

On the Mutual Information Distribution of OFDM-Based Spatial Multiplexing: Exact Variance and Outage Approximation

Matthew R. McKay, *Member, IEEE*,
 Peter J. Smith, *Senior Member, IEEE*,
 Himal A. Suraweera, *Member, IEEE*, and
 Iain B. Collings, *Senior Member, IEEE*

Abstract—This communication considers the distribution of the mutual information of frequency-selective spatially uncorrelated Rayleigh fading multiple-input–multiple-output (MIMO) channels. Results are presented for orthogonal frequency-division multiplexing (OFDM)-based spatial multiplexing. New exact closed-form expressions are derived for the variance of the mutual information. In contrast to previous results, our new expressions apply for systems with both arbitrary numbers of antennas and arbitrary-length channels. Simplified expressions are also presented for high and low signal-to-noise ratio (SNR) regimes. The analytical variance results are used to provide accurate analytical approximations for the distribution of the mutual information, and the outage capacity.

Index Terms—multiple-input–multiple-output (MIMO) systems, mutual information, orthogonal frequency-division multiplexing (OFDM).

I. INTRODUCTION

Multiple-input–multiple-output (MIMO) antenna technology has emerged as an effective technique for significantly improving the capacity of wireless communication systems. A great deal of work has been done on analyzing the MIMO capacity in various flat-fading channel scenarios, since the pioneering work of [1] and [2]. In particular, the mean (ergodic) capacity has now been comprehensively investigated (e.g., see [3]–[18] and references therein).

In addition, the outage capacity has also been investigated for flat-fading channels. This is an important capacity measure for systems with stringent delay constraints, and it also provides information about the system diversity [19]. With the exception of the exact two/three antenna results presented in [20] and [21], outage capacity analysis has typically involved approximating the distribution of the mutual information, since exact closed-form solutions are not forthcoming. It has been shown that the Gaussian distribution provides a good approximation in many cases [5], [8], [13], [22], [23].

In this communication, we consider frequency-selective MIMO channels, which are applicable for many current high data-rate wireless systems. We focus on MIMO orthogonal frequency-division

Manuscript received May 12, 2006; revised July 11, 2007. The material in this paper was presented in part at IEEE International Conference on Communications, Glasgow, U.K., June 2007.

M. R. McKay is with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong (e-mail: eemckay@ust.hk).

P. J. Smith is with the Department of Electrical and Computer Engineering, University of Canterbury, Private Bag 4800, Christchurch, New Zealand (e-mail: p.smith@elec.canterbury.ac.nz).

H. A. Suraweera was with the Department of Electrical and Computer Engineering, Monash University, Melbourne, Vic. 3800, Australia. He is now with the Center for Telecommunications and Microelectronics (T μ E), Victoria University, Melbourne City MC, Vic. 8001, Australia (e-mail: himal.suraweera@vu.edu.au).

I. B. Collings is with the ICT Centre, CSIRO, Marsfield, N.S.W. 2122, Australia (e-mail: iain.collings@csiro.au).

Communicated by H. Boche, Associate Editor for Communications.

Digital Object Identifier 10.1109/TIT.2008.924685

TABLE I
MIMO–OFDM PRESENCE IN STANDARDS

Standard	Technology
WLAN IEEE 802.11n	OFDM
WiMAX IEEE 802.16-2004	OFDM/OFDMA
WiMAX IEEE 802.16e	OFDMA
MBWA IEEE 802.20	OFDM
WRAN IEEE 802.22	OFDM
3GPP Release 8	OFDMA

multiplexing (OFDM) systems, since they form the underlying technology for a many emerging MIMO standards, as summarized in Table I, and consider spatial multiplexing transmission. Despite their key practical significance however, for these systems (and indeed frequency-selective MIMO channels in general), there are relatively few analytic MIMO capacity results. The ergodic capacity (average mutual information) was considered in [19] and [24]–[28], assuming Rayleigh and Rician channels, respectively, and was found to be easily obtained by summing the equivalent flat-fading ergodic MIMO capacity of each individual OFDM subcarrier. In contrast, the outage capacity does not decompose in this way.

Calculating the outage capacity for frequency-selective channels is difficult due to the nonnegligible correlations between subcarrier channel matrices. As such, the investigation of outage capacity has usually been performed using simulation studies [19], [29], [30]. It appears that the only current analytical outage capacity results for frequency-selective channels are presented in [31]–[33], all of which derive a Gaussian approximation for the mutual information distribution. The results in [31], however, are based on deriving exact expressions for the mutual information variance of single-input–single-output (SISO) channels only, whereas the results in [33] and [32] are based on approximating the mutual information variance using asymptotic methods. Specifically, [33] considers multiple-input–single-output (MISO) channels with asymptotically large channel lengths, whereas [32] considers MIMO channels with infinite numbers of transmit and receive antennas. We note also that for the extreme frequency-selective fading case, i.e., where the MIMO subcarrier matrices are independent across frequency, the variance of the mutual information could be easily calculated by adapting known MIMO flat-fading variance results given, for example, in [8] and [15]. For many practical systems, however, the subcarrier channels are typically highly correlated across frequency, and this approach cannot be applied.

In this communication, we consider MIMO–OFDM-based spatial multiplexing systems with finite numbers of antennas and operating over spatially uncorrelated Rayleigh fading channels with finite delay spreads. We first derive new exact closed-form expressions for the mutual information variance. We also give explicit reduced formulas for the specific cases of single-input–multiple-output (SIMO), MISO, and SISO systems. Moreover, simplified closed-form expressions are derived for the variance in the high and low signal-to-noise ratio (SNR) regimes.

Based on the new analytic variance results, along with known analytic mean results, we then present new approximations to the mutual information distribution of OFDM-based spatial multiplexing systems. In particular, we present a new closed-form Gaussian approximation, which is shown to be extremely accurate for many different system and channel scenarios. In the low SNR regime, we also present a new analytic Gamma approximation, which we show to be more accurate than the Gaussian approximation in this case.

Finally, we use the analytic Gaussian approximation to estimate the outage capacity. We find that the approximation is very accurate, and

show that for outage levels of practical interest, the outage capacity depends heavily on the delay spread of the channel.

This communication is organized as follows. In Section II, we describe the frequency-selective MIMO channel model, the OFDM-based spatial multiplexing signal model, and the associated mutual information. In Section III, we present the main analytical contributions of this communication, namely, analytical expressions for the variance of the mutual information. The proofs are relegated to the appendixes. In Section IV, we approximate the distribution of the mutual information and investigate the outage capacity.

The following notation is used throughout this communication. Matrices are represented with uppercase boldface and vectors with lowercase boldface. The superscripts $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^\dagger$ indicate matrix transpose, complex conjugate, and complex conjugate transpose, respectively. The matrix \mathbf{I}_p denotes a $p \times p$ identity matrix. We use $\det(\cdot)$ and $\text{tr}(\cdot)$ to represent the matrix determinant and trace operations, respectively. The operator $E[\cdot]$ denotes expectation and $\text{Var}(\cdot)$ denotes variance. The real Gaussian distribution with zero-mean and unit-variance is denoted $\mathcal{N}(0, 1)$, the corresponding complex circularly symmetric Gaussian distribution is denoted $\mathcal{CN}(0, 1)$, and the chi-square distribution with r degrees of freedom is denoted χ_r^2 .

II. OFDM-BASED SPATIAL MULTIPLEXING SYSTEMS

A. Channel and Signal Model

We consider a single-user OFDM-based spatial multiplexing system employing N_t transmit antennas, N_r receive antennas, and N subcarriers. The channel is assumed to be frequency-selective and is modeled as a length- L finite impulse response (FIR) filter (as in [19] and [25]), for which the discrete-time input–output relation is given by [25]

$$\mathbf{y}[q] = \sum_{p=0}^{L-1} \sigma_p \mathbf{H}[p] \mathbf{x}[q-p] + \mathbf{n}[q] \quad (1)$$

where $\mathbf{x}[q] \in \mathcal{C}^{N_t \times 1}$ is the signal vector transmitted at sample index q , $\mathbf{y}[q] \in \mathcal{C}^{N_r \times 1}$ is the corresponding received signal vector, and $\mathbf{n}[q] \in \mathcal{C}^{N_r \times 1}$ is the noise vector containing independent elements $\sim \mathcal{CN}(0, 1)$. Also, σ_p , for $p = 0, \dots, L-1$, represents the channel power delay profile and is normalized according to

$$\sum_{p=0}^{L-1} \sigma_p^2 = 1. \quad (2)$$

The $N_r \times N_t$ random matrices $\mathbf{H}[p]$, for $p = 0, \dots, L-1$, represent the MIMO channel impulse response. These matrices are assumed to be mutually uncorrelated, and are assumed to be known perfectly at the receiver but are unknown at the transmitter. The channel is assumed to be quasi-static, remaining constant for the duration of a codeword, but changing independently from codeword to codeword. Throughout this communication, we assume that the channel elements exhibit spatially uncorrelated Rayleigh fading,¹ in which case each $\mathbf{H}[p]$ contains independent elements $\sim \mathcal{CN}(0, 1)$.

At the transmitter, the time-domain input sequence $\mathbf{x}[q]$ is generated as N_t parallel OFDM symbols. The symbols for each antenna are OFDM modulated using an N -point inverse fast Fourier transform (IFFT) prior to transmission. At the receiver, OFDM demodulation is performed at each receive antenna using an N -point FFT. A key advantage of OFDM-based spatial multiplexing is that equalization is simple,

¹Note that a number of recent investigations have studied the impact of spatial correlation on MIMO capacity/mutual information (see, e.g., [8], [9], [12], and [13]). We do not follow this line of work here, however, since our primary focus is to study the impact of frequency-selective fading on mutual information, in which case the effect of correlation is observed across frequency.

since the frequency-selective MIMO channel is transformed into N orthogonal flat-fading MIMO subchannels via the IFFT/FFT processing.

To maintain orthogonality in the presence of intersymbol interference caused by multipath, OFDM systems typically employ a cyclic prefix extension. Assuming that the cyclic prefix is longer than the delay spread of the channel, we can write the equivalent frequency domain input–output model for OFDM-based spatial multiplexing as follows:

$$\mathbf{r}_k = \mathbf{H}_k \mathbf{a}_k + \mathbf{n}_k, \quad k = 0, \dots, N-1 \quad (3)$$

where \mathbf{a}_k is the transmitted vector for the k th subcarrier, assumed to be independent identically distributed (i.i.d.) Gaussian with covariance matrix $E[\mathbf{a}_k \mathbf{a}_k^\dagger] = \frac{\gamma}{N_t} \mathbf{I}_{N_t}$, \mathbf{r}_k is the received vector for the k th subcarrier, and \mathbf{n}_k is the corresponding complex additive white Gaussian noise (AWGN) vector satisfying $E[\mathbf{n}_k \mathbf{n}_k^\dagger] = \mathbf{I}_{N_r} \delta[k-\ell]$, where $\delta[\cdot]$ is the Kronecker-delta function. Also, \mathbf{H}_k is the k th subcarrier channel matrix given by

$$\mathbf{H}_k = \sum_{p=0}^{L-1} \sigma_p \mathbf{H}[p] \exp\left(-j2\pi \frac{k}{N} p\right) \quad (4)$$

containing independent entries $(\mathbf{H}_k)_{i,j} \sim \mathcal{CN}(0, 1)$. Note that due to the finite-length impulse response, correlation exists between different subcarrier channel matrices. Using (4), the correlation coefficients between the channel elements on two arbitrary subcarriers k and ℓ is easily derived as follows (see also [34]):

$$\begin{aligned} \rho_{k-\ell} &= E[(\mathbf{H}_k)_{i,j} (\mathbf{H}_\ell)_{i',j'}^*] \\ &= \sum_{p=0}^{L-1} \sigma_p^2 e^{-j2\pi(k-\ell)p/N} \delta[i-i'] \delta[j-j'] \end{aligned} \quad (5)$$

for all i, j, i', j' . As expected, these frequency correlation coefficients depend only on the difference between subcarriers (i.e., $k-\ell$) and not on the subcarriers themselves.

Note that with the above model, the SNR per receive antenna per subcarrier (henceforth referred to as “the SNR”) is given by γ .

B. Mutual Information

The focus of this communication is on the statistics of the mutual information of OFDM-based spatial multiplexing systems. It is now well known that the instantaneous mutual information in bits per second per Hertz (b/s/Hz) for a given channel realization is given by [19]

$$\mathcal{I}_{\text{ofdm}} = \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{I}_k \quad (6)$$

where \mathcal{I}_k is the instantaneous mutual information for the k th OFDM subcarrier, given by

$$\mathcal{I}_k = \log_2 \det \left(\mathbf{I}_{N_r} + \frac{\gamma}{N_t} \mathbf{H}_k \mathbf{H}_k^\dagger \right). \quad (7)$$

Note that the loss in mutual information due to the cyclic prefix has been neglected in (7). The mean (ergodic) mutual information is given by

$$E[\mathcal{I}_{\text{ofdm}}] = \frac{1}{N} \sum_{k=0}^{N-1} E[\mathcal{I}_k]. \quad (8)$$

It is obvious that (8) is equivalent to the ergodic mutual information of a flat-faded channel, for which case closed-form expressions are now available [6], [15], [35].

III. VARIANCE OF THE MUTUAL INFORMATION

In this section, we derive new closed-form expressions for the variance of the mutual information of OFDM-based spatial multiplexing. Our results are exact and apply for arbitrary finite system and channel parameters. We also present simplified expressions for the variance in the high and low SNR regimes, and give explicit reduced variance expressions for the cases of MISO, SIMO, and SISO systems. These results will be subsequently used in Section IV for providing accurate approximations to the mutual information distribution, and to the outage capacity.

A. Exact Analysis at All SNRs

The following theorem presents an exact expression for the variance of the mutual information of MIMO-OFDM systems.

Theorem 1: The variance of the mutual information of MIMO-OFDM systems is given by

$$\begin{aligned} \text{Var}(\mathcal{I}_{\text{ofdm}}) &= \frac{(\log_2(e))^2}{\Gamma_m(n)\Gamma_m(m)} \\ &\times \left(\frac{2}{N^2} \sum_{d=1}^{N-1} (N-d)\varphi(\rho_d) + \sum_{r=1}^m \sum_{s=1}^m \frac{\det(\mathbf{B}_{r,s})}{N} \right. \\ &\quad \left. - \frac{(\sum_{r=1}^m \det(\mathbf{A}_r))^2}{\Gamma_m(n)\Gamma_m(m)} \right) \end{aligned} \quad (9)$$

where $m = \min(N_r, N_t)$, $n = \max(N_r, N_t)$, and $\Gamma_m(\cdot)$ is the normalized complex multivariate gamma function defined as

$$\Gamma_m(n) = \prod_{i=1}^m \Gamma(n-i+1) \quad (10)$$

and

$$\varphi(\rho_d) = \begin{cases} \frac{(\sum_{r=1}^m \det(\mathbf{A}_r))^2}{\Gamma_m(n)\Gamma_m(m)}, & |\rho_d| = 0 \\ \sum_{r=1}^m \sum_{s=1}^m e^{2Nt/\gamma} \det(\mathbf{C}_{r,s}(\rho_d)), & 0 < |\rho_d| < 1 \\ \sum_{r=1}^m \sum_{s=1}^m \det(\mathbf{B}_{r,s}), & |\rho_d| = 1 \end{cases} \quad (11)$$

The matrix \mathbf{A}_r is $m \times m$, with (i, j) th element

$$(\mathbf{A}_r)_{i,j} = \begin{cases} b!, & \text{for } j \neq r \\ b!e^{Nt/\gamma} g_1(b+1), & \text{for } j = r. \end{cases} \quad (12)$$

The matrices $\mathbf{B}_{r,s}$ and $\mathbf{C}_{r,s}(\cdot)$ are $m \times m$ with (i, j) th elements given by (13) and (14), respectively, shown at the bottom of the page. Also,

$b = n + m - i - j$, $\tau = n - m$, $z = \tau + i + j - 1$, $u = \tau + i + t$, $v = \tau + j + t$, and $G_{3,4}^{4,0}(\cdot)$ is the Meijer-G function (see [36, eq. (9.301)] for definition)

$$g_1(z) = \sum_{h=1}^z E_h \left(\frac{Nt}{\gamma} \right) \quad (15)$$

$$g_2(z) = \sum_{h=1}^z E_h \left(\frac{Nt}{\gamma(1-|\rho_d|^2)} \right) \quad (16)$$

where $E_h(\cdot)$ is the exponential integral (see [37, eq. (5.1.12)] for definition). The function $\eta_{i,j}(\cdot, \cdot)$ is defined as

$$\begin{aligned} \eta_{i,j}(f(z), \rho_d) &= \Gamma(\tau + j) \sum_{t=0}^{j-1} \binom{j-1}{t} \left(\frac{1-|\rho_d|^2}{|\rho_d|^2} \right)^t \\ &\quad \times (\tau + j - t)_{i-1} f(z - t) \end{aligned} \quad (17)$$

for an arbitrary input function f , and $(\cdot)_r$ is the Pochhammer symbol

$$\begin{aligned} (a)_r &= a \cdot (a+1) \cdots (a+r-1) \\ &= \frac{\Gamma(a+r)}{\Gamma(a)}, \quad (a)_0 = 1. \end{aligned} \quad (18)$$

Proof: See Appendix I. \square

Note that the exact variance expression in Theorem 1 can be easily evaluated since it primarily involves simple polynomial and exponential terms, as well as standard functions such as exponential integrals and Meijer-G functions, both of which are implemented as built-in procedures in various mathematical software packages such as Maple and Mathematica. We also note that although Theorem 1 involves infinite series of exponential integrals, its numerical evaluation can be made more efficient by exploiting the following recurrence relations [37, eqs. (5.1.7) and (5.1.14)]:

$$\begin{aligned} E_1(z) &= -\text{Ei}(-z) \\ E_{n+1}(z) &= \frac{1}{n}(e^{-z} - zE_n(z)) \end{aligned} \quad (19)$$

for $z > 0$. As such, only a single exponential integral must be explicitly evaluated when summing these series. Moreover, it turns out that this infinite series converges quickly and can generally be evaluated with less than 20 terms. Therefore, the computational challenge associated with this series is very low.

