Capacity of MIMO Systems With Semicorrelated Flat Fading

Peter J. Smith, *Member, IEEE*, Sumit Roy, *Senior Member, IEEE*, and Mansoor Shafi, *Fellow, IEEE*

***Abstract*—The primary contribution of this work lies in the derivation of the *exact* characteristic function (and hence, the mean and variance) of the capacity of multiple-input multiple-output (MIMO) systems for semi-correlated flat-fading channels. A Gaussian approximation to the exact capacity results is suggested and evaluated for its accuracy. We show that over a range of correlation levels this approximation is adequate even for moderate numbers of transmit and receive antennas.**

***Index Terms*—Multiple-input multiple-output (MIMO) systems, Shannon capacity, spatial correlation.**

I. Introduction

There continues to be substantial interest in wireless communication systems that employ multiple transmit and receive antennas, due to their promise for dramatically increasing the capacity (or, equivalently, spectral efficiency) without requiring bandwidth expansion. The concept of such multiple-input multiple-output (MIMO) wireless systems was pioneered by Foschini and coworkers [1]–[4] and developed into the Bell Labs layered space–time (BLAST) architecture that report achieving spectral efficiencies in the range of 10–20 bits/s/Hz for typical configurations.[1] Derivation of capacity and other pertinent figures of merit for channels with independent gains should be credited to [5], [1], [3]; further results appear in [6], [7]. Throughout this correspondence, we assume a "quasi-stationary" channel [1] where capacity is interpreted as a random variable. Hence, we concentrate on the capacity distribution, leading to outage probabilities, rather than ergodic capacity.

Extensions of this work to correlated channels have begun to appear rapidly in the literature [8]–[12]. It is now well known that MIMO capacity is very sensitive to the presence of spatial fading correlation [8] which may be present at either or both ends of the radio link. As a point of notation, we denote a system with t transmit and r receive antennas as a $t \rightarrow r$ system. In fact, as shown by Shiu *et al.* [8], when the angular spread reduces, the correlation between the elements of the channel gain matrix increases. In turn, the capacity of an $n \rightarrow n$ MIMO system decreases and approaches closely that of a $1 \rightarrow n$ single-input multiple-output (SIMO) system.

The highest spectral efficiency of a point-to-point MIMO system is only achieved when there is uncorrelated fading among pairs of transmitter and receiver antennas. In practice, this can be achieved with sufficient spacing among the base station and (mobile or portable) terminal station antennas. At the base station, decorrelation is achieved using approx. 10λ separation between nearest elements in a linear array [11]. A mobile station operating in an urban outdoor area (say, in the central business district) or indoors in a home or office environment is likely to be surrounded by multiple scatterers that contribute to wide angle scattering. Therefore, only 0.5λ spacing may be adequate. At 2 GHz, this wavelength is 15 cm. This means that even a modest four-element antenna array at the base station will have a span of 4.5 m. Thus, contrary to popular belief, the spacing issue is of considerable significance at the base station where mounting of antennas is subject to strict environmental regulations. Note that design and mounting of closely spaced antennas on small form-factor portable devices (lap tops, PDAs, etc.) poses additional problems (to just using separation to decorrelate received signals) due to coupling via the substrate.

In this correspondence, we consider channels with correlation at one end only and we denote these channels as "semicorrelated" following [13]. Note that recent measurements conducted in downtown Helsinki, Finland show that the semicorrelated channel model is valid for certain urban environments [14].

In particular, system engineers would like to have the ability to exactly predict the influence on capacity due to the presence of correlation at either end and establish the onset of diminishing returns.

The main contributions of the correspondence are as follows:

- the derivation of the exact characteristic function of the capacity distribution under correlated fading at either end;
- differentiating the characteristic function to derive the first two moments of the capacity distribution;
- validating the above results for a number of correlation scenarios and array sizes (including equal and unequal t, r values).

