion coefficient of $\det(\boldsymbol{I}_n + \xi\boldsymbol{A})^{-1}$ such that

$$\begin{aligned}\det(\boldsymbol{I}_n + \xi\boldsymbol{A})^{-1} &= \prod_{i=1}^{\varrho(\boldsymbol{A})} (1 + \xi\alpha_{\langle i\rangle})^{-\tau_i(\boldsymbol{A})} \\ &= \sum_{i=1}^{\varrho(\boldsymbol{A})} \sum_{j=1}^{\tau_i(\boldsymbol{A})} \mathcal{X}_{i,j}(\boldsymbol{A})(1 + \xi\alpha_{\langle i\rangle})^{-j} \quad (128)\end{aligned}$$

where ξ is a scalar constant such that $\boldsymbol{I}_n + \xi\boldsymbol{A}$ is nonsingular. The (i,j)th characteristic coefficient $\mathcal{X}_{i,j}(\boldsymbol{A})$ can be determined by (129), at the bottom of the page, where $\varpi_{i,j} = \tau_i(\boldsymbol{A}) - j$.

Note that the characteristic coefficients are invariant with respect to the constant ξ and only a function of the spectra of $\boldsymbol{A}$. In addition, it can be seen from (128) with $\xi = 0$ that the sum of all the characteristic coefficients is equal to one. By definition, we have

$$\mathcal{X}_{1,j}(\boldsymbol{I}_n) = \begin{cases} 0, & j = 1, 2, \ldots, n-1 \\ 1, & j = n. \end{cases} \quad (130)$$

Example 6 (Constant Correlation Matrix): Consider a constant correlation matrix $\boldsymbol{\Phi}_n^{(c)}\rho$. Since the eigenvalues of $\boldsymbol{\Phi}_n^{(c)}\rho$ are $1 + (n-1)\rho$ and $1 - \rho$ with $n-1$ multiplicity, it is easy to show that the characteristic coefficients of $\boldsymbol{\Phi}_n^{(c)}\rho$, $\rho \in (0, 1)$, are

$$\mathcal{X}_{1,1}(\boldsymbol{\Phi}_n^{(c)}(\rho)) = \left(\frac{n\rho}{1 - \rho + n\rho}\right)^{1-n} \quad (131)$$

$$\mathcal{X}_{2,j}(\boldsymbol{\Phi}_n^{(c)}(\rho)) = -\frac{1-\rho}{1-\rho+n\rho} \cdot \left(\frac{n\rho}{1-\rho+n\rho}\right)^{j-n} \quad (132)$$

where $j = 1, 2, \ldots, n-1$.

Theorem 10: Let $\boldsymbol{X} \sim \tilde{\mathcal{N}}_{m,n}(\boldsymbol{0}_{m\times n}, \boldsymbol{\Sigma}, \boldsymbol{I}_n)$, $m \le n$, and $\sigma_1, \sigma_2, \ldots, \sigma_m$ be the eigenvalues of $\boldsymbol{\Sigma}$. Let $\boldsymbol{A}$ be a $\nu \times \nu$ positive-semidefinite matrix with the eigenvalues $\alpha_1, \alpha_2, \ldots, \alpha_\nu$. Then, for $\xi \ge 0$, we have

$$\begin{aligned}&\mathbb{E}\left\{\det(\boldsymbol{I}_{m\nu} + \xi\boldsymbol{A} \otimes \boldsymbol{X}\boldsymbol{X}^\dagger)^{-1}\right\} \\ &\qquad = \mathcal{A}^{-1} \det([\boldsymbol{\Omega}_1 \quad \boldsymbol{\Omega}_2 \quad \ldots \quad \boldsymbol{\Omega}_{\varrho(\boldsymbol{\Sigma})}]) \quad (133)\end{aligned}$$

where $\mathcal{A}$ is given in (121) and $\boldsymbol{\Omega}_k = (\Omega_{k,ij})$, $k = 1, 2, \ldots, \varrho(\boldsymbol{\Sigma})$, are $m \times \tau_k(\boldsymbol{\Sigma})$ matrices whose (i,j)th entry is given by

$$\begin{aligned}\Omega_{k,ij} = \sum_{p=1}^{\varrho(\boldsymbol{A})} \sum_{q=1}^{\tau_p(\boldsymbol{A})} \Big\{&\sigma_{\langle k\rangle}^{n-m+i+j-1}(n-m+i+j-2)! \\ &\times \mathcal{X}_{p,q}(\boldsymbol{A})\, {}_2F_0\left(n-m+i+j-1, q; -\xi\alpha_{\langle p\rangle}\sigma_{\langle k\rangle}\right)\Big\} \quad (134)\end{aligned}$$

where $\mathcal{X}_{p,q}(\boldsymbol{A})$ is the (p,q)th characteristic coefficient of $\boldsymbol{A}$.