The following corollary presents an exact variance expression for the mutual information of SIMO/MISO-OFDM systems (i.e., cases with $m = 1$, $n > 1$). To the best of our knowledge, this result is also new.

$$\begin{aligned} (\mathbf{B}_{r,s})_{i,j} &= \begin{cases} b!, & \text{for } j \neq r \text{ and } j \neq s \\ b!e^{Nt/\gamma} g_1(b+1), & \text{for } j = r \text{ or } j = s, \text{ and } r \neq s \\ 2(Nt/\gamma)^{b+1} e^{Nt/\gamma} \sum_{t=0}^b \binom{b}{t} (-1)^t G_{3,4}^{4,0} \left(Nt/\gamma \middle|_{0,t-b-1,t-b-1,t-b-1}^{t-b,t-b,t-b} \right), & \text{for } j = r = s \end{cases} \quad (13) \\ (\mathbf{C}_{r,s}(\rho_d))_{i,j} &= \begin{cases} \eta_{i,j}(1, \rho_d), & \text{for } i \neq r, j \neq s \\ \eta_{i,j}(g_1(z), \rho_d), & \text{for } i = r, j \neq s \\ |\rho_d|^{2(i-j)} \eta_{j,i}(g_1(z), \rho_d), & \text{for } i \neq r, j = s \\ \frac{(1-|\rho_d|^2)^z}{|\rho_d|^{2(j-1)}} e^{\frac{2Nt|\rho_d|^2}{\gamma(1-|\rho_d|^2)}} \sum_{t=0}^{\infty} \frac{|\rho_d|^{2t} \Gamma(u) \Gamma(v) g_2(u) g_2(v)}{t!(\tau+t)!}, & \text{for } i = r, j = s \end{cases} \quad (14) \end{aligned}$$

Corollary 1: The variance of the mutual information of SIMO/MISO-OFDM systems is given by

$$\begin{aligned} \text{Var}(\mathcal{I}_{\text{ofdm}}) &= \frac{(\log_2(e))^2}{\Gamma(n)} \left(\frac{2}{N^2} \sum_{d=1}^{N-1} (N-d) \varphi(\rho_d) - \Gamma(n) e^{2N_t/\gamma} (g_1(n))^2 \right. \\ &\quad \left. + \frac{2e^{N_t/\gamma}}{N} \left(\frac{N_t}{\gamma} \right)^n \sum_{t=0}^{n-1} \binom{n-1}{t} (-1)^{n-1-t} \right. \\ &\quad \left. \times G_{3,4}^{4,0} \left(N_t/\gamma \middle|_{0,-t,-t-1,-t-1}^{-t,-t,-t} \right) \right) \quad (20) \end{aligned}$$

where $\varphi(\rho_d)$ is defined in (21), shown at the bottom of the page.

The following corollary presents an exact variance expression for the mutual information of SISO-OFDM systems (i.e., $m = 1, n = 1$).

Corollary 2: The variance of the mutual information of SISO-OFDM systems is given by

$$\begin{aligned} \text{Var}(\mathcal{I}_{\text{ofdm}}) &= (\log_2(e))^2 \left(\frac{2}{N^2} \sum_{d=1}^{N-1} (N-d) \varphi(\rho_d) + \frac{2e^{1/\gamma}}{\gamma N} \right. \\ &\quad \left. \times G_{3,4}^{4,0} \left(1/\gamma \middle|_{0,-1,-1,-1}^{0,0,0} \right) - e^{2/\gamma} (g_1(1))^2 \right) \quad (22) \end{aligned}$$

where

$$\varphi(\rho_d) = \begin{cases} e^{2/\gamma} (g_1(1))^2, & \text{for } |\rho_d| = 0 \\ (1 - |\rho_d|^2) e^{\frac{2}{\gamma(1-|\rho_d|^2)}} \\ \quad \times \sum_{t=0}^{\infty} |\rho_d|^{2t} (g_2(1+t))^2, & \text{for } 0 < |\rho_d| < 1 \\ \frac{2e^{1/\gamma}}{\gamma N} G_{3,4}^{4,0} \left(1/\gamma \middle|_{0,-1,-1,-1}^{0,0,0} \right), & \text{for } |\rho_d| = 1 \end{cases} \quad (23)$$

Very recently, an equivalent expression for the SISO-OFDM variance given in (22) was presented in [31].² In contrast to (22), however, the equivalent result from [31] is not expressed in closed form, and it requires the evaluation of infinite series of incomplete gamma functions.

In Fig. 1, we compare the analytical variance expression (9) with the variance obtained via Monte Carlo simulation. Results are presented for two different $N_t \times N_r$ antenna configurations as a function of the channel length L . A uniform power delay profile is assumed (i.e., $\sigma_p^2 = 1/L$, for $p = 0, \dots, L-1$), eight subcarriers are used (simply by way of example, similar results are obtained for higher numbers of subcarriers), and the SNR is set to 10 dB. In all cases, we see a pre-

²Note that this expression was not explicitly stated in [31]. It can however be obtained by following the derivation of (47) and using [31, eqs. (12), (41), and (48)]

cise agreement between the simulated and analytic curves. Moreover, the variance is seen to be largest for the system with the least antennas, regardless of the channel length. For both antenna configurations, we see that the variance reduces with increasing L , and that this reduction is most significant for small L . For example, by increasing the channel length from $L = 1$ (flat-fading) to $L = 2$, the variance for both antenna configurations is more than halved.

In Fig. 2, we plot the analytical variance expression (9) and Monte Carlo simulation results for different SNRs, as a function of L . Again we see a precise agreement between the analytical and simulated results. From this figure, we see that for a given channel length the variance of the mutual information varies monotonically with the SNR. This increase is most significant for small values of L .

B. Analysis at High SNR

The following theorem presents a closed-form expression for the variance of the mutual information of MIMO-OFDM in the high SNR regime. This result is simpler than the exact general variance result given in Theorem 1, as it does not involve any infinite series.

Theorem 2: In the high SNR regime, the variance of the mutual information of MIMO-OFDM systems is given by

$$\begin{aligned} \text{Var}^\infty(\mathcal{I}_{\text{ofdm}}) &= (\log_2(e))^2 \left(\frac{2}{N^2} \sum_{d=1}^{N-1} (N-d) \tilde{\varphi}(\rho_d) + \frac{1}{N} \sum_{t=0}^{m-1} \psi'(n-t) \right. \\ &\quad \left. - \frac{N-1}{N} \left(\sum_{t=0}^{m-1} \psi(n-t) \right)^2 \right) \quad (24) \end{aligned}$$

where

$$\tilde{\varphi}(\rho_d) = \begin{cases} \left(\sum_{t=0}^{m-1} \psi(n-t) \right)^2, & \text{for } |\rho_d| = 0 \\ \frac{\sum_{r=1}^m \sum_{s=1}^m \det(\tilde{\mathbf{C}}_{r,s}(\rho_d))}{\Gamma_m(n) \Gamma_m(m)}, & \text{for } 0 < |\rho_d| < 1 \\ \sum_{t=0}^{m-1} \psi'(n-t) \\ \quad + \left(\sum_{t=0}^{m-1} \psi(n-t) \right)^2, & \text{for } |\rho_d| = 1 \end{cases} \quad (25)$$

where $\tilde{\mathbf{C}}_{r,s}(\rho_d)$ is an $m \times m$ matrix with (i, j) th element for the cases $i \neq r$ or $j \neq s$ given by

$$\left(\tilde{\mathbf{C}}_{r,s}(\rho_d) \right)_{i,j} = \begin{cases} \eta_{i,j}(1, \rho_d), & \text{for } i \neq r, j \neq s \\ \eta_{i,j}(\psi(z), \rho_d), & \text{for } i = r, j \neq s \\ |\rho_d|^{2(i-j)} \eta_{j,i}(\psi(z), \rho_d), & \text{for } i \neq r, j = s \end{cases} \quad (26)$$

and for the case ($i = r, j = s$) by

$$\begin{aligned} \left(\tilde{\mathbf{C}}_{r,s}(\rho_d) \right)_{i,j} &= h^2(\rho_d) |\rho_d|^{2(i-j)} \eta_{j,i}(1, \rho_d) + |\rho_d|^{2(i'-j)} \eta_{j',i'} \\ &\quad \times (\xi_{|\rho_d|^2}(z-1, \rho_d) + h(\rho_d) (\eta_{i,j}(H(z-1), \rho_d) \\ &\quad - \ln(1 - |\rho_d|^2) \eta_{i,j}(1, \rho_d) + |\rho_d|^{2(i-j)} \\ &\quad \times (\eta_{j,i}(H(z-1), \rho_d) - \ln(1 - |\rho_d|^2) \eta_{j,i}(1, \rho_d))) \quad (27) \end{aligned}$$

$$\varphi(\rho_d) = \begin{cases} \Gamma(n) e^{2N_t/\gamma} (g_1(n))^2, & \text{for } |\rho_d| = 0 \\ (1 - |\rho_d|^2)^n e^{\frac{2N_t}{\gamma(1-|\rho_d|^2)}} \sum_{t=0}^{\infty} \frac{|\rho_d|^{2t} \Gamma(n+t) (g_2(n+t))^2}{t!}, & \text{for } 0 < |\rho_d| < 1 \\ 2e^{N_t/\gamma} \left(\frac{N_t}{\gamma} \right)^n \sum_{t=0}^{n-1} \binom{n-1}{t} (-1)^{n-1-t} G_{3,4}^{4,0} \left(N_t/\gamma \middle|_{0,-t,-t-1,-t-1}^{-t,-t,-t} \right), & \text{for } |\rho_d| = 1 \end{cases} \quad (21)$$

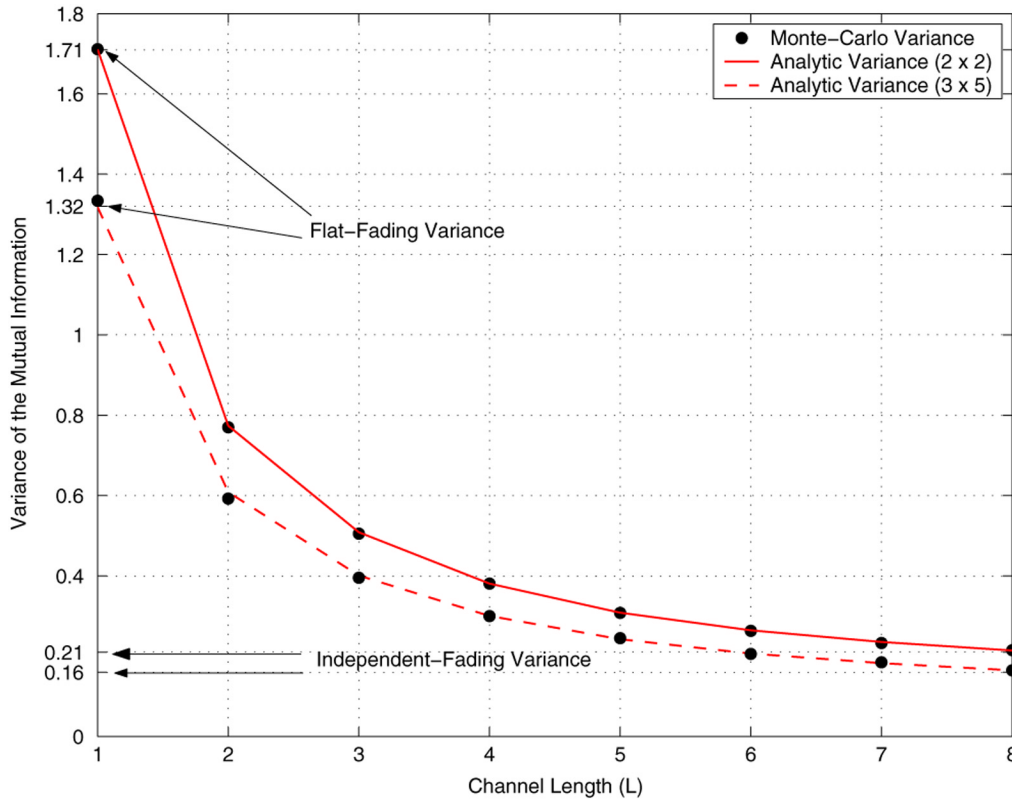


Fig. 1. Variance of the mutual information of MIMO-OFDM for different $N_t \times N_r$ antenna configurations and different channel lengths (uniform power delay profile). The “analytic variance” curves are based on (9). Eight subcarriers are considered, with SNR of 10 dB.

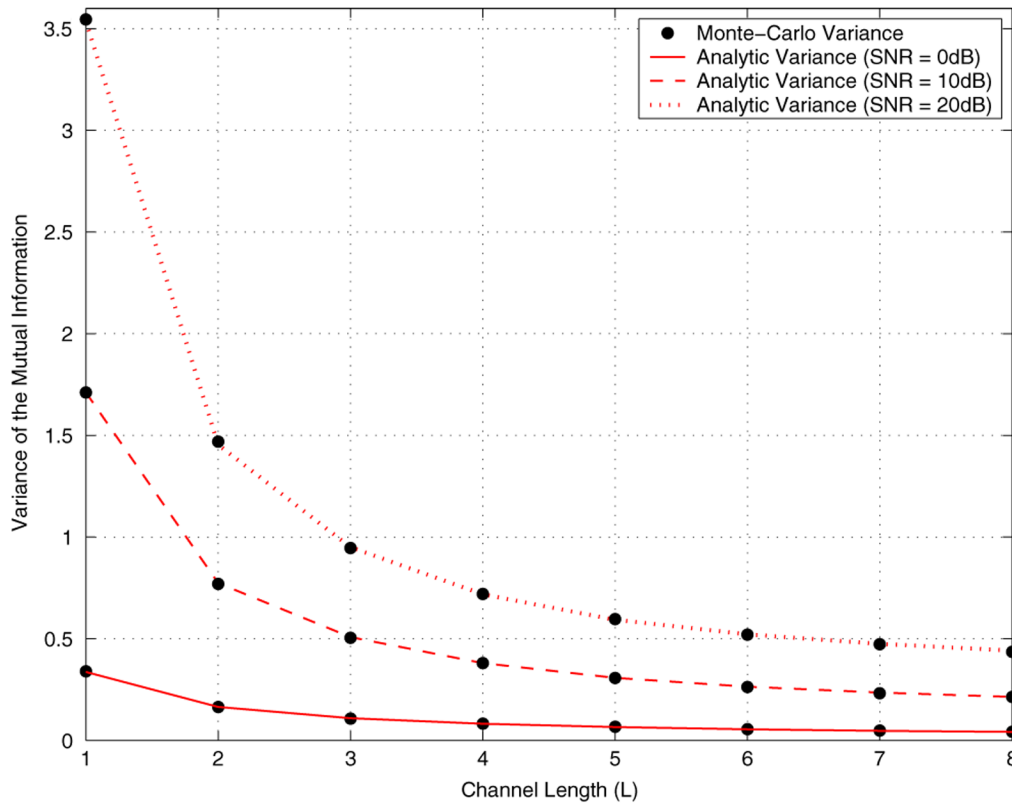


Fig. 2. Variance of the mutual information of MIMO-OFDM for different SNRs and different channel lengths (uniform power delay profile). The “analytic variance” curves are based on (9). A 2×2 system is considered with eight subcarriers.

where $i' = \max(i, j)$ and $j' = \min(i, j)$. Also, $\eta_{i,j}(\cdot)$ is defined in (17) in Theorem 1, $\xi(\cdot)$ is defined in (126), $h(\cdot)$ is given by

$$h(\rho_d) = \ln(1 - |\rho_d|^2) - \mathcal{K} \quad (28)$$

and $\mathcal{K} = 0.5772 \dots$ is the Euler–Mascheroni constant. The function $H(\cdot)$ denotes the harmonic number

$$H(z) = \begin{cases} \sum_{\ell=1}^z \frac{1}{\ell}, & \text{for } z > 0 \\ 0, & \text{for } z = 0 \end{cases} \quad (29)$$

and $\psi(\cdot)$ is the digamma function defined as [37, eq. (6.3.2)]

$$\psi(n - t) = H(n - t - 1) - \mathcal{K} \quad (30)$$

with first derivative $\psi'(\cdot)$ corresponding to the polygamma function [37, eq. (6.4.1)].

Proof: See Appendix III. \square

The following two corollaries present very simple high SNR variance expressions for the special case of SIMO/MISO and SISO systems, respectively.

Corollary 3: The variance of the mutual information of SIMO/MISO–OFDM systems at high SNR is given by

$$\begin{aligned} \text{Var}^\infty(\mathcal{I}_{\text{ofdm}}) &= (\log_2(e))^2 \left(\frac{2}{N^2} \sum_{d=1}^{N-1} (N-d) \tilde{\varphi}(\rho_d) + \frac{\psi'(n)}{N} \right) \quad (31) \end{aligned}$$

where $\tilde{\varphi}(\rho_d)$ is given by (32), shown at the bottom of the page, with $\text{Li}_2(\cdot)$ denoting the *dilogarithm* function [37, eq. (27.7.1)]. \square

Corollary 4: The variance of the mutual information of SISO–OFDM systems at high SNR is given by

$$\begin{aligned} \text{Var}^\infty(\mathcal{I}_{\text{ofdm}}) &= (\log_2(e))^2 \\ &\times \left(\frac{2}{N^2} \sum_{d=1}^{N-1} (N-d) \text{Li}_2(1 - |\rho_d|^2) + \frac{\pi^2}{6N} \right). \quad (33) \end{aligned}$$

It is important to note that the results in Theorem 2 and Corollaries 3 and 4 do not depend on the SNR. Therefore, a main insight that we can draw from these expressions is that the variance of the MIMO–OFDM mutual information converges to a deterministic limit as the SNR increases, which we have now quantified precisely. This phenomenon is illustrated in Fig. 3, where we plot the variance of the MIMO–OFDM mutual information for different $N_t \times N_r$ antenna configurations, and for different SNRs. The “analytic variance (high SNR)” dashed lines are based on (24) for the 2×3 case, (31) for the 1×2 case, and (33) for the 1×1 case. The “analytic variance (exact)” curves are based on (9) for the 2×3 case, (20) for the 1×2 case, and (22) for the 1×1 case. Monte Carlo simulated variance curves are also presented for further verification. We see that the results converge quickly in all cases.

C. Analysis at Low SNR

The following theorem presents a very simple closed-form expression for the variance of the mutual information of MIMO–OFDM in the low SNR regime.

Theorem 3: In the low SNR regime, the variance of the mutual information of MIMO–OFDM systems is given by

$$\text{Var}^0(\mathcal{I}_{\text{ofdm}}) = (\log_2(e))^2 \frac{\gamma^2 N_r}{N N_t} \left(1 + 2 \sum_{d=1}^{N-1} \frac{N-d}{N} |\rho_d|^2 \right). \quad (34)$$

Proof: See Appendix IV. \square

The following corollary gives upper and lower bounds (as a function of the frequency correlation coefficients) for the variance of the MIMO–OFDM mutual information in the low SNR regime.

Corollary 5: In the low SNR regime, the variance of the mutual information of MIMO–OFDM systems satisfies

$$\begin{aligned} \frac{1}{N} &\leq \frac{\text{Var}^0(\mathcal{I}_{\text{ofdm}})}{\text{Var}^0(\mathcal{I}_{\text{flat}})} \\ &= \frac{1}{N} \left(1 + 2 \sum_{d=1}^{N-1} \frac{N-d}{N} |\rho_d|^2 \right) \leq 1 \quad (35) \end{aligned}$$

where $\text{Var}^0(\mathcal{I}_{\text{flat}})$ denotes the mutual information variance for an i.i.d. flat-fading Rayleigh MIMO channel. The left-hand side is an equality for $|\rho_d| = 0$ (independent fading across all frequency subcarriers), and the right-hand side is an equality for $|\rho_d| = 1$ (identical fading across all subcarriers, i.e., flat-fading).

Proof: The proof follows by using

$$0 \leq |\rho_d| \leq 1 \quad (36)$$

in (34) and noting that

$$\text{Var}^0(\mathcal{I}_{\text{flat}}) = (\log_2(e))^2 \frac{\gamma^2 N_r}{N_t} \quad (37)$$

which is found by directly setting $N = 1$ in (34). \square

It is interesting to note from (35) that in the low SNR regime, the scaling of the MIMO–OFDM variance with respect to the flat-fading variance depends only on the channel delay profile, and it is independent of the number of transmit and receive antennas.

For the particular case of a uniform power delay profile (i.e., with $\sigma_p^2 = 1/L$ for all $p = 0, \dots, L-1$), we can obtain a simple insightful expression for the variance ratio in (35), as given below.

Corollary 6: For a uniform power delay profile, (35) becomes

$$\begin{aligned} \frac{1}{N} &\leq \frac{\text{Var}^0(\mathcal{I}_{\text{ofdm}})}{\text{Var}^0(\mathcal{I}_{\text{flat}})} \\ &= \frac{1}{N} \left(1 + 2 \sum_{d=1}^{N-1} \frac{N-d}{N} \left(\frac{\sin(\frac{\pi d L}{N})}{L \sin(\frac{\pi d}{N})} \right)^2 \right) \leq 1 \quad (38) \end{aligned}$$

$$\tilde{\varphi}(\rho_d) = \begin{cases} 0, & \text{for } |\rho_d| = 0 \\ \text{Li}_2(1 - |\rho_d|^2) - H^2(n-1) + 2 \sum_{b=1}^{n-1} \frac{H(b-1)}{b} \\ \quad + \sum_{b=1}^{n-1} \frac{1}{b} \left(\left(\frac{|\rho_d|^{2b-1}}{|\rho_d|^{2b}} \right)^b \ln(1 - |\rho_d|^2) - \sum_{t=0}^{b-2} \frac{\left(\frac{|\rho_d|^{2t-1}}{|\rho_d|^{2t}} \right)^{t+1}}{b-t-1} \right), & \text{for } 0 < |\rho_d| < 1 \\ \psi'(n), & \text{for } |\rho_d| = 1 \end{cases} \quad (32)$$

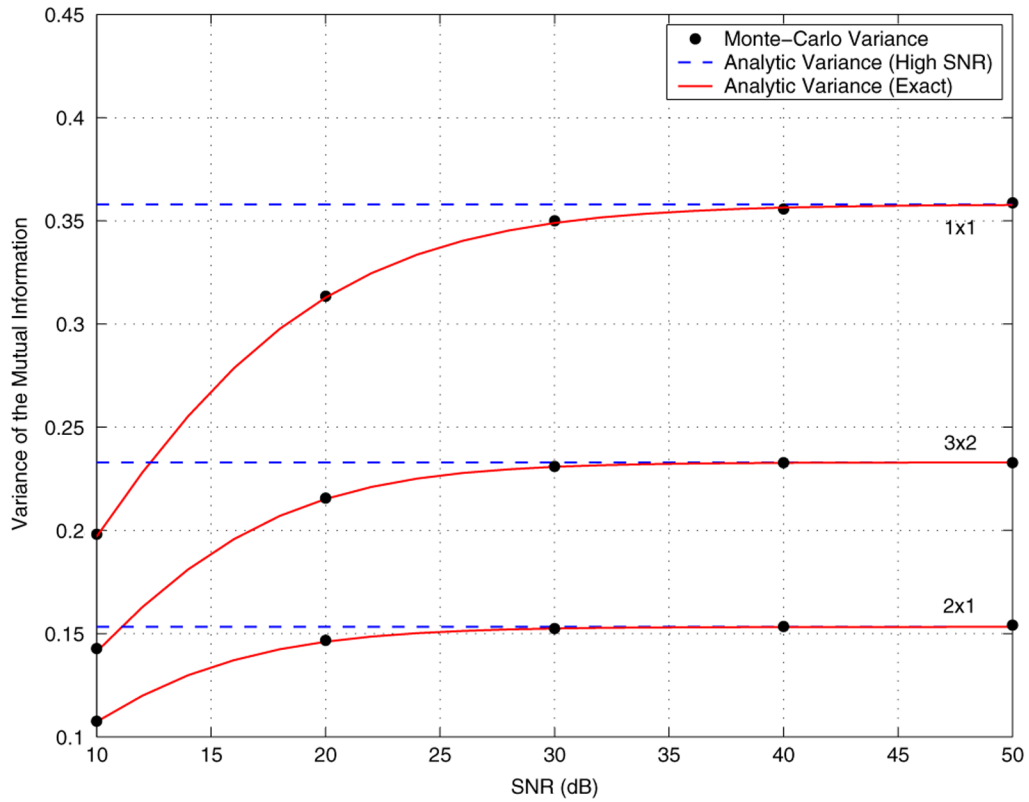


Fig. 3. Variance of the mutual information of MIMO-OFDM for different $N_t \times N_r$ antenna configurations, and for different SNRs. The “analytic variance (high SNR)” lines are based on (24) for the 2×3 case, (31) for the 1×2 case, and (33) for the 1×1 case. The “analytic variance (exact)” curves are based on (9) for the 2×3 case, (20) for the 1×2 case, and (22) for the 1×1 case. Sixteen subcarriers are considered, and the channel follows an eight-path uniform power delay profile.

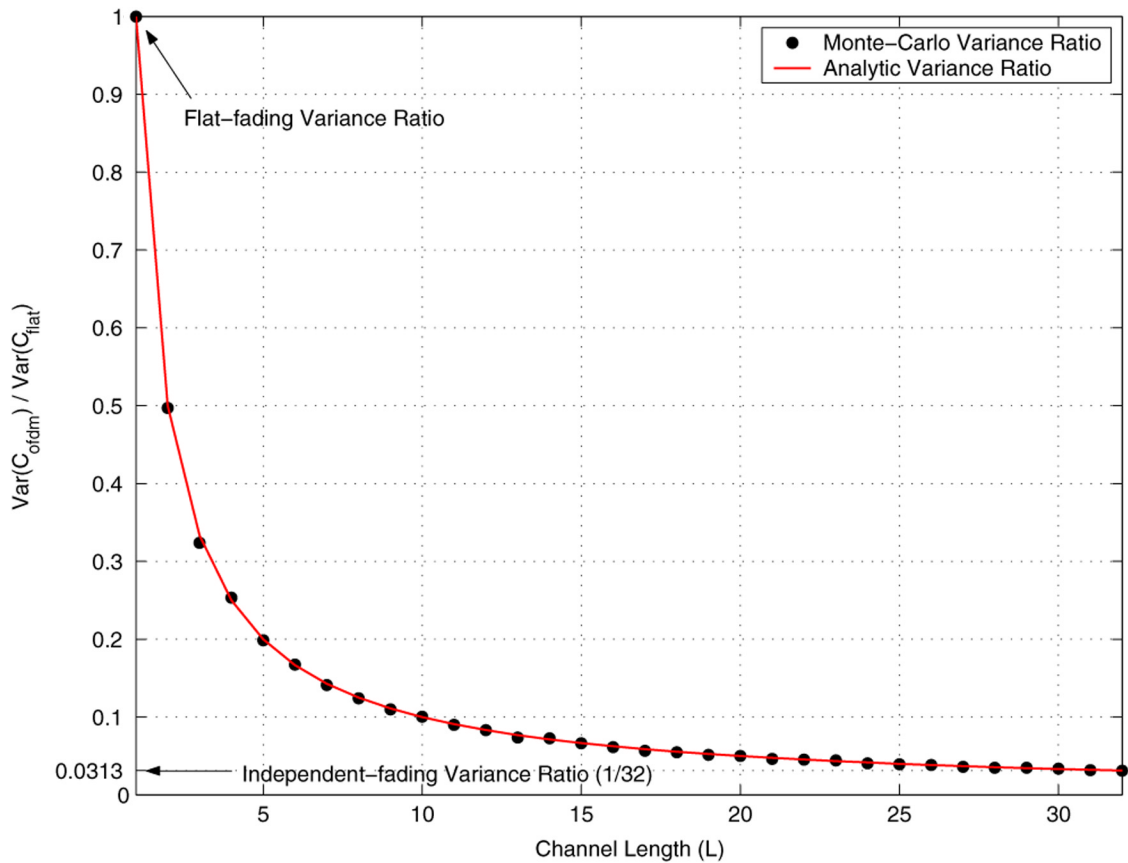


Fig. 4. Ratio of the MIMO-OFDM mutual information variance and the flat-fading MIMO mutual information variance at low SNR for different channel lengths (uniform power delay profile). The “analytic variance ratio” curve is based on (38). A 2×2 system at -25 dB is considered with 32 subcarriers.

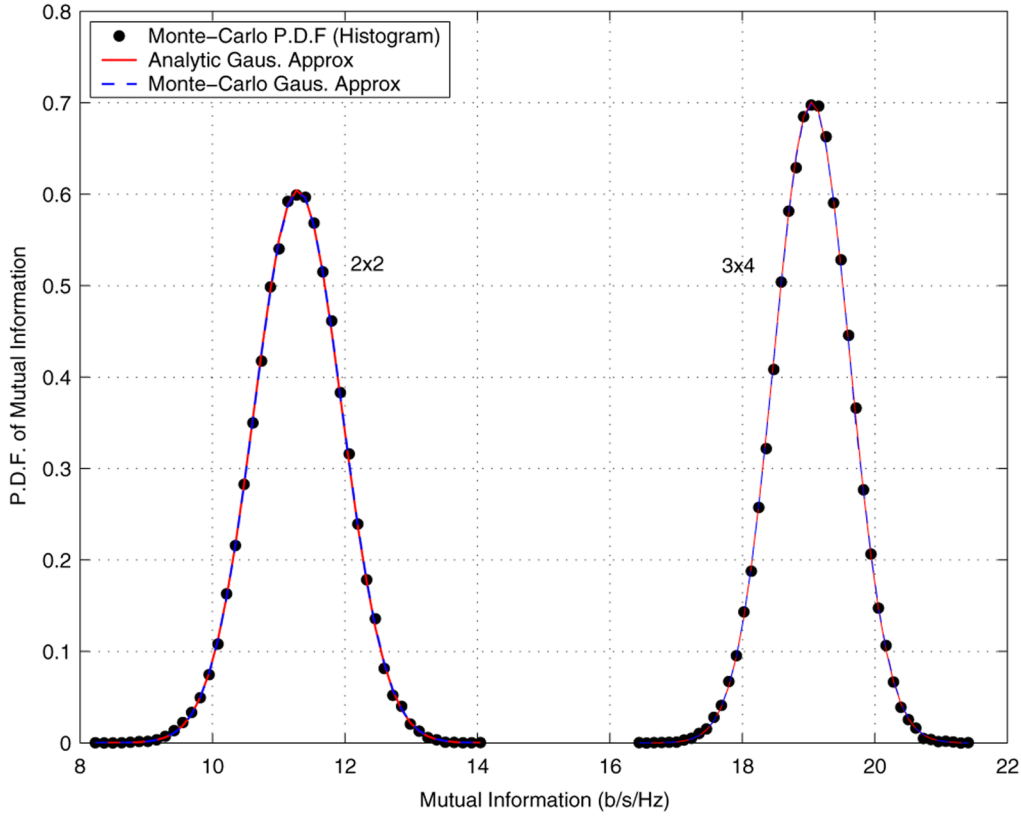


Fig. 5. The pdf of MIMO-OFDM mutual information for different $N_t \times N_r$ antenna configurations. The “analytic Gaussian approximation” curves are based on the exact mean formula (48) and exact variance formula (9). Sixty four subcarriers are considered with SNR of 20 dB. The channel follows an eight-path exponential power delay profile with $K_{\text{exp}} = 4$.

where the left-hand side is an equality for $L = N$, and the right-hand side is an equality for the case $L = 1$.

Proof: The proof follows trivially from (35) after noting that the frequency correlation coefficients (5) in this case can be expressed as [38]

$$\rho_d = \frac{\sin\left(\frac{\pi dL}{N}\right)}{L \sin\left(\frac{\pi d}{N}\right)} e^{j\frac{\pi d}{N}}. \quad (39)$$

□

The summation in (38) is of a similar type to that in [32, eq. (60)], which gave an asymptotic expression for the variance for large antenna numbers and involved the same squared-ratio terms. As mentioned in [32], as L increases, the ratio becomes more peaked as a function of d , thereby decreasing the overall sum. Thus, from (38), we see that the variance of the mutual information varies inversely with the channel delay spread in the low SNR regime. This agrees with previous observations seen via simulation studies in [19], and for the regime of large antenna numbers in [32]. These results are further corroborated in Fig. 4, where (38) is plotted as a function of the channel length L .