Proof: From Theorem 8, we have (135) at the bottom of the page, where (a) follows from the fact that the integrand is symmetric in $\lambda_1, \lambda_2, \ldots, \lambda_m$ and (b) follows from the generalized Cauchy–Binet formula [22, Appendix], [23, Lemma 2], yielding the (i,j)th entry of $m \times \tau_k(\boldsymbol{\Sigma})$ matrices $\boldsymbol{\Omega}_k$, $k = 1, 2, \ldots, \varrho(\boldsymbol{\Sigma})$, as

$$\begin{aligned}\Omega_{k,ij} = \int_0^\infty &\prod_{p=1}^{\varrho(\boldsymbol{A})} (1 + \xi\alpha_{\langle p\rangle}\lambda)^{-\tau_p(\boldsymbol{A})} \\ &\times \lambda^{n-m+i+j-2} e^{-\lambda/\sigma_{\langle k\rangle}} d\lambda. \quad (136)\end{aligned}$$

$$\begin{aligned}\mathcal{X}_{i,j}(\boldsymbol{A}) &= \frac{1}{\varpi_{i,j}!\,\alpha_{\langle i\rangle}^{\varpi_{i,j}}} \cdot \left[\frac{d^{\varpi_{i,j}}}{d\upsilon^{\varpi_{i,j}}}\left(1 + \upsilon\alpha_{\langle i\rangle}\right)^{\tau_i(\boldsymbol{A})} \det(\boldsymbol{I}_n + \upsilon\boldsymbol{A})^{-1}\right]\Bigg|_{\upsilon = -1/\alpha_{\langle i\rangle}} \\ &= \frac{(-1)^{\varpi_{i,j}}}{\alpha_{\langle i\rangle}^{\varpi_{i,j}}} \sum_{\substack{k_1 + k_2 + \cdots + k_{\varrho(\boldsymbol{A})} = \varpi_{i,j} \\ k_l \in \{0, \mathbb{N}\} \text{ for } \forall l \ne i \\ k_i = 0}} \prod_{\substack{l=1 \\ l \ne i}}^{\varrho(\boldsymbol{A})} \binom{\tau_l(\boldsymbol{A}) + k_l - 1}{k_l} \frac{\alpha_{\langle l\rangle}^{k_l}}{\left(1 - \frac{\alpha_{\langle l\rangle}}{\alpha_{\langle i\rangle}}\right)^{\tau_l(\boldsymbol{A}) + k_l}} \quad (129)\end{aligned}$$

$$\begin{aligned}&\mathbb{E}\left\{\det(\boldsymbol{I}_{m\nu} + \xi\boldsymbol{A} \otimes \boldsymbol{X}\boldsymbol{X}^\dagger)^{-1}\right\} \\ &= \mathbb{E}\left\{\prod_{p=1}^{\varrho(\boldsymbol{A})} \det(\boldsymbol{I}_m + \xi\alpha_{\langle p\rangle}\boldsymbol{X}\boldsymbol{X}^\dagger)^{-\tau_p(\boldsymbol{A})}\right\} \\ &= \int \cdots \int_{0 < \lambda_m \le \cdots \le \lambda_1 < \infty} \prod_{k=1}^m \prod_{p=1}^{\varrho(\boldsymbol{A})} (1 + \xi\alpha_{\langle p\rangle}\lambda_k)^{-\tau_p(\boldsymbol{A})} p_{\boldsymbol{\lambda}}(\lambda_1, \lambda_2, \ldots, \lambda_m) d\lambda_1 \cdots d\lambda_m \\ &\overset{(a)}{=} \frac{1}{m!\mathcal{A}} \underbrace{\int_0^\infty \cdots \int_0^\infty}_{m\text{-fold}} \det_{1 \le i,j \le m}\left(\lambda_j^{i-1}\right) \det([\boldsymbol{G}_1 \quad \boldsymbol{G}_2 \quad \ldots \quad \boldsymbol{G}_{\varrho(\boldsymbol{\Sigma})}]) \prod_{k=1}^m \left\{\lambda_k^{n-m} \prod_{p=1}^{\varrho(\boldsymbol{A})} (1 + \xi\alpha_{\langle p\rangle}\lambda_k)^{-\tau_p(\boldsymbol{A})}\right\} d\lambda_1 \cdots d\lambda_m \\ &\overset{(b)}{=} \mathcal{A}^{-1} \det([\boldsymbol{\Omega}_1 \quad \boldsymbol{\Omega}_2 \quad \ldots \quad \boldsymbol{\Omega}_{\varrho(\boldsymbol{\Sigma})}]) \quad (135)\end{aligned}$$

Using a partial fraction decomposition, (136) can be written as

$$\Omega_{k,ij} = \sum_{p=1}^{\varrho(\boldsymbol{A})} \sum_{q=1}^{\tau_p(\boldsymbol{A})} \mathcal{X}_{p,q}(\boldsymbol{A}) \int_0^\infty (1+\xi\alpha_{\langle p\rangle}\lambda)^{-q} \times \lambda^{n-m+i+j-2} e^{-\lambda/\sigma_{\langle k\rangle}} d\lambda \quad (137)$$

where the characteristic coefficients $\mathcal{X}_{p,q}(\boldsymbol{A})$ is given by (129). We complete the proof of the theorem by evaluating the integral in (137) with the help of the following integral identity:

$$\int_0^\infty (1+ax)^{\mu-1} x^{n-1} e^{-x/b} dx = b^n (n-1)!\, {}_2F_0(n, -\mu+1; -ab) \quad (138)$$

where $a, b > 0$, $n \in \mathbb{N}$, and $\mu \in \mathbb{C}$. □

Corollary 4: Let $\boldsymbol{X} \sim \tilde{\mathcal{N}}_{m,n}(\boldsymbol{0}_{m\times n}, \boldsymbol{\Sigma}, \boldsymbol{I}_n)$, $m \le n$. Then, for $\nu \in \mathbb{N}$, we have