IV. OUTAGE APPROXIMATION OF MIMO-OFDM-BASED SPATIAL MULTIPLEXING

We now use the analytic expressions from the previous section to present and investigate approximations for the distribution of mutual information. We then use the approximations to estimate outage capacity.

Unless otherwise stated, for all results in this section, we model the channel according to the exponential power delay profile [39]

$$\sigma_p^2 = \begin{cases} \frac{1 - e^{-1/K_{\text{exp}}}}{1 - e^{-L/K_{\text{exp}}}} e^{-p/K_{\text{exp}}}, & \text{for } 0 \leq p < L \\ 0, & \text{otherwise} \end{cases} \quad (40)$$

where K_{exp} is a parameter that characterizes the rate of decay of the power delay profile as a function of p , and is loosely related to the root mean square (RMS) delay spread [39].

A. Gaussian and Gamma Approximations

We first investigate the accuracy of a Gaussian approximation for various system configurations and channel scenarios.

Fig. 5 presents the analytical Gaussian approximation for the MIMO-OFDM mutual information probability density function (pdf) based on the exact mean and variance expressions in (48) and (1), respectively, as well as empirically generated pdfs (Monte Carlo histogram), for different antenna configurations. A 64-subcarrier system is considered with SNR of 20 dB. We see that the analytic curves match the true distribution almost perfectly for both antenna configurations. We also present curves for a simulation-based Gaussian approximation (based on the mean and variance of the Monte Carlo generated histograms) for further verification. Note that these curves are indistinguishable from our new analytical Gaussian approximation curves.

Fig. 6 compares the analytical Gaussian approximation with empirically generated pdf curves, for different channel RMS delay spreads. Again we see that the analytic Gaussian approximation is accurate in

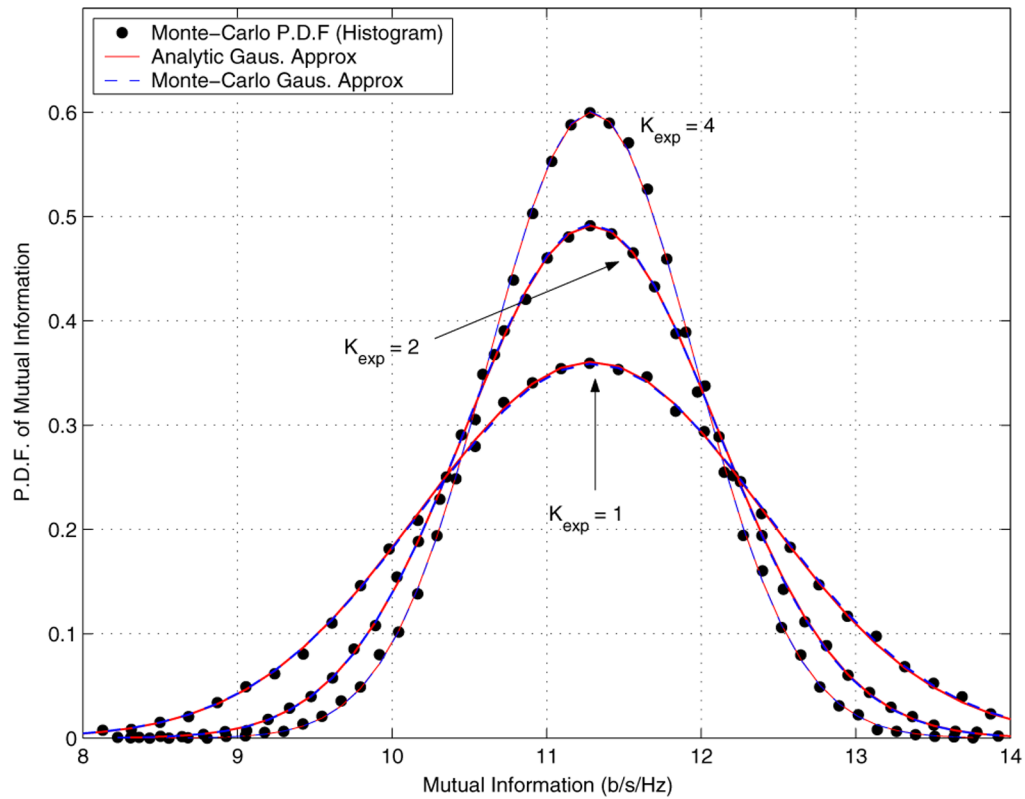


Fig. 6. The pdf of MIMO-OFDM mutual information for different RMS delay spreads (channels follow an eight-path exponential power delay profile with different K_{exp}). The “analytic Gaussian approximation” curves are based on the exact mean formula (48) and exact variance formula (9). Sixteen subcarriers are considered with SNR of 20 dB.

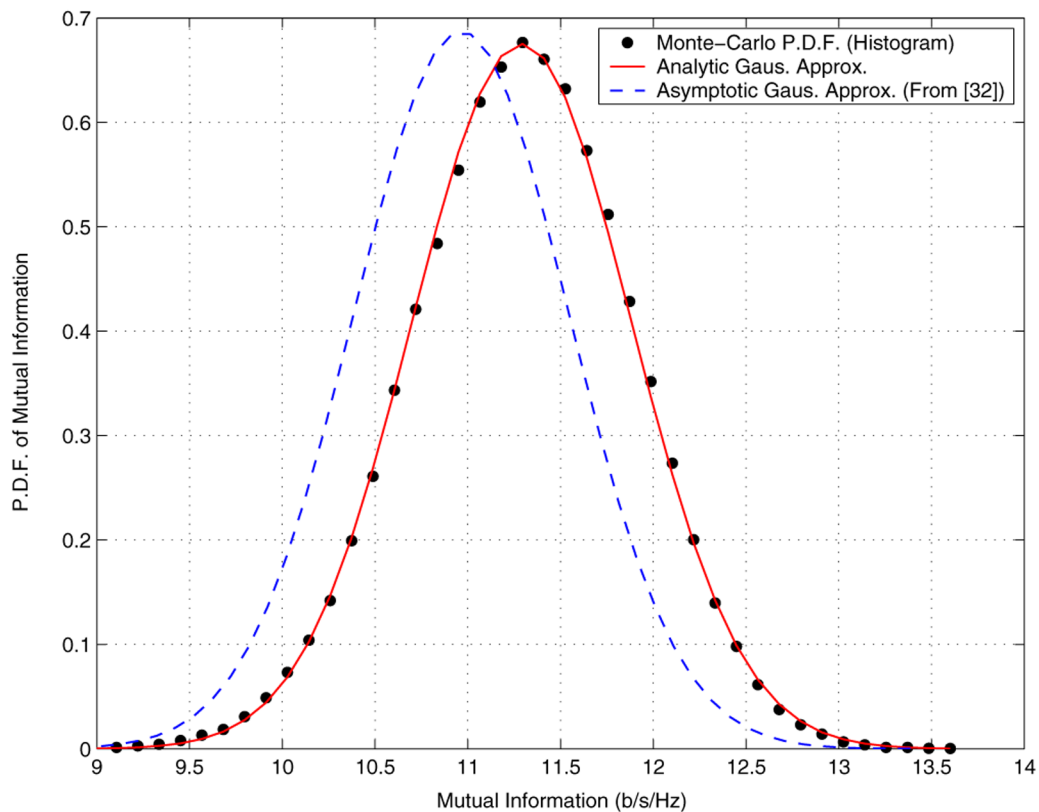


Fig. 7. The pdf of MIMO-OFDM mutual information. The “analytic Gaussian approximation” curve is based on the exact mean formula (48) and exact variance formula (9). The “asymptotic Gaussian approximation (from [32])” curve is based on [32, eqs. (59) and (60)]. The 2×2 antennas and 32 subcarriers are considered with 20-dB SNR. The channel follows an eight-path uniform power delay profile.

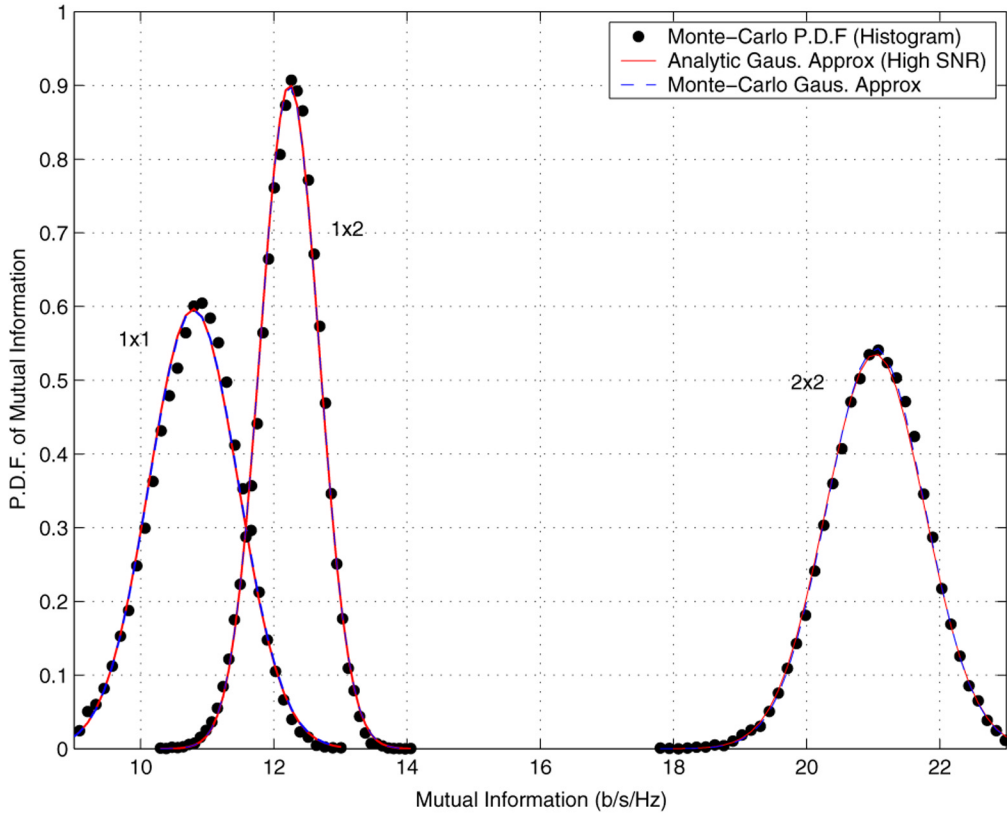


Fig. 8. The pdf of MIMO-OFDM mutual information at high SNR. The “analytic Gaussian approximation (high SNR)” curves are based on the high SNR mean formula [3, Th. 2], and the variance formula (24) for the 2×2 case, (31) for the 1×2 case, and (33) for the 1×1 case. Sixteen subcarriers are considered with SNR of 35 dB. The channel follows an eight-path exponential power delay profile with $K_{\text{exp}} = 4$.

all cases. Moreover, we see a significant reduction in the variance of the mutual information as the RMS delay spread increases (i.e., as K_{exp} increases). Again note that the Monte Carlo Gaussian approximation is indistinguishable from our new analytical Gaussian approximation curves.

Fig. 7 compares our new analytic Gaussian approximation with the asymptotic Gaussian approximation previously derived in [32], formally derived under the assumption of asymptotically large antenna numbers. To our knowledge, this is the only other comparable analytical result in the literature that applies for arbitrary-length frequency-selective MIMO channels. In the figure, we consider a 2×2 system at 20-dB SNR. The channel has a uniform power-delay profile, for which simple approximations to the mean and variance of the mutual information were explicitly presented in [32, eqs. (59) and (60)]. Clearly, although the approximation in [32] was shown to be quite accurate for some practical scenarios, Fig. 7 shows that our analytic Gaussian approximation is more accurate (although it is only shown for 32 subcarriers, the same observation has been made for all systems investigated).

Fig. 8 presents the distribution of the mutual information at high SNRs, comparing MIMO, SIMO, and SISO systems. The analytic Gaussian approximation curves are based on a high SNR mean formula from [3, Th. 2], and the high SNR variance formula (24) for the MIMO case, (31) for the SIMO case, and (33) for the SISO case. We see that the analytic Gaussian approximation is accurate in all cases. Again note that the Monte Carlo Gaussian approximation is indistinguishable from our new analytical Gaussian approximation curves.

Fig. 9 presents the distribution of the mutual information at low SNRs. The analytic Gaussian approximation curve is generated based

on the low SNR mean formula obtained by combining (131) and (133) and the low SNR variance formula (34). In this case we see that a Gaussian distribution no longer accurately predicts the mutual information pdf. This can be explained by examining (131), where we see that at low SNRs the mutual information for each subcarrier is a function of $\text{tr}(\mathbf{H}_k \mathbf{H}_k^\dagger)$, which for i.i.d. Rayleigh fading is $\sim \chi_{2N_r N_t}^2$. Hence, the overall mutual information (8) is distributed as the sum of N correlated $\chi_{2N_r N_t}^2$ random variables, which (for small N) is clearly quite different from Gaussian.

Motivated by this observation, we propose to approximate the mutual information pdf at low SNR with a Gamma distribution. Note that a Gamma approximation was previously considered in the context of flat-fading channels in [40]. The Gamma pdf is given by

$$f(x) = \frac{\theta^r x^{r-1} e^{-\theta x}}{\Gamma(r)}, \quad x \geq 0 \tag{41}$$

where r is the shape parameters and θ is the scale parameter. By matching the first two moments, a Gamma approximation for the mutual information pdf of MIMO-OFDM is obtained by evaluating

$$r = \frac{E[\mathcal{I}_{\text{ofdm}}]}{\text{Var}(\mathcal{I}_{\text{ofdm}})} \tag{42}$$

and

$$\theta = \frac{E^2[\mathcal{I}_{\text{ofdm}}]}{\text{Var}(\mathcal{I}_{\text{ofdm}})}. \tag{43}$$

This analytic Gamma approximation is plotted in Fig. 9, based on the same low SNR analytic mean and variance formulas as used

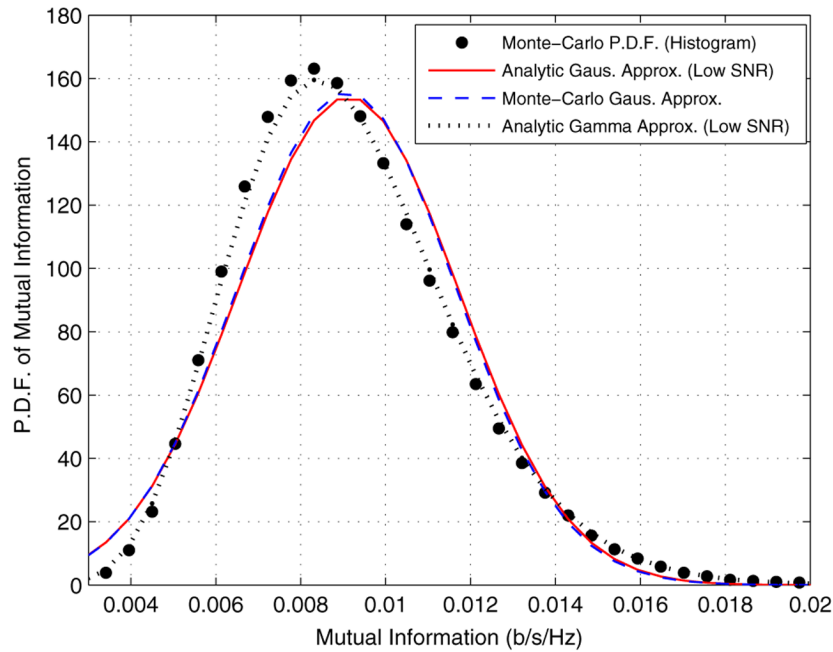


Fig. 9. The pdf of MIMO-OFDM mutual information at low SNR. The “analytic Gaussian approximation (low SNR)” and “analytic Gamma approximation (low SNR)” curves are based on the low SNR mean formula obtained by combining (131) and (133) and the low SNR variance formula (34). A 2×2 system with 32 subcarriers is considered with SNR of -25 dB. The channel follows a four-path exponential power delay profile with $K_{\text{exp}} = 2$.

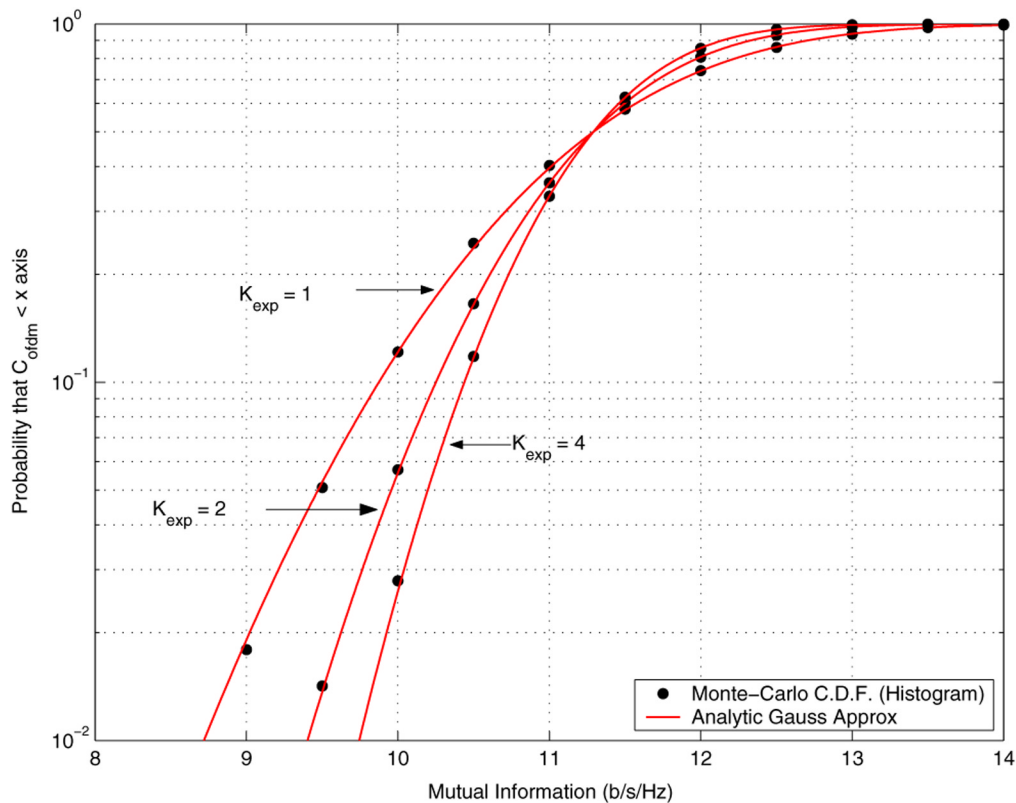


Fig. 10. The cdf of MIMO-OFDM mutual information for different RMS delay spreads (channels follow an eight-path exponential power delay profile with different K_{exp}). The “analytic Gaussian approximation” curves are based on the exact mean formula (48) and exact variance formula (9). A 2×2 system with 16 subcarriers is considered with SNR of 20 dB.

for the low SNR Gaussian approximation above. We clearly see that the Gamma approximation is much more accurate than the Gaussian approximation in this low SNR regime, and it follows the simulated pdf very closely.

B. Outage Capacity

The outage capacity $\mathcal{I}_{\text{out},q}$ is defined as the maximum information rate guaranteed to be supported for $100(1 - q)\%$ of the channel

realizations,³ i.e.,

$$P(\mathcal{I}_{\text{ofdm}} \leq \mathcal{I}_{\text{out},q}) = q \quad (44)$$

where q denotes the outage probability, and is thus directly obtained by inverting the cumulative distribution function (cdf) of $\mathcal{I}_{\text{ofdm}}$. If the distribution of the mutual information is Gaussian, then the outage capacity can be computed from the derived mean and variance as [33, eq. (26)]

$$\mathcal{I}_{\text{out},q} = E[\mathcal{I}_{\text{ofdm}}] - \sqrt{\text{Var}(\mathcal{I}_{\text{ofdm}})}Q^{-1}(q) \quad (45)$$

where $Q(\cdot)$ is the Gaussian Q -function.

Fig. 10 plots the outage probability for channels with different RMS delay spreads. The ‘‘analytic Gaussian approximation’’ curves are generated by approximating the cdf in (44) as a Gaussian distribution, and using the exact mean and variance formulas in (48) and (1) respectively. Clearly, this analytic Gaussian approximation matches closely with the empirically generated cdf (Monte Carlo histogram) in all cases. Moreover, we see that for outage probabilities of practical interest (e.g., $q = 1\%$), increasing the RMS delay spread can yield a significant improvement in outage capacity.

V. CONCLUSION

This communication has considered the mutual information distribution of frequency-selective MIMO channels, in the context of OFDM-based spatial multiplexing systems. Exact closed-form expressions were presented for the mutual information variance, applying for arbitrary finite system and channel parameters. These results were used to provide accurate analytical approximations for the distribution of mutual information, and the outage capacity. We observed that for most scenarios a Gaussian approximation is accurate, while also noting that for low SNR a Gamma approximation yielded even higher accuracy.

APPENDIX I PROOF OF THEOREM 1

Proof: By definition, the variance of the mutual information is given by

$$\text{Var}(\mathcal{I}_{\text{ofdm}}) = E[\mathcal{I}_{\text{ofdm}}^2] - E^2[\mathcal{I}_{\text{ofdm}}]. \quad (46)$$

Noting that $E[\mathcal{I}_{\text{ofdm}}] = E[\mathcal{I}_{\text{flat}}]$, and using (6), we have

$$\begin{aligned} \text{Var}(\mathcal{I}_{\text{ofdm}}) &= E\left[\frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{\ell=0}^{N-1} \mathcal{I}_k \mathcal{I}_\ell\right] - E^2[\mathcal{I}_{\text{flat}}] \\ &= \frac{1}{N^2} \left(\sum_{k=0}^{N-1} \sum_{\ell=0, \ell \neq k}^{N-1} E[\mathcal{I}_k \mathcal{I}_\ell] + \sum_{k=0}^{N-1} E[\mathcal{I}_k^2] \right) - E^2[\mathcal{I}_{\text{flat}}] \\ &= \frac{1}{N^2} \left(\sum_{k=0}^{N-1} \sum_{\ell=0, \ell \neq k}^{N-1} E[\mathcal{I}_k \mathcal{I}_\ell] \right) + \frac{1}{N} E[\mathcal{I}_{\text{flat}}^2] - E^2[\mathcal{I}_{\text{flat}}] \end{aligned} \quad (47)$$

³Strictly speaking, computing the outage capacity would require performing a numerical optimization over all possible input distributions, as discussed in [1]. Here, however, we adopt a common slight abuse of terminology and use the term outage capacity to denote the outage rate for the case of OFDM-based spatial multiplexing systems with equal power Gaussian inputs.

where $\mathcal{I}_{\text{flat}}$ denotes the mutual information of a flat-fading channel. Note that the last line followed by noting that, under the assumptions in Section II-A, the channel statistics for each subcarrier (and therefore, the mutual information statistics) are identical [19], and moreover, these statistics are equal to that of a flat-fading i.i.d. Rayleigh channel. The first and second moments of the mutual information for flat-fading channels has been previously derived in terms of incomplete gamma functions in [15, eqs. (29) and (31)]. Using [37, eq. (6.5.9)], we perform some basic manipulations to express these results in alternative simplified forms as follows:

$$E[\mathcal{I}_{\text{flat}}] = \frac{\log_2(e)}{\Gamma_m(n)\Gamma_m(m)} \sum_{r=1}^m \det(\mathbf{A}_r) \quad (48)$$

$$E[\mathcal{I}_{\text{flat}}^2] = \frac{(\log_2(e))^2}{\Gamma_m(n)\Gamma_m(m)} \sum_{r=1}^m \sum_{s=1}^m \det(\mathbf{B}_{r,s}) \quad (49)$$

where \mathbf{A}_r and $\mathbf{B}_{r,s}$ are defined in (12) and (13), respectively.

The challenge is to evaluate the cross correlation of the mutual information across frequency subcarriers $E[\mathcal{I}_k \mathcal{I}_\ell]$ which, using (7), is given by

$$\begin{aligned} E[\mathcal{I}_k \mathcal{I}_\ell] &= E\left[\log_2 \det\left(\mathbf{I}_{N_r} + \frac{\gamma}{N_t} \mathbf{H}_k \mathbf{H}_k^\dagger\right) \right. \\ &\quad \left. \times \log_2 \det\left(\mathbf{I}_{N_r} + \frac{\gamma}{N_t} \mathbf{H}_\ell \mathbf{H}_\ell^\dagger\right)\right] \\ &= E\left[\sum_{i=1}^m \log_2\left(1 + \frac{\gamma}{N_t} \lambda_i\right) \sum_{j=1}^m \log_2\left(1 + \frac{\gamma}{N_t} \omega_j\right)\right] \end{aligned} \quad (50)$$

where $\underline{\lambda} = \{\lambda_i\}_{i=1}^m$ and $\underline{\omega} = \{\omega_j\}_{j=1}^m$ are the nonzero eigenvalues of $\mathbf{H}_k \mathbf{H}_k^\dagger$ and $\mathbf{H}_\ell \mathbf{H}_\ell^\dagger$, respectively. Defining

$$\alpha(x) = \log_2\left(1 + \frac{\gamma}{N_t} x\right) \quad (51)$$

we have

$$\begin{aligned} E[\mathcal{I}_k \mathcal{I}_\ell] &= E\left[\sum_{i=1}^m \sum_{j=1}^m \alpha(\lambda_i) \alpha(\omega_j)\right] \\ &= \sum_{i=1}^m \sum_{j=1}^m E[\alpha(\lambda_i) \alpha(\omega_j)]. \end{aligned} \quad (52)$$

Now, to evaluate the expectations in (52), we first simplify the problem by exploiting the symmetry with respect to the λ_i 's and ω_j 's. To this end, let λ and ω be randomly (uniformly) chosen eigenvalues from $\underline{\lambda}$ and $\underline{\omega}$, respectively. Then, clearly

$$\Pr(\lambda = \lambda_i, \omega = \omega_j) = \frac{1}{m^2} \quad (53)$$

for any given $i \in \{1, \dots, m\}, j \in \{1, \dots, m\}$. Hence, we can also write

$$\begin{aligned} E[\alpha(\lambda) \alpha(\omega)] &= \sum_{i=1}^m \sum_{j=1}^m \Pr(\lambda = \lambda_i, \omega = \omega_j) \\ &\quad \times E[\alpha(\lambda) \alpha(\omega) | \lambda = \lambda_i, \omega = \omega_j] \\ &= \frac{1}{m^2} \sum_{i=1}^m \sum_{j=1}^m E[\alpha(\lambda_i) \alpha(\omega_j)] \end{aligned} \quad (54)$$

where the second line follows from (53). Therefore, by directly comparing (54) with (52), it follows that

$$E[\mathcal{I}_k \mathcal{I}_\ell] = m^2 E[\alpha(\lambda) \alpha(\omega)]. \quad (55)$$

We point out that the simplification from (52)–(55) is particularly important, since in order to evaluate the expectation in (55), clearly we

only require the distribution of a pair of arbitrarily selected eigenvalues λ and ω . This turns out to be much more convenient than dealing with the distributions of the individual pairs of ordered eigenvalues, i.e., λ_i and ω_j , required to directly evaluate (52).

The joint pdf of λ and ω is presented in Lemma 1 in Appendix II. From this lemma, we see that $f(\lambda, \omega)$, and correspondingly, $E[\mathcal{I}_k \mathcal{I}_\ell]$ in (55), only depends on k and ℓ through their absolute difference, i.e., since $f(\lambda, \omega)$ only depends on k and ℓ via $|\rho_{k-\ell}|$, and from (5)

$$|\rho_{k-\ell}| = |\rho_{\ell-k}^*| = |\rho_{\ell-k}|. \quad (56)$$

Therefore, the left-hand summation in (47) can be written as

$$\sum_{k=0}^{N-1} \sum_{\ell=0, \ell \neq k}^{N-1} E[\mathcal{I}_k \mathcal{I}_\ell] = 2 \sum_{d=1}^{N-1} (N-d) E[\mathcal{I}_0 \mathcal{I}_d]. \quad (57)$$

Note that for subcarrier spacings d for which the frequency matrices are independent (i.e., $\rho_d = 0$) or completely correlated (i.e., $\rho_d = 1$), the expectations in (57) are evaluated trivially as

$$\begin{aligned} E[\mathcal{I}_0 \mathcal{I}_d] &= E^2[\mathcal{I}_{\text{flat}}], & \rho_d &= 0 \\ E[\mathcal{I}_0 \mathcal{I}_d] &= E[\mathcal{I}_{\text{flat}}^2], & \rho_d &= 1. \end{aligned} \quad (58)$$

For the case $0 < |\rho_d| < 1$, such a direct evaluation is not possible, and we use (55) in Lemma 1 and (67) to evaluate the expectations in (57) as follows:

$$\begin{aligned} E[\mathcal{I}_0 \mathcal{I}_d] &= m^2 \int_0^\infty \int_0^\infty \frac{\alpha(\lambda) \alpha(\omega) |\rho_d|^{-m(n-1)}}{\Gamma_m(n) \Gamma_m(m) m^2 (1-|\rho_d|^2)^m} \\ &\quad \times \sum_{r=1}^m \sum_{s=1}^m \det(\mathbf{D}_{r,s}(\lambda, \omega)) d\lambda d\omega \\ &= \frac{|\rho_d|^{-m(n-1)}}{\Gamma_m(n) \Gamma_m(m) (1-|\rho_d|^2)^m} \sum_{r=1}^m \sum_{s=1}^m \det(\mathbf{D}_{r,s}) \end{aligned} \quad (59)$$

for $0 < |\rho_d| < 1$, where $\mathbf{D}_{r,s}$ is an $m \times m$ matrix with (i, j) th element defined in (60), shown at the bottom of the page.

Using the identity [11]

$$\int_0^\infty \ln(1 + \alpha\lambda) \lambda^{q-1} e^{-b\lambda} d\lambda = \Gamma(q) e^{b/\alpha} b^{-q} \sum_{h=1}^q E_h \left(\frac{b}{\alpha} \right) \quad (61)$$

we can evaluate $b(i, j)$ and $c(i, j)$ in closed form as

$$\begin{aligned} b(i, j) &= \frac{\log_2(e) e^{Nt/\gamma} \Gamma(\tau + j) |\rho_d|^\tau}{(1-|\rho_d|^2)^{-j}} \\ &\quad \times \sum_{t=0}^{j-1} \binom{j-1}{t} \left(\frac{|\rho_d|^2}{1-|\rho_d|^2} \right)^t \frac{\Gamma(u)}{(\tau+t)!} g_1(u) \end{aligned} \quad (62)$$

and

$$\begin{aligned} c(i, j) &= \frac{\log_2(e) e^{Nt/\gamma} \Gamma(\tau + i) |\rho_d|^\tau}{(1-|\rho_d|^2)^{-i}} \\ &\quad \times \sum_{t=0}^{i-1} \binom{i-1}{t} \left(\frac{|\rho_d|^2}{1-|\rho_d|^2} \right)^t \frac{\Gamma(v)}{(\tau+t)!} g_1(v) \end{aligned} \quad (63)$$

respectively. We evaluate the remaining integral $a(i, j)$ by using the power series expansion

$$I_\tau(x) = \sum_{k=0}^{\infty} \left(\frac{x}{2} \right)^{\tau+2k} \frac{1}{k!(\tau+k)!} \quad (64)$$

and integrating term by term using (61) to obtain

$$\begin{aligned} a(i, j) &= (\log_2(e))^2 e^{\frac{2Nt}{\gamma(1-|\rho_d|^2)}} |\rho_d|^\tau (1-|\rho_d|^2)^{\tau+i+j} \\ &\quad \times \sum_{t=0}^{\infty} \frac{|\rho_d|^{2t} \Gamma(u) \Gamma(v)}{t!(\tau+t)!} g_2(u) g_2(v). \end{aligned} \quad (65)$$

Substituting (62), (63), and (65) into (60), we perform some basic algebraic manipulations to write (59) as follows:

$$E[\mathcal{I}_0 \mathcal{I}_d] = \frac{(\log_2(e))^2 e^{2Nt/\gamma}}{\Gamma_m(n) \Gamma_m(m)} \sum_{r=1}^m \sum_{s=1}^m \det(\mathbf{C}_{r,s}(\rho_d)) \quad (66)$$

for $0 < |\rho_d| < 1$. The proof is completed by substituting (66) and (58) into (57), and then substituting (57), (49), and (48) into (47) and simplifying. \square

APPENDIX II

JOINT PDF OF ARBITRARILY SELECTED EIGENVALUES OF SUBCARRIER MATRICES

Lemma 1: Let λ and ω be arbitrarily selected nonzero eigenvalues of the subcarrier matrices $\mathbf{H}_k \mathbf{H}_k^\dagger$ and $\mathbf{H}_\ell \mathbf{H}_\ell^\dagger$, respectively. Then, the joint pdf of λ and ω is given by

$$\begin{aligned} f(\lambda, \omega) &= \frac{|\rho_d|^{-m(n-1)}}{\Gamma_m(n) \Gamma_m(m) m^2 (1-|\rho_d|^2)^m} \\ &\quad \times \sum_{r=1}^m \sum_{s=1}^m \det(\mathbf{D}_{r,s}(\lambda, \omega)) \end{aligned} \quad (67)$$

where $d = k - \ell$, $\tau = n - m$, and $\mathbf{D}_{r,s}(\lambda, \omega)$ is an $m \times m$ matrix with (i, j) th element given by (68), shown at the bottom of the next page, where $I_\tau(\cdot)$ is the modified Bessel function of the first kind [37, eq. (9.6.10)].