$$\mathbb{E}\left\{\det(\boldsymbol{I}_m + \xi\boldsymbol{X}\boldsymbol{X}^\dagger)^{-\nu}\right\} = \frac{\det(\boldsymbol{\Omega})}{\prod_{i=1}^m (n-i)!\,(i-1)!} \quad (139)$$

where $\boldsymbol{\Omega} = (\Omega_{ij})$ is the $m \times m$ Hankel matrix whose (i,j)th entry is given by

$$\Omega_{ij} = (n-m+i+j-2)!\, {}_2F_0(n-m+i+j-1, \nu; -\xi). \quad (140)$$

Proof: It follows immediately from Theorem 10 with $\boldsymbol{\Sigma} = \boldsymbol{I}_m$, $\boldsymbol{A} = \boldsymbol{I}_\nu$, $\varrho(\boldsymbol{\Sigma}) = 1$, $\tau_1(\boldsymbol{\Sigma}) = m$, $\sigma_{\langle 1\rangle} = 1$, $\varrho(\boldsymbol{A}) = 1$, $\tau_1(\boldsymbol{A}) = \nu$, and $\alpha_{\langle 1\rangle} = 1$. □

Theorem 11: Let $\boldsymbol{X} \sim \tilde{\mathcal{N}}_{m,n}(\boldsymbol{0}_{m\times n}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$, σ_i, $i = 1, 2, \ldots, m$, and ψ_j, $j = 1, 2, \ldots, n$, be the eigenvalues of $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$, respectively. Then, for $\xi \ge 0$, we have

$$\mathbb{E}\left\{\det(\boldsymbol{I}_m + \xi\boldsymbol{X}\boldsymbol{X}^\dagger)^{-1}\right\} = \sum_{p=1}^{\varrho(\boldsymbol{\Sigma})} \sum_{q=1}^{\varrho(\boldsymbol{\Psi})} \sum_{i=1}^{\tau_p(\boldsymbol{\Sigma})} \sum_{j=1}^{\tau_q(\boldsymbol{\Psi})} \mathcal{X}_{p,i}(\boldsymbol{\Sigma}) \mathcal{X}_{q,j}(\boldsymbol{\Psi})\, {}_2F_0\left(i, j; -\xi\sigma_{\langle p\rangle}\psi_{\langle q\rangle}\right) \quad (141)$$

where $\mathcal{X}_{p,i}(\boldsymbol{\Sigma})$ and $\mathcal{X}_{q,j}(\boldsymbol{\Psi})$ are the (p,i)th and (q,j)th characteristic coefficients of $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$, respectively.

Proof: It follows from Lemmas 1 and 2 that

$$\begin{aligned}\det(\boldsymbol{I}_m + \xi\boldsymbol{X}\boldsymbol{X}^\dagger)^{-1} &= \mathbb{E}_{\boldsymbol{y}_1}\left\{\operatorname{etr}(-\xi\boldsymbol{X}^\dagger\boldsymbol{y}_1\boldsymbol{y}_1^\dagger\boldsymbol{X})\right\} \\ &= \mathbb{E}_{\boldsymbol{y}_1,\boldsymbol{y}_2}\left\{\operatorname{etr}(\xi\boldsymbol{y}_1^\dagger\boldsymbol{X}\boldsymbol{y}_2 - \boldsymbol{y}_2^\dagger\boldsymbol{X}^\dagger\boldsymbol{y}_1)\right\}\end{aligned} \quad (142)$$

where $\boldsymbol{y}_1 \sim \tilde{\mathcal{N}}_{m,1}(\boldsymbol{0}_{m\times 1}, \boldsymbol{I}_m, 1)$ and $\boldsymbol{y}_2 \sim \tilde{\mathcal{N}}_{n,1}(\boldsymbol{0}_{n\times 1}, \boldsymbol{I}_n, 1)$. Denoting the left-hand side of (141) by $\text{LHS}_{(141)}$ and using (142), we have

$$\begin{aligned}\text{LHS}_{(141)} &= \mathbb{E}_{\boldsymbol{y}_1,\boldsymbol{y}_2}\left\{\mathbb{E}_{\boldsymbol{X}}\left\{\operatorname{etr}(\xi\boldsymbol{y}_2\boldsymbol{y}_1^\dagger\boldsymbol{X} - \boldsymbol{X}^\dagger\boldsymbol{y}_1\boldsymbol{y}_2^\dagger)\right\}\right\} \\ &= \mathbb{E}_{\boldsymbol{y}_1,\boldsymbol{y}_2}\left\{\exp(-\xi\boldsymbol{y}_1^\dagger\boldsymbol{\Sigma}\boldsymbol{y}_1\boldsymbol{y}_2^\dagger\boldsymbol{\Psi}\boldsymbol{y}_2)\right\}.\end{aligned} \quad (143)$$

Now, introducing a delta function to decouple the expectations for $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ in (143) yields (144) at the bottom of the page, where (a) is obtained by replacing the delta function with its Fourier representation, (b) follows from Lemma 1, and (c) is obtained from Definition 4. Using the integral identity, for $a > 0$, $\ell \in \mathbb{N}$, and $z \in \mathbb{R}$

$$\int_{-\infty}^\infty e^{\jmath\omega z}(1+\jmath a\omega)^{-\ell} d\omega = \frac{\pi z^{\ell-1} e^{-\sqrt{z^2}/a}}{a^\ell (\ell-1)!}(1+\operatorname{sign}(z)) \quad (145)$$