Proof: From (5), we see that $\mathbf{H}_k \mathbf{H}_k^\dagger$ and $\mathbf{H}_\ell \mathbf{H}_\ell^\dagger$ are (frequency) correlated Wishart matrices. In [41], the joint ordered eigenvalue density for matrices of this general form was evaluated for cases where the

$$\mathbf{D}_{r,s}(\lambda, \omega)_{i,j} = \begin{cases} a(i, j) \triangleq \int_0^\infty \int_0^\infty \lambda^{\frac{\tau}{2}+i-1} \omega^{\frac{\tau}{2}+j-1} e^{-\frac{\lambda}{1-|\rho_d|^2}} e^{-\frac{\omega}{1-|\rho_d|^2}} I_\tau \left(\frac{2|\rho_d|}{1-|\rho_d|^2} \sqrt{\lambda\omega} \right) \alpha(\lambda) \alpha(\omega) d\lambda d\omega, & \text{for } i = r, j = s \\ b(i, j) \triangleq \int_0^\infty \frac{\Gamma(\tau+j) |\rho_d|^\tau}{(1-|\rho_d|^2)^{-j}} e^{-\lambda} \lambda^{\tau+i-1} \sum_{t=0}^{j-1} \binom{j-1}{t} \left(\frac{|\rho_d|^2 \lambda}{1-|\rho_d|^2} \right)^t \frac{\alpha(\lambda)}{(\tau+t)!} d\lambda, & \text{for } i = r, j \neq s \\ c(i, j) \triangleq \int_0^\infty \frac{\Gamma(\tau+i) |\rho_d|^\tau}{(1-|\rho_d|^2)^{-i}} e^{-\omega} \omega^{\tau+j-1} \sum_{t=0}^{i-1} \binom{i-1}{t} \left(\frac{|\rho_d|^2 \omega}{1-|\rho_d|^2} \right)^t \frac{\alpha(\omega)}{(\tau+t)!} d\omega, & \text{for } i \neq r, j = s \\ \frac{\Gamma(\tau+j) |\rho_d|^\tau}{(1-|\rho_d|^2)^{-j}} \sum_{t=0}^{j-1} \binom{j-1}{t} \left(\frac{|\rho_d|^2}{1-|\rho_d|^2} \right)^t \frac{(\tau+i+t-1)!}{(\tau+t)!}, & \text{for } i \neq r, j \neq s \end{cases} \quad (68)$$

correlation coefficient was real. Extending this result to complex correlation coefficients, and to unordered eigenvalues, we obtain the joint eigenvalue density

$$f_u(\underline{\lambda}, \underline{\omega}) = \frac{1}{m!^2 \Gamma_m(n) \Gamma_m(m) (1 - |\rho_d|^2)^m} |\rho_d|^{-m(n-1)} \times \exp\left(-\frac{\sum_{t=1}^m (\lambda_t + \omega_t)}{1 - |\rho_d|^2}\right) \Delta_m(\underline{\lambda}) \Delta_m(\underline{\omega}) \times \det\left((\lambda_i \omega_j)^{\frac{\alpha_i}{2}} I_\tau \left(\frac{2|\rho_d|}{1 - |\rho_d|^2} \sqrt{\lambda_i \omega_j}\right)\right) \quad (69)$$

where $\Delta_m(\cdot)$ is a Vandermonde determinant, defined as

$$\Delta_m(\underline{\lambda}) = \prod_{i < j} (\lambda_j - \lambda_i) = \det(\lambda_i^{j-1}). \quad (70)$$

Note that the extension from ordered to unordered eigenvalues simply involved the addition of the leading $1/m!^2$ factor in (69), whereas the extension from real to complex correlation coefficients is trivial, and the proof is omitted.

To evaluate (67), we marginalize (69) as follows:

$$f(\lambda, \omega) = \int_0^\infty \cdots \int_0^\infty f_u(\underline{\lambda}, \underline{\omega}) d\underline{\lambda}_1 d\underline{\omega}_1 \quad (71)$$

where $d\underline{\lambda}_1 = d\lambda_2 \cdots d\lambda_m$, $d\underline{\omega}_1 = d\omega_2 \cdots d\omega_m$, and we have let $\lambda_1 = \lambda$ and $\omega_1 = \omega$. We evaluate these integrals by first expanding the Vandermonde determinants in (69) according to

$$\Delta_m(\underline{\lambda}) \Delta_m(\underline{\omega}) = \sum_{\alpha} (-1)^{\text{per}(\alpha)} \prod_{i=1}^m \lambda_i^{\alpha_i - 1} \times \sum_{\beta} (-1)^{\text{per}(\beta)} \prod_{j=1}^m \omega_j^{\beta_j - 1} \quad (72)$$

where the sums are over all permutations $\alpha = \{\alpha_1, \dots, \alpha_m\}$ and $\beta = \{\beta_1, \dots, \beta_m\}$ of $\{1, \dots, m\}$, and $(-1)^{\text{per}(\alpha)}$ and $(-1)^{\text{per}(\beta)}$ denote the signs of the permutations. Substituting (72) and (69) into (71) yields

$$f(\lambda, \omega) = \int_0^\infty \cdots \int_0^\infty \frac{|\rho_d|^{-m(n-1)}}{\Gamma_m(n) \Gamma_m(m) (m!)^2 (1 - |\rho_d|^2)^m} \times \exp\left(-\frac{\sum_{t=1}^m (\lambda_t + \omega_t)}{1 - |\rho_d|^2}\right) \sum_{\alpha} (-1)^{\text{per}(\alpha)}$$

$$\times \prod_{i=1}^m \lambda_i^{\alpha_i - 1} \sum_{\beta} (-1)^{\text{per}(\beta)} \prod_{j=1}^m \omega_j^{\beta_j - 1} \times \det\left((\lambda_i \omega_j)^{\frac{\alpha_i}{2}} I_\tau \left(\frac{2|\rho_d|}{1 - |\rho_d|^2} \sqrt{\lambda_i \omega_j}\right)\right) d\underline{\lambda}_1 d\underline{\omega}_1 = \frac{|\rho_d|^{-m(n-1)}}{\Gamma_m(n) \Gamma_m(m) (m!)^2 (1 - |\rho_d|^2)^m} \sum_{\alpha} \sum_{\beta} (-1)^{\text{per}(\alpha) + \text{per}(\beta)} \times \int_0^\infty \cdots \int_0^\infty \det(a(\lambda_i, \omega_j, \alpha_i, \beta_j)) d\underline{\lambda}_1 d\underline{\omega}_1 \quad (73)$$

where

$$a(\lambda_i, \omega_j, \alpha_i, \beta_j) = \lambda_i^{\frac{\alpha_i}{2} + \alpha_i - 1} \omega_j^{\frac{\beta_j}{2} + \beta_j - 1} e^{-\frac{\lambda_i}{1 - |\rho_d|^2}} \times e^{-\frac{\omega_j}{1 - |\rho_d|^2}} I_\tau \left(\frac{2|\rho_d|}{1 - |\rho_d|^2} \sqrt{\lambda_i \omega_j}\right). \quad (74)$$

Expanding the determinants, integrating term by term, and reforming determinants, we obtain

$$f(\lambda, \omega) = \frac{|\rho_d|^{-m(n-1)}}{\Gamma_m(n) \Gamma_m(m) (m!)^2 (1 - |\rho_d|^2)^m} \times \sum_{\alpha} \sum_{\beta} (-1)^{\text{per}(\alpha) + \text{per}(\beta)} \det(\tilde{\mathbf{D}}_{\alpha, \beta}(\lambda, \omega)) \quad (75)$$

where $\tilde{\mathbf{D}}_{\alpha, \beta}(\cdot)$ is an $m \times m$ matrix with (i, j) th element given by (76), shown at the bottom of the page. Reordering rows and columns yields

$$\det(\tilde{\mathbf{D}}_{\alpha, \beta}(\lambda, \omega)) = (-1)^{\text{per}(\alpha) + \text{per}(\beta)} \det(\mathbf{D}_{\alpha_1, \beta_1}(\lambda, \omega)) \quad (77)$$

where

$$(\mathbf{D}_{\alpha_1, \beta_1}(\lambda, \omega))_{i, j} = \begin{cases} a(\lambda, \omega, i, j), & \text{for } i = \alpha_1, j = \beta_1 \\ b(\lambda, i, j), & \text{for } i = \alpha_1, j \neq \beta_1 \\ c(\omega, i, j), & \text{for } i \neq \alpha_1, j = \beta_1 \\ d(i, j), & \text{for } i \neq \alpha_1, j \neq \beta_1. \end{cases} \quad (78)$$

Applying (77) in (75), we can further simplify as follows:

$$f(\lambda, \omega) = \frac{|\rho_d|^{-m(n-1)}}{\Gamma_m(n) \Gamma_m(m) (m!)^2 (1 - |\rho_d|^2)^m} \times \sum_{\alpha} \sum_{\beta} \det(\mathbf{D}_{\alpha_1, \beta_1}(\lambda, \omega)) = \frac{|\rho_d|^{-m(n-1)} ((m-1)!)^2}{\Gamma_m(n) \Gamma_m(m) (m!)^2 (1 - |\rho_d|^2)^m}$$

$$(\mathbf{D}_{r, s}(\lambda, \omega))_{i, j} = \begin{cases} \lambda^{\frac{\alpha_i}{2} + i - 1} \omega^{\frac{\beta_j}{2} + j - 1} e^{-\frac{\lambda}{1 - |\rho_d|^2}} e^{-\frac{\omega}{1 - |\rho_d|^2}} I_\tau \left(\frac{2|\rho_d|}{1 - |\rho_d|^2} \sqrt{\lambda \omega}\right), & \text{for } i = r, j = s \\ \frac{\Gamma(\tau + j) |\rho_d|^\tau}{(1 - |\rho_d|^2)^{-j}} e^{-\lambda} \lambda^{\tau + i - 1} \sum_{t=0}^{j-1} \binom{j-1}{t} \left(\frac{|\rho_d|^2 \lambda}{1 - |\rho_d|^2}\right)^t \frac{1}{(\tau + t)!}, & \text{for } i = r, j \neq s \\ \frac{\Gamma(\tau + i) |\rho_d|^\tau}{(1 - |\rho_d|^2)^{-i}} e^{-\omega} \omega^{\tau + j - 1} \sum_{t=0}^{i-1} \binom{i-1}{t} \left(\frac{|\rho_d|^2 \omega}{1 - |\rho_d|^2}\right)^t \frac{1}{(\tau + t)!}, & \text{for } i \neq r, j = s \\ \frac{\Gamma(\tau + j) |\rho_d|^\tau}{(1 - |\rho_d|^2)^{-j}} \sum_{t=0}^{j-1} \binom{j-1}{t} \left(\frac{|\rho_d|^2}{1 - |\rho_d|^2}\right)^t \frac{(\tau + i + t - 1)!}{(\tau + t)!}, & \text{for } i \neq r, j \neq s \end{cases} \quad (68)$$

$$(\tilde{\mathbf{D}}_{\alpha, \beta}(\lambda, \omega))_{i, j} = \begin{cases} a(\lambda, \omega, \alpha_i, \beta_j), & \text{for } i = 1, j = 1 \\ b(\lambda, \alpha_i, \beta_j) \triangleq \int_0^\infty a(\lambda, \omega_j, \alpha_i, \beta_j) d\omega_j, & \text{for } i = 1, j \neq 1 \\ c(\omega, \alpha_i, \beta_j) \triangleq \int_0^\infty a(\lambda_i, \omega, \alpha_i, \beta_j) d\lambda_i, & \text{for } i \neq 1, j = 1 \\ d(\alpha_i, \beta_j) \triangleq \int_0^\infty \int_0^\infty a(\lambda_i, \omega_j, \alpha_i, \beta_j) d\lambda_i d\omega_j, & \text{for } i \neq 1, j \neq 1 \end{cases} \quad (76)$$

$$\begin{aligned}
& \times \sum_{\alpha_1=1}^m \sum_{\beta_1=1}^m \det(\mathbf{D}_{\alpha_1, \beta_1}(\lambda, \omega)) \\
& = \frac{|\rho_d|^{-m(n-1)}}{\Gamma_m(n)\Gamma_m(m)m^2(1-|\rho_d|^2)^m} \\
& \times \sum_{r=1}^m \sum_{s=1}^m \det(\mathbf{D}_{r,s}(\lambda, \omega)). \quad (79)
\end{aligned}$$

The result now follows by combining (74), (78), and (79), and by evaluating the integrals $b(\cdot)$, $c(\cdot)$, and $d(\cdot)$ inside the remaining determinant, using the identities [41]

$$\begin{aligned}
& \int_0^\infty x^{a+\frac{t}{2}-1} e^{-cx} I_t(2\sqrt{fx}) dx \\
& = \frac{(t+a-1)!}{c^{\frac{t}{2}+a}} \left(\frac{f}{c}\right)^{\frac{t}{2}} e^{\frac{f}{c}} \sum_{r=0}^{a-1} \binom{a-1}{r} \frac{\left(\frac{f}{c}\right)^r}{(t+r)!} \quad (80)
\end{aligned}$$

for integers a and t , and [36]

$$\int_0^\infty x^t e^{-ax} dx = \Gamma(t+1)a^{-(t+1)} \quad (81)$$

for integer $t \geq 0$.

APPENDIX III PROOF OF THEOREM 2

Proof: We start by noting that at high SNR, (7) approaches

$$\mathcal{I}_k = \log_2 \det \left(\frac{\gamma}{N_t} \mathbf{W}_k \right) \quad (82)$$

where it is an $m \times m$ complex Wishart matrix given by

$$\mathbf{W}_k = \begin{cases} \mathbf{H}_k \mathbf{H}_k^H, & \text{for } N_r \leq N_t \\ \mathbf{H}_k^H \mathbf{H}_k, & \text{for } N_r > N_t. \end{cases} \quad (83)$$

Substituting (82) into (47) and using (57), we write the variance of the MIMO-OFDM mutual information at high SNR as follows:

$$\begin{aligned}
& \text{Var}^\infty(\mathcal{I}_{\text{ofdm}}) \\
& = \left(\frac{2}{N^2} \sum_{d=1}^{N-1} (N-d) E \left[\log_2 \det \left(\frac{\gamma}{N_t} \mathbf{W}_0 \right) \log_2 \det \left(\frac{\gamma}{N_t} \mathbf{W}_d \right) \right] \right. \\
& \left. + \frac{1}{N} E \left[\left(\log_2 \det \left(\frac{\gamma}{N_t} \mathbf{W}_0 \right) \right)^2 \right] - E^2 \left[\log_2 \det \left(\frac{\gamma}{N_t} \mathbf{W}_0 \right) \right] \right). \quad (84)
\end{aligned}$$

Noting that

$$\log_2 \det \left(\frac{\gamma}{N_t} \mathbf{W}_0 \right) = m \log_2 \left(\frac{\gamma}{N_t} \right) + \log_2 \det(\mathbf{W}_0) \quad (85)$$

we apply some simple algebra to (84) and find that the terms involving γ cancel perfectly, leaving

$$\begin{aligned}
& \text{Var}^\infty(\mathcal{I}_{\text{ofdm}}) = \left(\frac{2}{N^2} \sum_{d=1}^{N-1} (N-d) E [\log_2 \det(\mathbf{W}_0) \log_2 \det(\mathbf{W}_d)] \right. \\
& \left. + \frac{1}{N} E [(\log_2 \det(\mathbf{W}_0))^2] - E^2 [\log_2 \det(\mathbf{W}_0)] \right). \quad (86)
\end{aligned}$$

Since \mathbf{W}_0 is a complex Wishart matrix, we invoke results from [3] to give

$$E[\log_2 \det(\mathbf{W}_0)] = \log_2(e) \sum_{t=0}^{m-1} \psi(n-t) \quad (87)$$

$$\begin{aligned}
& E[(\log_2 \det(\mathbf{W}_0))^2] \\
& = (\log_2(e))^2 \left(\sum_{t=0}^{m-1} \psi'(n-t) + \sum_{t=0}^{m-1} \psi(n-t) \right). \quad (88)
\end{aligned}$$

We now consider the remaining expectation $E[\log_2 \det(\mathbf{W}_0) \log_2 \det(\mathbf{W}_d)]$ in (86). For the extreme cases of $\rho_d = 0$ and $\rho_d = 1$, this is directly obtained from (87) and (88), respectively. The main challenge is to obtain a closed-form finite-sum expression for $0 < |\rho_d| < 1$.