(144) can be written as

$$\text{LHS}_{(141)} = \sum_{p=1}^{\varrho(\boldsymbol{\Sigma})} \sum_{q=1}^{\varrho(\boldsymbol{\Psi})} \sum_{i=1}^{\tau_p(\boldsymbol{\Sigma})} \sum_{j=1}^{\tau_q(\boldsymbol{\Psi})} \frac{\mathcal{X}_{p,i}(\boldsymbol{\Sigma})\mathcal{X}_{q,j}(\boldsymbol{\Psi})}{\sigma_{\langle p\rangle}^i (i-1)!} \times \int_0^\infty (1+\xi\psi_{\langle q\rangle}z)^{-j} z^{i-1} e^{-z/\sigma_{\langle p\rangle}} dz. \quad (146)$$

Finally, we obtain the desired result (141) by evaluating the integral in (146) with the help of (138). □

Appendix III
Proofs

A. Proof of Theorem 1

We first prove Theorem 1 for M-ary phase shift keying (M-PSK) signaling. The SEP of the OSTBC with M-PSK constellation can be expressed as [41], [42]

$$P_{\text{e, MPSK}} = \frac{1}{\pi}\int_0^\Theta \phi_{\gamma_{\text{STBC}}}\left(\frac{g}{\sin^2\theta}; \bar{\gamma}\right) d\theta \quad (147)$$

$$\begin{aligned}\text{LHS}_{(141)} &= \mathbb{E}_{\boldsymbol{y}_1,\boldsymbol{y}_2}\left\{\int_{-\infty}^\infty e^{-\xi z\boldsymbol{y}_2^\dagger\boldsymbol{\Psi}\boldsymbol{y}_2}\delta(z - \boldsymbol{y}_1^\dagger\boldsymbol{\Sigma}\boldsymbol{y}_1)dz\right\} \\ &\overset{\text{(a)}}{=} \frac{1}{2\pi}\mathbb{E}_{\boldsymbol{y}_1,\boldsymbol{y}_2}\left\{\int_{-\infty}^\infty\int_{-\infty}^\infty e^{-\xi z\boldsymbol{y}_2^\dagger\boldsymbol{\Psi}\boldsymbol{y}_2} e^{\jmath\left(z-\boldsymbol{y}_1^\dagger\boldsymbol{\Sigma}\boldsymbol{y}_1\right)\omega} d\omega dz\right\} \\ &= \frac{1}{2\pi}\int_{-\infty}^\infty\int_{-\infty}^\infty e^{\jmath\omega z}\mathbb{E}_{\boldsymbol{y}_1}\left\{\operatorname{etr}(-\jmath\omega\boldsymbol{\Sigma}\boldsymbol{y}_1\boldsymbol{y}_1^\dagger)\right\}\mathbb{E}_{\boldsymbol{y}_2}\left\{\operatorname{etr}(-\xi z\boldsymbol{\Psi}\boldsymbol{y}_2\boldsymbol{y}_2^\dagger)\right\} d\omega dz \\ &\overset{\text{(b)}}{=} \frac{1}{2\pi}\int_{-\infty}^\infty\int_{-\infty}^\infty e^{\jmath\omega z}\det(\boldsymbol{I}_m + \jmath\omega\boldsymbol{\Sigma})^{-1}\det(\boldsymbol{I}_n + \xi z\boldsymbol{\Psi})^{-1} d\omega dz \\ &\overset{\text{(c)}}{=} \frac{1}{2\pi}\sum_{p=1}^{\varrho(\boldsymbol{\Sigma})} \sum_{q=1}^{\varrho(\boldsymbol{\Psi})} \sum_{i=1}^{\tau_p(\boldsymbol{\Sigma})} \sum_{j=1}^{\tau_q(\boldsymbol{\Psi})} \mathcal{X}_{p,i}(\boldsymbol{\Sigma})\mathcal{X}_{q,j}(\boldsymbol{\Psi})\int_{-\infty}^\infty\int_{-\infty}^\infty e^{\jmath\omega z}(1+\jmath\sigma_{\langle p\rangle}\omega)^{-i}(1+\xi\psi_{\langle q\rangle}z)^{-j} d\omega dz\end{aligned} \quad (144)$$

$$\mathbb{E}\left\{\|\boldsymbol{H}\|_{\mathrm{F}}^{4}\right\} = \mathbb{E}_{\boldsymbol{\Xi}_1,\boldsymbol{\Xi}_2}\left\{\mathrm{tr}^2\left(\frac{1}{n_{\mathrm{S}}}\boldsymbol{\Xi}_1\boldsymbol{\Xi}_2\boldsymbol{\Xi}_2^{\dagger}\boldsymbol{\Xi}_1^{\dagger}\right)\right\}$$
$$= \left(\frac{n_{\mathrm{R}}}{n_{\mathrm{S}}}\right)^2 \mathrm{tr}\left(\boldsymbol{\Phi}_{\mathrm{T}}^2\right)\mathrm{tr}\left(\boldsymbol{\Phi}_{\mathrm{S}}^2\right) + \mathrm{tr}\left(\boldsymbol{\Phi}_{\mathrm{T}}^2\right)\mathrm{tr}\left(\boldsymbol{\Phi}_{\mathrm{R}}^2\right) + \left(\frac{n_{\mathrm{T}}}{n_{\mathrm{S}}}\right)^2 \mathrm{tr}\left(\boldsymbol{\Phi}_{\mathrm{R}}^2\right)\mathrm{tr}\left(\boldsymbol{\Phi}_{\mathrm{S}}^2\right) + (n_{\mathrm{T}}n_{\mathrm{R}})^2. \tag{155}$$