We start by following the same procedure as used in (50)–(60) in the proof of Theorem 1, which yields

$$\begin{aligned}
& E[\log_2 \det(\mathbf{W}_0) \log_2 \det(\mathbf{W}_d)] \\
& = \frac{|\rho_d|^{-m(n-1)}}{\Gamma_m(n)\Gamma_m(m)(1-|\rho_d|^2)^m} \sum_{r=1}^m \sum_{s=1}^m \det(\bar{\mathbf{D}}_{r,s}) \quad (89)
\end{aligned}$$

for $0 < |\rho_d| < 1$, where $\bar{\mathbf{D}}_{r,s}$ is an $m \times m$ matrix with entries corresponding to (60), but with the $\alpha(\cdot)$ functions replaced with

$$\tilde{\alpha}(x) = \log_2(x). \quad (90)$$

□ We now evaluate the integrals for the elements of $\bar{\mathbf{D}}_{r,s}$ corresponding to $b(i, j)$ and $c(i, j)$ in (60), using the identity [36, eq. (4.352.1)]

$$\begin{aligned}
& \int_0^\infty x^{q-1} e^{-bx} \ln(x) dx \\
& = \frac{\Gamma(q)}{b^q} (\psi(q) - \ln(b)), \quad q > 0, \quad b > 0. \quad (91)
\end{aligned}$$

This gives

$$\begin{aligned}
& b(i, j) = \frac{\log_2(e) \Gamma(\tau+j) |\rho_d|^\tau}{(1-|\rho_d|^2)^{-j}} \\
& \times \sum_{t=0}^{j-1} \binom{j-1}{t} \left(\frac{|\rho_d|^2}{1-|\rho_d|^2} \right)^t \frac{\Gamma(u) \psi(u)}{(\tau+t)!} \quad (92)
\end{aligned}$$

and

$$\begin{aligned}
& c(i, j) = \frac{\log_2(e) \Gamma(\tau+i) |\rho_d|^\tau}{(1-|\rho_d|^2)^{-i}} \\
& \times \sum_{t=0}^{i-1} \binom{i-1}{t} \left(\frac{|\rho_d|^2}{1-|\rho_d|^2} \right)^t \frac{\Gamma(v) \psi(v)}{(\tau+t)!}. \quad (93)
\end{aligned}$$

To evaluate the remaining integrals in $\bar{\mathbf{D}}_{r,s}$, i.e., for the elements $a(i, j)$, we use (64) and (91) to obtain

$$a(i, j) = (\log_2(e))^2 |\rho_d|^\tau (1-|\rho_d|^2)^{\tau+i+j} \quad (94)$$

$$\begin{aligned}
& \times \sum_{t=0}^{\infty} \frac{|\rho_d|^{2t} \Gamma(u) \Gamma(v)}{t!(\tau+t)!} (H(u-1) + h(\rho_d)) \\
& \times (H(v-1) + h(\rho_d)). \quad (95)
\end{aligned}$$

Next we use (92)–(94) in (89), and perform some basic simplifications to obtain

$$\begin{aligned}
& E[\log_2 \det(\mathbf{W}_0) \log_2 \det(\mathbf{W}_d)] \\
& = \frac{(\log_2(e))^2}{\Gamma_m(n)\Gamma_m(m)} \sum_{r=1}^m \sum_{s=1}^m \det(\tilde{\mathbf{C}}_{r,s}(\rho_d)) \quad (96)
\end{aligned}$$

where $\tilde{\mathbf{C}}_{r,s}(\rho_d)$ is an $m \times m$ matrix with (i, j) th element given by (97), shown at the bottom of the next page. Expression (24) follows by using (96), (88), and (87) in (86).

To complete the proof, we must express the infinite summation in (97) in the simplified finite-sum form of (27). This simplification requires significant algebraic manipulations, which we now detail. Start by recalling the definitions $u = t + \tau + i$ and $v = t + \tau + j$, and writing the infinite sum in (97) as follows:

$$\left(\check{C}_{r,s}(\rho_d)\right)_{i,j} = \frac{(1 - |\rho_d|^2)^z}{|\rho_d|^{2(j-1)}} \mathcal{S}(|\rho_d|^2) \quad (98)$$

where

$$\begin{aligned} \mathcal{S}(x) &\triangleq \sum_{t=0}^{\infty} \frac{x^t (\tau + t + i - 1)! (\tau + t + i - 1)!}{t! (\tau + t)!} \\ &\times (h(\sqrt{x}) + H(\tau + t + i - 1)) \\ &\times (h(\sqrt{x}) + H(\tau + t + j - 1)). \end{aligned} \quad (99)$$

Note that the series (99), and those that follow below, are convergent for $|x| < 1$ [a condition which holds in (98)].

Now, (99) can be written as

$$\begin{aligned} \mathcal{S}(x) &= h^2(\sqrt{x}) \mathcal{S}_1(1, 1, x) + h(\sqrt{x}) (\mathcal{S}_1(H(i), 1, x) \\ &+ \mathcal{S}_1(1, H(j), x) + \mathcal{S}_1(H(i), H(j), x)) \end{aligned} \quad (100)$$

where

$$\begin{aligned} \mathcal{S}_1(f_1(i), f_2(j), x) &\triangleq \sum_{t=0}^{\infty} \frac{x^t (\tau + t + i - 1)! (\tau + t + j - 1)!}{t! (\tau + t)!} \\ &\times f_1(\tau + t + i - 1) f_2(\tau + t + j - 1) \end{aligned} \quad (101)$$

for arbitrary functions f_1 and f_2 . We now consider each of the infinite sums in (100) in turn.

First consider $\mathcal{S}_1(1, 1, x)$. Following a similar general approach to that used in [42], we perform the following sequence of operations⁴

$$\begin{aligned} \mathcal{S}_1(1, 1, x) &= \sum_{t=0}^{\infty} \frac{x^t (\tau + t + i - 1)! (\tau + t + j - 1)!}{t! (\tau + t)!} \\ &= \frac{d^{\tau+i-1}}{dx^{\tau+i-1}} \sum_{t=0}^{\infty} \frac{x^{t+\tau+i-1} (\tau + t + j - 1)!}{(\tau + t)!} \\ &= \frac{d^{\tau+i-1}}{dx^{\tau+i-1}} \sum_{t=\tau}^{\infty} \frac{x^{t+i-1} (t + j - 1)!}{t!} \\ &= \frac{d^{\tau+i-1}}{dx^{\tau+i-1}} \sum_{t=0}^{\infty} \frac{x^{t+i-1} (t + j - 1)!}{t!} \\ &= \frac{d^{\tau+i-1}}{dx^{\tau+i-1}} x^{i-1} \sum_{t=0}^{\infty} \frac{x^t (t + j - 1)!}{t!} \end{aligned}$$

⁴Note that for this particular case, a finite expression could be also found by directly matching the infinite series to a hypergeometric function, and using associated identities; something which cannot be done in the other cases.

$$\begin{aligned} &= \frac{d^{\tau+i-1}}{dx^{\tau+i-1}} \left(x^{i-1} \frac{d^{j-1}}{dx^{j-1}} \sum_{t=0}^{\infty} x^{t+j-1} \right) \\ &= \frac{d^{\tau+i-1}}{dx^{\tau+i-1}} \left(x^{i-1} \frac{d^{j-1}}{dx^{j-1}} \sum_{t=0}^{\infty} x^t \right). \end{aligned} \quad (102)$$

Via application of the Leibnitz formula, it can be shown that

$$\begin{aligned} \mathcal{S}_1(1, 1, x) &= \sum_{b=0}^{i-1} \binom{\tau + i - 1}{\tau + b} \frac{(i-1)!}{b!} x^b \frac{d^{\tau+j+b-1}}{dx^{\tau+j+b-1}} \left(\sum_{t=0}^{\infty} x^t \right) \\ &= \Gamma(\tau + i) \sum_{b=0}^{i-1} \binom{i-1}{b} \frac{x^b}{(\tau + b)!} \frac{d^{\tau+j+b-1}}{dx^{\tau+j+b-1}} \left(\sum_{t=0}^{\infty} x^t \right). \end{aligned} \quad (103)$$

Now noting that

$$\sum_{t=0}^{\infty} x^t = \frac{1}{1-x}, \quad |x| < 1 \quad (104)$$

with derivatives

$$\frac{d^r}{dx^r} \left(\sum_{t=0}^{\infty} x^t \right) = \frac{r!}{(1-x)^{r+1}} \quad (105)$$

we can write (103) as follows:

$$\begin{aligned} \mathcal{S}_1(1, 1, x) &= \frac{\Gamma(\tau + i)}{(1-x)^{\tau+j}} \sum_{b=0}^{i-1} \binom{i-1}{b} \left(\frac{x}{1-x} \right)^b \frac{(\tau + j + b - 1)!}{(\tau + b)!} \\ &= \frac{\Gamma(\tau + i) x^{i-1}}{(1-x)^z} \sum_{b=0}^{i-1} \binom{i-1}{b} \left(\frac{1-x}{x} \right)^b (\tau + i - b)_{j-1} \\ &= \frac{x^{i-1} \eta_{j,i}(1, \sqrt{x})}{(1-x)^z}. \end{aligned} \quad (106)$$

Now consider $\mathcal{S}_1(H(i), 1, x)$. Following the same sequence of operations as in (102) and (103), we find that

$$\begin{aligned} \mathcal{S}_1(H(i), 1, x) &= \frac{d^{\tau+j-1}}{dx^{\tau+j-1}} \left(x^{j-1} \frac{d^{i-1}}{dx^{i-1}} \sum_{t=0}^{\infty} x^t H(t) \right) \\ &= \Gamma(\tau + j) \sum_{b=0}^{j-1} \binom{j-1}{b} \\ &\times \frac{x^b}{(\tau + b)!} \frac{d^{\tau+i+b-1}}{dx^{\tau+i+b-1}} \left(\sum_{t=0}^{\infty} x^t H(t) \right). \end{aligned} \quad (107)$$

Now we use [43]

$$\sum_{t=0}^{\infty} x^t H(t) = \frac{-\ln(1-x)}{1-x}, \quad |x| < 1 \quad (108)$$

$$\left(\check{C}_{r,s}(\rho_d)\right)_{i,j} = \begin{cases} \eta_{i,j}(1, \rho_d), & \text{for } i \neq r, j \neq s \\ \eta_{i,j}(\psi(z), \rho_d), & \text{for } i = r, j \neq s \\ |\rho_d|^{2(i-j)} \eta_{j,i}(\psi(z), \rho_d), & \text{for } i \neq r, j = s \\ \frac{(1-|\rho_d|^2)^z}{|\rho_d|^{2(j-1)}} \sum_{t=0}^{\infty} \frac{|\rho_d|^{2t} \Gamma(u) \Gamma(v) (h(\rho_d) + H(u-1)) (h(\rho_d) + H(v-1))}{t! (\tau + t)!}, & \text{for } i = r, j = s \end{cases} \quad (97)$$

and the corresponding derivatives

$$\frac{d^r}{dx^r} \left(\sum_{t=0}^{\infty} x^t H(t) \right) = \frac{r!}{(1-x)^{r+1}} (H(r) - \ln(1-x)) \quad (109)$$

to write (107) as follows:

$$\begin{aligned} \mathcal{S}_1(H(i), 1, x) &= \frac{\Gamma(\tau+j)}{(1-x)^{\tau+i}} \sum_{b=0}^{j-1} \binom{j-1}{b} \left(\frac{x}{1-x}\right)^b \frac{(\tau+i+b-1)!}{(\tau+b)!} \\ &\quad \times (H(\tau+i+b-1) - \ln(1-x)) \\ &= \frac{x^{j-1} \Gamma(\tau+j)}{(1-x)^z} \sum_{b=0}^{j-1} \binom{j-1}{b} \left(\frac{1-x}{x}\right)^b (\tau+j-b)_{i-1} \\ &\quad \times (H(z-1-b) - \ln(1-x)) \\ &= \frac{x^{j-1}}{(1-x)^z} (\eta_{i,j}(H(z-1), \sqrt{x}) - \ln(1-x) \eta_{i,j}(1, \sqrt{x})). \quad (110) \end{aligned}$$

Now consider $\mathcal{S}_1(1, H(j), x)$. Using exactly the same approach as for $\mathcal{S}_1(H(i), 1, x)$, we obtain

$$\mathcal{S}_1(1, H(j), x) = \frac{x^{i-1}}{(1-x)^z} (\eta_{j,i}(H(z-1), \sqrt{x}) - \ln(1-x) \eta_{j,i}(1, \sqrt{x})). \quad (111)$$

Finally, consider $\mathcal{S}_1(H(i), H(j), x)$. We follow the same sequence of operations as in (102) and (103). In this case, it is convenient to take the successive derivatives based on the order of i and j . In particular, with $i' = \max(i, j)$ and $j' = \min(i, j)$, we obtain

$$\begin{aligned} \mathcal{S}_1(H(i), H(j), x) &= \frac{d^{\tau+i'-1}}{dx^{\tau+i'-1}} \left(x^{i'-1} \frac{d^{j'-1}}{dx^{j'-1}} \mathcal{S}_2(x) \right) \\ &= \Gamma(\tau+i') \sum_{b=0}^{i'-1} \binom{i'-1}{b} \frac{x^b}{(\tau+b)!} \frac{d^{\tau+j'+b-1}}{dx^{\tau+j'+b-1}} \mathcal{S}_2(x) \quad (112) \end{aligned}$$

where

$$\mathcal{S}_2(x) \triangleq \sum_{t=0}^{\infty} x^t H(t) H(t+i'-j'). \quad (113)$$

In this case, in contrast to the previous summations in (104) and (108), the infinite summation in (113) cannot be directly expressed in a finite form. To evaluate this series in finite form, we start by using (29) to write

$$\begin{aligned} \mathcal{S}_2(x) &= \sum_{t=1}^{\infty} x^t H(t) \left(H(t) + \sum_{q=1}^{i'-j'} \frac{1}{t+q} \right) \\ &= \sum_{t=1}^{\infty} x^t H(t)^2 + \mathcal{S}_3(x) \\ &= \frac{L_{i_2}(1-x) + \ln^2(1-x)}{1-x} + \mathcal{S}_3(x) \quad (114) \end{aligned}$$

where $L_{i_2}(\cdot)$ is the *dilogarithm* function [37, eq. (27.7.1)], and $\mathcal{S}_3(\cdot)$ is given by

$$\mathcal{S}_3(x) \triangleq \sum_{q=1}^{i'-j'} \sum_{t=1}^{\infty} \frac{x^t H(t)}{t+q}. \quad (115)$$

Note that the last line in (114) followed by using an identity from [43]. We now manipulate $\mathcal{S}_3(\cdot)$ as follows:

$$\begin{aligned} \mathcal{S}_3(x) &= \sum_{q=1}^{i'-j'} \frac{1}{x^q} \sum_{t=1}^{\infty} \frac{x^{t+q} H(t)}{t+q} \quad (116) \\ &= \sum_{q=1}^{i'-j'} \frac{1}{x^q} \sum_{t=1}^{\infty} \int x^{t+q-1} H(t) dx \\ &= \sum_{q=1}^{i'-j'} \frac{1}{x^q} \int x^{q-1} \left(\sum_{t=1}^{\infty} x^t H(t) \right) dx \\ &= \sum_{q=1}^{i'-j'} -\frac{1}{x^q} \int \frac{x^{q-1} \ln(1-x)}{1-x} dx. \quad (117) \end{aligned}$$

For $q > 1$, consider

$$\begin{aligned} \frac{x^{q-1}}{1-x} &= -x^{q-2} + \frac{x^{q-2}}{1-x} = \dots \\ &= \frac{1}{1-x} - \sum_{v=1}^{q-1} x^{v-1}, \quad q > 1 \quad (118) \end{aligned}$$

so, therefore

$$\begin{aligned} \mathcal{S}_3(x) &= \sum_{q=2}^{i'-j'} \frac{1}{x^q} \sum_{v=1}^{q-1} \int x^{v-1} \ln(1-x) dx \\ &\quad - \sum_{q=1}^{i'-j'} \frac{1}{x^q} \int \frac{\ln(1-x)}{1-x} dx. \quad (119) \end{aligned}$$

Using [36, eq. (2.729)]⁵

$$\begin{aligned} \int y^m \ln(1-y) dy &= \frac{1}{m+1} \left((y^{m+1} - 1) \ln(1-y) - \sum_{k=1}^{m+1} \frac{y^{m-k+2}}{m-k+2} \right) + \text{const} \quad (120) \end{aligned}$$

and noting that

$$\begin{aligned} \int \frac{\ln(1-x)}{1-x} dx &= - \int \ln(1-x) \frac{d}{dx} \ln(1-x) dx \\ &= -\frac{\ln^2(1-x)}{2} + \text{const} \quad (121) \end{aligned}$$

we can now express $\mathcal{S}_3(x)$ in finite form as follows:

$$\begin{aligned} \mathcal{S}_3(x) &= \sum_{q=2}^{i'-j'} \frac{1}{x^q} \sum_{v=1}^{q-1} \frac{1}{v} \left((x^v - 1) \ln(1-x) - \sum_{t=1}^v \frac{x^t}{t} \right) \\ &\quad + \frac{\ln^2(1-x)}{2} \sum_{q=1}^{i'-j'} \frac{1}{x^q}. \quad (122) \end{aligned}$$

Note that it can be easily verified, using (116), that the integration constant generated in going from (119) to (122) is zero. After much algebraic manipulation, it can be shown that (122) reduces to

$$\begin{aligned} \mathcal{S}_3(x) &= \frac{\ln(1-x)}{2} \sum_{q=1}^{i'-j'} \frac{\ln(1-x)}{x^q} \\ &\quad + \sum_{q=1}^{i'-j'-1} \left(\frac{\ln(1-x) H(i'-j'-q)}{x^q} - \frac{\ln(1-x) H(q)}{x^{q+1}} \right. \\ &\quad \left. - \frac{1}{x^q} \sum_{r=1}^{i'-j'-q} \frac{H(r+q-1) - H(r-1)}{r} \right). \quad (123) \end{aligned}$$

⁵There is a missing (-1) factor in this reference.