where $\Theta = \pi - \pi/M$ and $g = \sin^2(\pi/M)$. From (147), we can obtain the upper bound as

$$P_{\mathrm{e,\ MPSK}} \le \left(1 - \frac{1}{M}\right)\phi_{\gamma_{\mathrm{STBC}}}(g;\bar{\gamma}) \tag{148}$$

which becomes tighter as $\bar{\gamma}$ increases [49], and hence yields

$$d_{\mathrm{STBC}} = \lim_{\bar{\gamma}\to\infty} \frac{-\log \phi_{\gamma_{\mathrm{STBC}}}(g;\bar{\gamma})}{\log \bar{\gamma}}. \tag{149}$$

Therefore, the asymptotic behavior of the MGF $\phi_{\gamma_{\mathrm{STBC}}}(s;\bar{\gamma})$ at large $\bar{\gamma}$ reveals a high-SNR slope of the SEP curve.

Suppose that $\bar{\gamma}$ is sufficiently large. For $n_{\mathrm{T}} \le n_{\mathrm{R}}$, it follows from (7) that

$$\log \phi_{\gamma_{\mathrm{STBC}}}(g;\bar{\gamma}) \approx -\underbrace{\mathrm{rank}\left(\boldsymbol{\Xi}_1^{\dagger}\boldsymbol{\Xi}_1\boldsymbol{\Phi}_{\mathrm{S}}\otimes\boldsymbol{\Phi}_{\mathrm{T}}\right)}_{n_{\mathrm{T}}\cdot\min(n_{\mathrm{R}},n_{\mathrm{S}})}\log\bar{\gamma} + \mathrm{constant}. \tag{150}$$

Similarly, using (8), we have for $n_{\mathrm{T}} > n_{\mathrm{R}}$

$$\log \phi_{\gamma_{\mathrm{STBC}}}(g;\bar{\gamma}) \approx -\underbrace{\mathrm{rank}\left(\boldsymbol{\Phi}_{\mathrm{R}}\otimes\boldsymbol{\Xi}_2\boldsymbol{\Xi}_2^{\dagger}\right)}_{n_{\mathrm{R}}\cdot\min(n_{\mathrm{T}},n_{\mathrm{S}})}\log\bar{\gamma} + \mathrm{constant}'. \tag{151}$$

Hence

$$d_{\mathrm{STBC}} = \min(n_{\mathrm{T}},n_{\mathrm{R}})\cdot\min\{\max(n_{\mathrm{T}},n_{\mathrm{R}}),n_{\mathrm{S}}\} \tag{152}$$

from which (10) follows immediately. For a general case of arbitrary two-dimensional signaling constellation with polygonal decision boundaries, the SEP can be written as a convex combination of terms akin to (147) [83]. Hence, we can easily generalize the proof to the case of any two-dimensional signaling constellation.

B. Proofs of Properties 1–3

1) Proof of Property 1: Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of $\boldsymbol{\Phi}$. Then, the correlation figure $\zeta(\boldsymbol{\Phi})$ defined in Definition 2 can be written as

$$\zeta(\boldsymbol{\Phi}) = \frac{1}{n^2}\sum_{k=1}^{n}\lambda_k^2 \tag{153}$$

which is symmetric in $\lambda_1, \lambda_2, \ldots, \lambda_n$ and holds Schur's condition (61). Hence, we complete the proof.

2) Proof of Property 2: Since $\prod_{i=1}^{m}\zeta(\boldsymbol{\Phi}_i) = \zeta(\bigotimes_{i=1}^{m}\boldsymbol{\Phi}_i)$, it follows immediately from Property 1.

3) Proof of Property 3: Let $\lambda_1^{(i)}, \lambda_2^{(i)}, \ldots, \lambda_n^{(i)}$ be the eigenvalues of $\boldsymbol{\Phi}_i$ $(i = 1, 2, \ldots, m)$. Then, $\sum_{i=1}^{m}\zeta(\boldsymbol{\Phi}_i)$ can be written as

$$\sum_{i=1}^{m}\zeta(\boldsymbol{\Phi}_i) = \sum_{i=1}^{m}\sum_{k=1}^{n_i}\left(\frac{\lambda_k^{(i)}}{n_i}\right)^2 \tag{154}$$

which is symmetric in
$$\left\{\frac{1}{n_i}\lambda_1^{(i)}, \frac{1}{n_i}\lambda_2^{(i)}, \ldots, \frac{1}{n_i}\lambda_{n_i}^{(i)}\right\}_{i=1}^{m}$$
and holds Schur's condition (61). Since
$$\left\{\frac{1}{n_i}\lambda_1^{(i)}, \frac{1}{n_i}\lambda_2^{(i)}, \ldots, \frac{1}{n_i}\lambda_{n_i}^{(i)}\right\}_{i=1}^{m}$$
are the eigenvalues of $\bigoplus_{i=1}^{m}\frac{1}{n_i}\boldsymbol{\Phi}_i$, we complete the proof.

C. Proof of Theorem 2

Using Theorem 7 in Appendix II, we get (155), shown at the top of the page. Combining (24), (25), and (155), together with the fact that $\mathbb{E}\{\|\boldsymbol{H}\|_{\mathrm{F}}^2\} = n_{\mathrm{T}}n_{\mathrm{R}}$, yields (31).