Now substituting (123) into (114), we can express $\mathcal{S}_2(x)$ as the finite sum

$$\begin{aligned} \mathcal{S}_2(x) &= \frac{\text{Li}_2(1-x) + \ln^2(1-x)}{1-x} + \frac{\ln(1-x)}{2} \sum_{q=1}^{i'-j'} \frac{\ln(1-x)}{x^q} \\ &+ \sum_{q=1}^{i'-j'-1} \left(\frac{\ln(1-x)H(i'-j'-q)}{x^q} - \frac{\ln(1-x)H(q)}{x^{q+1}} \right. \\ &\quad \left. - \frac{1}{x^q} \sum_{r=1}^{i'-j'-q} \frac{H(r+q-1) - H(r-1)}{r} \right). \end{aligned} \quad (124)$$

The corresponding derivatives can be obtained after tedious algebra as follows:

$$\frac{d^r}{dx^r} \mathcal{S}_2(x) = \frac{r!}{(1-x)^{r+1}} \xi_x(r) \quad (125)$$

where

$$\begin{aligned} \xi_x(r) &= \text{Li}_2(1-x) + \ln^2(1-x) - 2H(r) \ln(1-x) \\ &+ \sum_{b=1}^r \left(\frac{2H(b-1) - f_{1,b-1}(x)}{b} \right) \\ &+ \frac{1}{2} \sum_{q=1}^{\delta} \left(\ln(1-x) f_{q,r}(x) - \sum_{b=0}^{r-1} \frac{f_{q,b}(x)}{r-b} \right) \\ &+ \sum_{q=1}^{\delta-1} \left(H(\delta-q) f_{q,r}(x) - H(q) f_{q+1,r}(x) + \mu_{q,r}(x) K(q) \right) \end{aligned} \quad (126)$$

where $\delta = i' - j'$, and recall that $\text{Li}_2(\cdot)$ is the *dilogarithm* function [37, eq. (27.7.1)]. Also, $K(\cdot)$ is a constant given by

$$K(q) = \sum_{t=1}^{\delta-q} \frac{H(t+q-1) - H(t-1)}{t} \quad (127)$$

and

$$f_{q,r}(x) = \sum_{t=0}^{r-1} \frac{\mu_{q,t}(x)}{r-t} - \mu_{q,r}(x) \ln(1-x) \quad (128)$$

where

$$\mu_{q,r}(x) = \binom{q+r-1}{r} \frac{(x-1)^{r+1}}{x^{r+q}}. \quad (129)$$

Substituting (125) into (112), we obtain

$$\begin{aligned} \mathcal{S}_1(H(i), H(j), x) &= \frac{\Gamma(\tau+i')}{(1-x)^{\tau+j'}} \sum_{b=0}^{i'-1} \binom{i'-1}{b} \left(\frac{x}{1-x} \right)^b \\ &\times \frac{(\tau+j'+b-1)!}{(\tau+b)!} \xi_x(\tau+j'+b-1) \\ &= \frac{x^{i'-1} \Gamma(\tau+i')}{(1-x)^z} \sum_{b=0}^{i'-1} \binom{i'-1}{b} \left(\frac{1-x}{x} \right)^b \\ &\times (\tau+i'-b)_{j'-1} \xi_x(z-b-1) \\ &= \frac{x^{i'-1} \eta_{j',i'}(\xi_x(z-1), \sqrt{x})}{(1-x)^z}. \end{aligned} \quad (130)$$

Finally, substituting (130), (111), (110), and (106) into (100), and then combining with (98) and simplifying, we obtain the desired finite-sum expression in (27). \square

APPENDIX IV PROOF OF THEOREM 3

Proof: We start by following [44] and [45] and applying a first-order Taylor approximation to (7) near $\gamma = 0$ to give

$$\mathcal{I}_k \approx \log_2(e) \frac{\gamma}{N_t} \text{tr}(\mathbf{H}_k \mathbf{H}_k^\dagger). \quad (131)$$

Note that, as also mentioned in [44] and [45], we emphasize that this result is only accurate for the low SNR regime; in general, requiring that the condition $\|(\gamma/N_t) \mathbf{H}_k \mathbf{H}_k^\dagger\| < 1$ is satisfied.

Now, substituting (131) into (47) and using (57), we write the variance of the MIMO-OFDM mutual information at low SNR as follows:

$$\begin{aligned} \text{Var}^0(\mathcal{I}_{\text{ofdm}}) &= (\log_2(e))^2 \left(\frac{\gamma}{N_t} \right)^2 \\ &\times \left(\frac{2}{N^2} \sum_{d=1}^{N-1} (N-d) E[\text{tr}(\mathbf{H}_0 \mathbf{H}_0^\dagger) \text{tr}(\mathbf{H}_d \mathbf{H}_d^\dagger)] \right. \\ &\quad \left. + \frac{1}{N} E[\text{tr}^2(\mathbf{H}_{\text{flat}} \mathbf{H}_{\text{flat}}^\dagger)] - E^2[\text{tr}(\mathbf{H}_{\text{flat}} \mathbf{H}_{\text{flat}}^\dagger)] \right) \end{aligned} \quad (132)$$

where \mathbf{H}_{flat} is a flat-fading i.i.d. Rayleigh fading channel matrix. From [7], we have the following results:

$$E[\text{tr}(\mathbf{H}_{\text{flat}} \mathbf{H}_{\text{flat}}^\dagger)] = N_r N_t \quad (133)$$

$$E[\text{tr}^2(\mathbf{H}_{\text{flat}} \mathbf{H}_{\text{flat}}^\dagger)] = N_r N_t (1 + N_r N_t). \quad (134)$$

For the remaining expectation in (132), we write

$$\begin{aligned} &E[\text{tr}(\mathbf{H}_0 \mathbf{H}_0^\dagger) \text{tr}(\mathbf{H}_d \mathbf{H}_d^\dagger)] \\ &= E \left[\sum_{i=1}^{N_r} \sum_{j=1}^{N_t} |(\mathbf{H}_0)_{i,j}|^2 \sum_{k=1}^{N_r} \sum_{\ell=1}^{N_t} |(\mathbf{H}_d)_{k,\ell}|^2 \right] \\ &= \sum_{i=1}^{N_r} \sum_{j=1}^{N_t} E[|(\mathbf{H}_0)_{i,j}|^2 |(\mathbf{H}_d)_{i,j}|^2] + (N_r N_t)^2 - N_r N_t \end{aligned} \quad (135)$$

where the second line followed by noting that $E[|(\mathbf{H}_0)_{i,j}|^2 |(\mathbf{H}_d)_{k,\ell}|^2] = 1$ for all $(i,j) \neq (k,\ell)$. Now using (5), it can be easily shown that

$$\begin{aligned} &E[|(\mathbf{H}_0)_{i,j}|^2 |(\mathbf{H}_d)_{i,j}|^2] \\ &= |\rho_d|^2 E[|(\mathbf{H}_d)_{i,j}|^4] + (1 - |\rho_d|^2) E[|(\mathbf{E})_{i,j}|^2] \\ &= 1 + |\rho_d|^2. \end{aligned} \quad (136)$$

Substituting (136) into (135), we find that

$$E[\text{tr}(\mathbf{H}_0 \mathbf{H}_0^\dagger) \text{tr}(\mathbf{H}_d \mathbf{H}_d^\dagger)] = N_r N_t (|\rho_d|^2 + N_r N_t). \quad (137)$$

The theorem now follows by substituting (137), (134), and (133) into (132) and then performing some basic simplifications. \square

REFERENCES

- [1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Commun.*, pp. 585–595, Nov.–Dec. 1999.

- [2] G. J. Foschini, "Layered space-time architecture for wireless communications in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, Autumn 1996.
- [3] A. Grant, "Rayleigh fading multi-antenna channels," *EURASIP J. Appl. Signal Process.*, Special Issue on Space-Time Coding (Part I), pp. 316–329, Mar. 2002.
- [4] 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.
- [5] P. J. Smith, S. Roy, and M. Shafi, "Capacity of MIMO systems with semicorrelated flat fading," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2781–2788, Oct. 2003.
- [6] 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.
- [7] 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.
- [8] A. L. Moustakas, S. H. Simon, and A. M. Sengupta, "MIMO capacity through correlated channels in the presence of correlated interferers and noise: A (not so) large N analysis," *IEEE Trans. Inf. Theory*, vol. 45, no. 10, pp. 2545–2561, Oct. 2003.
- [9] M. Kiessling and J. Speidel, "Mutual information of MIMO channels in correlated Rayleigh fading environments—A general solution," in *Proc. IEEE Int. Conf. Commun.*, Paris, France, Jun. 2004, pp. 814–818.
- [10] S. H. Simon, A. L. Moustakas, and L. Marinelli, "Capacity and character expansions: Moment generating function and other exact results for MIMO correlated channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5336–5351, Dec. 2006.
- [11] G. Alfano, A. Lozano, A. M. Tulino, and S. Verdú, "Mutual information and eigenvalue distribution of MIMO Ricean channels," in *Proc. Int. Symp. Inf. Theory Appl.*, Parma, Italy, Oct. 10–13, 2004.
- [12] A. Lozano, A. M. Tulino, and S. Verdú, "High-SNR power offset in multi-antenna communication," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4134–4151, Dec. 2005.
- [13] M. R. McKay and I. B. Collings, "General capacity bounds for spatially correlated Rician MIMO channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3121–3145, Sep. 2005.
- [14] A. M. Tulino, A. Lozano, and S. Verdú, "Impact of antenna correlation on the capacity of multi-antenna channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2491–2509, Jul. 2005.
- [15] M. Kang and M.-S. Alouini, "Capacity of MIMO Rician channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 112–122, Jan. 2006.
- [16] M. R. McKay and I. B. Collings, "Improved general lower bound for spatially-correlated Rician MIMO capacity," *IEEE Commun. Lett.*, vol. 10, no. 3, pp. 162–164, Mar. 2006.
- [17] M. R. McKay, P. J. Smith, and I. B. Collings, "New properties of complex noncentral quadratic forms and bounds on MIMO mutual information," in *Proc. IEEE Int. Symp. Inf. Theory*, Seattle, WA, Jul. 2006, pp. 1209–1213.
- [18] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Found. Trends Commun. Inf. Theory*, vol. 1, no. 1, pp. 1–182, 2004.
- [19] H. Bölcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 225–234, Feb. 2002.
- [20] P. J. Smith, L. M. Garth, and S. Loyka, "Exact capacity distributions for MIMO systems with small numbers of antennas," *IEEE Commun. Lett.*, vol. 7, no. 10, pp. 481–483, Oct. 2003.
- [21] P. J. Smith and L. M. Garth, "Exact capacity distribution for dual MIMO systems in Ricean fading," *IEEE Commun. Lett.*, vol. 8, no. 1, pp. 18–20, Jan. 2004.
- [22] P. J. Smith and M. Shafi, "On a Gaussian approximation to the capacity of wireless MIMO systems," in *Proc. IEEE Int. Conf. Commun.*, New York, NY, Apr. 2002, vol. 1, pp. 406–410.
- [23] Z. Wang and G. B. Giannakis, "Outage mutual information of space-time MIMO channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 4, pp. 657–662, Apr. 2004.
- [24] J. Wang and K. Yao, "Capacity scaling in OFDM based spatial multiplexing systems," in *Proc. IEEE Veh. Technol. Conf.*, Vancouver, BC, Canada, Sep. 2002, pp. 28–32.
- [25] O. Oyman, R. U. Nabar, H. Bölcskei, and A. J. Paulraj, "Characterizing the statistical properties of mutual information in MIMO channels," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2784–2795, Nov. 2003.
- [26] K. Liu, V. Raghavan, and A. M. Sayeed, "Capacity scaling and spectral efficiency in wide-band correlated MIMO channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2504–2526, Oct. 2003.
- [27] L. S. Pillutla and S. K. Jayaweera, "MIMO capacity of an OFDM-based system under Ricean fading," in *Proc. IEEE Veh. Technol. Conf.*, Milan, Italy, May 2004, pp. 618–622.
- [28] M. R. McKay and I. B. Collings, "On the capacity of frequency-flat and frequency-selective Rician MIMO channels with single-ended correlation," *IEEE Trans. Wireless Commun.*, vol. 5, no. 8, pp. 2038–2043, Aug. 2006.
- [29] P. L. Kuffe, A. B. Sesay, and J. McRory, "Capacity of MIMO-OFDM systems in spatially correlated indoor fading channels," in *IEEE Veh. Technol. Conf.*, Los Angeles, CA, Sep. 2004, pp. 129–133.
- [30] H. A. Suraweera, J. T. Y. Ho, T. Sivaumaran, and J. Armstrong, "An approximated Gaussian analysis and results on the capacity distribution for MIMO-OFDM," in *Proc. IEEE Int. Symp. Personal Indoor Mobile Radio Commun.*, Berlin, Germany, Sep. 2005, pp. 211–215.
- [31] A. Clarke, P. J. Smith, and D. P. Taylor, "Instantaneous capacity of OFDM on Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 355–361, Jan. 2007.
- [32] A. L. Moustakas and S. H. Simon, "On the outage capacity of correlated multiple-path MIMO channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 3887–3903, Nov. 2007.
- [33] G. Barriac and U. Madhow, "Characterizing outage rates for space-time communication over wideband channels," *IEEE Trans. Commun.*, vol. 52, no. 12, pp. 2198–2207, Dec. 2004.
- [34] Y. Li, L. J. Cimini, and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 902–915, Jul. 1998.
- [35] M. Dohler, "Virtual antenna arrays," Ph.D. dissertation, King's College London, Univ. London, London, U.K., 2003.
- [36] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 4th ed. San Diego, CA: Academic, 1965.
- [37] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 4th ed. New York: Dover, 1965.
- [38] E. Ko and D. Hong, "Improved space-time block-coding with frequency diversity for OFDM systems," in *Proc. IEEE Int. Conf. Commun.*, Paris, France, 2004, pp. 3217–3220.
- [39] S. H. Müller-Weinfurter, "Coding approaches for multiple antenna transmission in fast fading and OFDM," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2442–2450, Oct. 2002.
- [40] M. Kang, L. Yang, and M.-S. Alouini, "How accurate are the Gaussian and Gamma approximations to the outage capacity of MIMO channels," in *Proc. 6th Baiona Workshop Signal Process. Commun.*, Baiona, Spain, Sep. 2003.
- [41] P. J. Smith and L. M. Garth, "Distribution and characteristic functions for correlated complex Wishart matrices," *J. Multivariate Anal.*, vol. 98, no. 4, pp. 661–677, Apr. 2007.
- [42] S. Wang and A. Abdi, "On the second-order statistics of the instantaneous mutual information in Rayleigh fading channels," *IEEE Trans. Inf. Theory* 2006 [Online]. Available: <http://arxiv.org/pdf/cs/0603027>, submitted for publication
- [43] B. de Neumann, "An interesting result arising from the analysis of diversity receivers," *Bull. IMA*, vol. 27, pp. 48–50, Mar. 1991.
- [44] M. Kiessling, J. Speidel, and A. Boronka, "Asymptotics of ergodic MIMO capacity in correlated Rayleigh fading environments," in *Proc. IEEE Veh. Technol. Conf.*, Genoa, Italy, May 2004, pp. 843–847.
- [45] L. W. Hanlen and A. J. Grant, "Optimal transmit covariance for ergodic MIMO channels," *IEEE Trans. Inf. Theory* 2005 [Online]. Available: <http://arxiv.org/abs/cs/0510060>, submitted for publication