D. Proof of Theorem 3

In this case, the ergodic capacity (or Shannon-sense mean capacity) is given by the well-known expression [2]–[4]

$$C(\bar{\gamma}) = \mathbb{E}\left\{\log_2\det\left(\boldsymbol{I}_{n_{\mathrm{R}}} + \frac{\bar{\gamma}}{n_{\mathrm{T}}}\boldsymbol{H}\boldsymbol{H}^{\dagger}\right)\right\}\ \text{bits/s/Hz} \tag{156}$$

which is achieved by the complex Gaussian input $\boldsymbol{X} \sim \tilde{\mathcal{N}}_{n_{\mathrm{T}},N_{\mathrm{c}}}\left(\boldsymbol{0}_{n_{\mathrm{T}}\times N_{\mathrm{c}}}, \frac{\mathcal{P}}{n_{\mathrm{T}}}\boldsymbol{I}_{n_{\mathrm{T}}}, \boldsymbol{I}_{N_{\mathrm{c}}}\right)$.

From [53, eq. (35)] and [53, Theorem 9], we get

$$\left.\frac{E_{\mathrm{b}}}{N_0}\right|_{\min} = \frac{n_{\mathrm{T}}\log_e 2}{\mathbb{E}\left\{\|\boldsymbol{H}\|_{\mathrm{F}}^2\right\}} = \frac{\log_e 2}{n_{\mathrm{R}}} \tag{157}$$

and

$$S_0 = \frac{2\left(\mathbb{E}\left\{\|\boldsymbol{H}\|_{\mathrm{F}}^2\right\}\right)^2}{\mathbb{E}\left\{\mathrm{tr}\left[(\boldsymbol{H}\boldsymbol{H}^{\dagger})^2\right]\right\}} = \frac{2(n_{\mathrm{T}}n_{\mathrm{R}}n_{\mathrm{S}})^2}{\mathbb{E}_{\boldsymbol{\Xi}_1,\boldsymbol{\Xi}_2}\left\{\mathrm{tr}\left[(\boldsymbol{\Xi}_1\boldsymbol{\Xi}_2\boldsymbol{\Xi}_2^{\dagger}\boldsymbol{\Xi}_1^{\dagger})^2\right]\right\}}. \tag{158}$$

Using Definition 2 and Theorem 7 in Appendix II, (158) can be expressed in terms of the correlation figures of $\boldsymbol{\Phi}_{\mathrm{T}}$, $\boldsymbol{\Phi}_{\mathrm{R}}$, and $\boldsymbol{\Phi}_{\mathrm{S}}$ as in (39).

E. Proof of Theorem 4

Due to the channel decoupling property of OSTBCs, the Shannon capacity (bits/s/Hz) of OSTBC MIMO channels can be written as

$$C_{\mathrm{STBC}}(\bar{\gamma}) = \mathcal{R}\cdot\mathbb{E}\left\{\log_2\left(1 + \frac{\bar{\gamma}\|\boldsymbol{H}\|_{\mathrm{F}}^2}{n_{\mathrm{T}}\mathcal{R}}\right)\right\} \tag{159}$$

which is achieved by complex Gaussian inputs $x_k \sim \mathcal{CN}\left(0, \frac{\mathcal{P}}{n_{\mathrm{T}}\mathcal{R}}\right)$. From [53, eq. (35)], [53, Theorem 9], and the first two derivatives of (159) at $\bar{\gamma} = 0$, it is easy to show (44) and

$$S_0^{\mathrm{STBC}} = \frac{2\mathcal{R}}{\kappa\left(\|\boldsymbol{H}\|_{\mathrm{F}}\right)} \tag{160}$$

from which and Theorem 2, (45) follows readily.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their helpful feedback on the paper. They also wish to thank J. S. Kwak and I. Keliher for their comments and 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, Jun. 1987.

[2] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

[3] 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.

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

[5] B. Wang, J. Zhang, and A. Høst-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 29–43, Jan. 2005.

[6] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[7] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[8] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1456–1467, Jul. 1999.

[9] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block coding for wireless communications: Performance results," *IEEE J. Select. Areas Commun.*, vol. 17, no. 3, pp. 451–460, Mar. 1999.

[10] O. Tirkkonen and A. Hottinen, "Square-matrix embeddable space-time block codes for complex signal constellations," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 384–395, Feb. 2002.

[11] W. Su and X.-G. Xia, "Two generalized complex orthogonal space-time block codes of rates $7/11$ and $3/5$ for 5 and 6 transmit antennas," *IEEE Trans. Inform. Theory*, vol. 49, no. 1, pp. 313–316, Jan. 2003.

[12] H. Wang and X.-G. Xia, "Upper bounds of rates of complex orthogonal space-time block codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2788–2796, Oct. 2003.

[13] X.-B. Liang and X.-G. Xia, "On the nonexistence of rate-one generalized complex orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 49, no. 11, pp. 2984–2989, Nov. 2003.

[14] H. Jafarkhani, "A quasi-orthogonal space-time block code," *IEEE Trans. Commun.*, vol. 49, no. 1, pp. 1–4, Jan. 2001.

[15] B. A. Sethuraman, B. S. Rajan, and V. Shashidhar, "Full-diversity, high-rate space-time block codes from division algebras," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2596–2616, Oct. 2003.

[16] W. Su and X.-G. Xia, "Signal constellations for quasi-orthogonal space-time block codes with full diversity," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2331–2347, Oct. 2004.

[17] 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.

[18] 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. Inf. Theory*, vol. 48, no. 3, pp. 637–650, Mar. 2002.

[19] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO channel model with experimental validation," *IEEE J. Select. Areas Commun.*, vol. 20, no. 6, pp. 1211–1226, Aug. 2002.

[20] M. T. Ivrlač, W. Utschick, and J. A. Nossek, "Fading correlations in wireless MIMO communication systems," *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 819–828, Jun. 2003.

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

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

[23] H. Shin, M. Z. Win, J. H. Lee, and M. Chiani, "On the capacity of doubly correlated MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 8, pp. 2253–2265, Aug. 2006.

[24] D. Gesbert, H. Bölcskei, D. A. Gore, and A. J. Paulraj, "Outdoor MIMO wireless channels: Models and performance prediction," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 1926–1934, Dec. 2002.

[25] D. Chizhik, G. J. Foschini, M. J. Gans, and R. A. Valenzuela, "Keyholes, correlations, and capacities of multielement transmit and receive antennas," *IEEE Trans. Wireless Commun.*, vol. 1, no. 2, pp. 361–368, Apr. 2002.

[26] A. F. Molisch, M. Steinbauer, M. Toeltsch, E. Bonek, and R. S. Thomä, "Capacity of MIMO systems based on measured wireless channels," *IEEE J. Select. Areas Commun.*, vol. 20, no. 3, pp. 561–569, Apr. 2002.

[27] A. F. Molisch, "A generic model for MIMO wireless propagation channels in macro- and microcells," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 61–70, Jan. 2004.

[28] S. Loyka and A. Kouki, "On MIMO channel capacity, correlations, and keyholes: Analysis of degenerate channels," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 1886–1888, Dec. 2002.

[29] P. Almers, F. Tufvesson, and A. F. Molisch, "Measurement of keyhole effect in a wireless multiple-input multiple-output (MIMO) channel," *IEEE Commun. Lett.*, vol. 7, no. 8, pp. 373–375, Aug. 2003.

[30] X. W. Cui and Z. M. Feng, "Lower capacity bound for MIMO correlated fading channels with keyhole," *IEEE Commun. Lett.*, vol. 8, no. 8, pp. 500–502, Aug. 2004.

[31] A. G. Burr, "Capacity bounds and estimates for the finite scatterers MIMO wireless channel," *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 812–818, Jun. 2003.

[32] M. Debbah and R. R. Müller, "MIMO channel modeling and the principle of maximum entropy," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1667–1690, May 2005.

[33] A. S. Y. Poon, R. W. Brodersen, and D. N. C. Tse, "Degrees of freedom in multiple-antenna channels: A signal space approach," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 523–536, Feb. 2005.

[34] A. S. Y. Poon, D. N. C. Tse, and R. W. Brodersen, "Impact of scattering on the capacity, diversity, and progation range of multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1087–1100, Mar. 2006.

[35] E. A. Jorswieck and A. Sezgin, "Impact of spatial correlation on the performance of orthogonal space-time block codes," *IEEE Commun. Lett.*, vol. 8, no. 1, pp. 21–23, Jan. 2004.

[36] J. Wang, M. K. Simon, M. P. Fitz, and K. Yao, "On the performance of space-time codes over spatially correlated Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 52, no. 6, pp. 877–881, Jun. 2004.

[37] M. Fozunbal, S. W. McLaughlin, R. W. Schafer, and J. M. Landsberg, "On space-time coding in the presence of spatial-temporal correlation," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1910–1926, Sep. 2004.

[38] R. U. Nabar, H. Bölcskei, and A. J. Paulraj, "Diversity and outage performance in space-time block coded Ricean MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2519–2532, Sep. 2005.

[39] A. Maaref and S. Aïssa, "Exact capacity and symbol-error probability analysis for STBC in spatially correlated MIMO nakagami fading channels," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM'05)*, St. Louis, MO, Nov. 2005, pp. 3649–3653.

[40] A. Maaref and S. Aïssa, "Capacity of space-time block codes in MIMO Rayleigh fading channels with adaptive transmission and estimation errors," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, Sep. 2005.

[41] H. Shin and J. H. Lee, "Effect of keyholes on the symbol error rate of space-time block codes," *IEEE Commun. Lett.*, vol. 7, no. 1, pp. 27–29, Jan. 2003.

[42] H. Shin and J. H. Lee, "Performance analysis of space—time block codes over keyhole Nakagami-m fading channels," *IEEE Trans. Veh. Technol.*, vol. 53, no. 2, pp. 351–362, Mar. 2004.

[43] A. M. N. Nasrabadi, H. R. Bahrami, and S. H. Jamali, "Space-time trellis codes for keyhole channels: Performance criterion and code design," *Electron. Lett.*, vol. 40, no. 1, pp. 53–55, Jan. 2004.

[44] S. Sanayei, A. Hedayat, and A. Nosratinia, "Space-time codes in keyhole channels: Analysis and design," in *Proc. IEEE Global Commununications Conf. (GLOBECOM'04)*, Dallas, TX, Nov. 2004, pp. 3768–3772.

[45] T. K. Y. Lo, "Maximum ratio transmission," *IEEE Trans. Commun.*, vol. 47, no. 10, pp. 1458–1461, Oct. 1999.

[46] M. Kang and M.-S. Alouini, "Largest eigenvalue of complex Wishart matrices and performance analysis of MIMO MRC systems," *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 418–426, Apr. 2003.

[47] P. A. Dighe, R. K. Mallik, and S. S. Jamuar, "Analysis of transmit-receive diversity in Rayleigh fading," *IEEE Trans. Commun.*, vol. 51, no. 4, pp. 694–703, Apr. 2003.

[48] A. J. Grant, "Performance analysis of transmit beamforming," *IEEE Trans. Commun.*, vol. 53, no. 4, pp. 738–744, Apr. 2005.

[49] M. K. Simon and M.-S. Alouini, *Digital Communication Over Fading Channels: A Unified Approach to Performance Analysis*. New York: Wiley, 2000.

[50] M. Z. Win and J. H. Winters, "Analysis of hybrid selection/maximal-ratio combining in Rayleigh fading," *IEEE Trans. Commun.*, vol. 47, no. 12, pp. 1773–1776, Dec. 1999.

[51] M. Z. Win and Z. A. Kostić, "Impact of spreading bandwidth on Rake reception in dense multipath channels," *IEEE J. Sel. Areasa Commun.*, vol. 17, no. 10, pp. 1794–1806, Oct. 1999.

[52] M. Z. Win, "Distribution-invariant monotonicity theorems on multichannel diversity," *IEEE Trans. Wireless Commun.*, to be published.

[53] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.

[54] A. Lozano, A. M. Tulino, and S. Verdú, "Multiple-antenna capacity in the low-power regime," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2527–2544, Oct. 2003.

[55] M. Chiani, M. Z. Win, and A. Zanella, "Error probability for optimum combining of M-ary PSK signals in the presence of interference and noise," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1949–1957, Nov. 2003.

[56] M. Chiani, M. Z. Win, and A. Zanella, "On optimum combining of M-PSK signals with unequal-power interferers and noise," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 44–47, Jan. 2005.

[57] H. Özelik, M. Herdin, W. Weichselberger, J. Wallace, and E. Bonek, "Deficiencies of 'Kronecker' MIMO radio channel model," *Electon. Lett.*, vol. 39, no. 16, pp. 1209–1210, Aug. 2003.

[58] W. Weichselberger, M. Herdin, H. Özelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation of both link ends," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 90–100, Jan. 2006.

[59] U. Charash, "Reception through Nakagami fading multipath channels with random delays," *IEEE Trans. Commun.*, vol. COM-27, no. 4, pp. 657–670, Apr. 1979.

[60] Z. Wang and G. B. Giannakis, "A simple and general parameterization quantifying performance in fading channels," *IEEE Trans. Commun.*, vol. 51, no. 8, pp. 1389–1398, Aug. 2003.

[61] M.-S. Alouini and M. K. Simon, "Dual diversity over correlated lognormal fading channels," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 1946–1959, Dec. 2002.

[62] B. Holter and G. E. Øien, "On the amount of fading in MIMO diversity systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2498–2507, Sep. 2005.

[63] K. Pearson, "Das Fehlergesetz und Seine Verallgemeinerungen Durch Fechner und Pearson. A Rejoinder," *Biometrika*, vol. 4, no. 1/2, pp. 169–212, Jun. 1905.

[64] S. Shamai (Shitz) and S. Verdú, "The impact of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1302–1327, May 2001.

[65] I. Schur, "Über eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie," *Sitzungsbericht der Berliner Mathematischen Gesellschaft*, vol. 22, pp. 9–20, 1923.

[66] L. A. Shepp, AT&T Labs., 1999, private conversation.

[67] T. Ando, "Majorizations, doubly stochastic matrices, and comparison of eigenvalues," *Linear Algebra Appl.*, vol. 118, pp. 163–248, Jun. 1989.

[68] R. B. Bapat and V. S. Sunder, "On majorization and Schur products," *Linear Algebra Appl.*, vol. 72, pp. 107–117, Dec. 1985.

[69] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. New York: Academic, 1979.

[70] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1991.

[71] S. A. Jafar and A. Goldsmith, "Multiple-antenna capacity in correlated Rayleigh fading with channel covariance information," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 990–997, May 2005.

[72] J. N. Pierce and S. Stein, "Multiple diversity with nonindependent fading," *Proc. IRE*, vol. 48, no. 1, pp. 89–104, Jan. 1960.

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

[74] A. T. James, "The distribution of the latent roots of the covariance matrix," *Ann. Math. Statist.*, vol. 31, no. 1, pp. 151–158, Mar. 1960.

[75] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Ann. Math. Statist.*, vol. 35, no. 2, pp. 475–501, Jun. 1964.

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

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

[78] A. Takemura, *Zonal Polynomials*. Hayward, CA: Inst. Math. Stat., 1984, vol. 4, Lecture Note—Monograph Series.

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

[80] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, 2nd ed. New York: Oxford Univ. Press, 1999.

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

[82] G. L. Turin, "The characteristic function of Hermitian quadratic forms in complex normal variables," *Biometrika*, vol. 47, pp. 199–201, Jun. 1960.

[83] X. Dong, N. C. Beaulieu, and P. H. Wittke, "Error probabilities of two-dimensional M-ary signaling in fading," *IEEE Trans. Commun.*, vol. 47, no. 3, pp. 352–355, Mar. 1999.