Despite their obvious complexity, the exact results have value in allowing a calibration of simulations with the corresponding closed-form results. Once implemented, of course, results are produced more quickly than simulations and without simulation error. In an earlier paper [15], we showed that the capacity of a MIMO system for uncorrelated channels can be successfully approximated by a Gaussian distribution. In this correspondence, we show that for the correlated cases considered here, the Gaussian approximation is still effective. This result follows from simulation of MIMO capacities of correlated channels and comparison with a Gaussian distribution where the mean and variance were computed using the exact results derived in this correspondence. For the uncorrelated case, the distribution of the standardized capacity converges to Gaussian under various limiting regimes [9], [15]. For the correlated case, we conjecture that this will also occur under suitable constraints on the correlation structure. In this correspondence, however, we simply demonstrate the accuracy of a Gaussian approximation by experimentation. Note that the characteristic function for the uncorrelated case has been derived in [16].

The channel model assumed for correlated fading is the well-known separable model where separate processes induce correlation at the transmitter and receiver. This approach is now considered an acceptable basis for research into correlated fading [8]–[12], [17] and also appears in the standards literature [18]. Recent work in this area [8]–[10], [12], [19] gives capacity results under limiting scenarios, but to the best of our knowledge no exact results are available, with the exception of some work reported in [12].

The correspondence is organized as follows: Section II formulates the problem. Section III contains the derivation of the exact characteristic function of the capacity and its mean and variance. Section IV

Manuscript received October 30, 2002; revised April 27, 2003.

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).

S. Roy is with Wireless Technology, Communications and Interconnect Laboratories, Intel Corporation, Hillsboro, OR 97124, USA on leave from the Department of Electrical Engineering, University of Washington, Seattle, WA 98195 USA (e-mail: roy@ee.washington.edu).

M. Shafi is with Telecom New Zealand, Wellington, New Zealand (e-mail: mansoor.shafi@telecom.co.nz).

Communicated by B. M. Hochwald, Guest Editor.

Digital Object Identifier 10.1109/TIT.2003.817472

[1]Just recently, Lucent Technologies announced chip sets that achieve 19.2 Mbits/s for third-generation (3G) cellular networks using four transmit and receive antennas.

0018-9448/03$17.00 © 2003 IEEE

provides a Gaussian approximation based on the exact results and some numerical checks of the analysis. Finally, in Section V, some conclusions are given.

II. Problem Formulation

Consider a MIMO system where each user transmits simultaneously via t antennas and reception is via r antennas. The total power of the complex transmitted signal $\boldsymbol{s}$ is constrained to P regardless of the number of antennas. The received signal $\boldsymbol{r}$ is given by

$$\boldsymbol{r} = \boldsymbol{X}\boldsymbol{s} + \boldsymbol{n} \tag{1}$$

where $\boldsymbol{n}$ is a complex r-dimensional additive white Gaussian noise (AWGN) vector, with statistically independent components of identical power σ^2 at each of the r receive branches and $\boldsymbol{X}$ is an $r \times t$ (complex) matrix of independent and identically distributed (i.i.d.) circular Gaussian variables with zero mean and unit total variance as is appropriate for independent and identical fading channels. Without loss of generality, we assume $\sigma^2 = 1$ for the remainder of the correspondence. The capacity of such a MIMO system with no channel state information at the transmitter is given by [5], [1], [3]

$$C = \log_2 \left(\det \left(\boldsymbol{I}_r + \frac{P}{t} \boldsymbol{X}\boldsymbol{X}^\dagger \right) \right) \tag{2}$$

In the case of correlated fading, the most common models assume that the fading is induced by separate physical processes at the transmitter and receiver [8]. This leads to the channel matrix being modeled as $\boldsymbol{AXB}$ where $\boldsymbol{A}$ induces correlation at the receiver end, $\boldsymbol{B}$ induces correlation at the transmitter side, and $\boldsymbol{X}$ is as defined above. Using this model we can write the capacity as

$$\begin{aligned} C &= \log_2 \left(\det \left(\boldsymbol{I}_r + \frac{P}{t} \boldsymbol{AXBB}^\dagger \boldsymbol{X}^\dagger \boldsymbol{A}^\dagger \right) \right) \\ &= \log_2 \left(\det \left(\boldsymbol{I}_r + \frac{P}{t} \boldsymbol{AX}\boldsymbol{\Phi}_B \boldsymbol{\Gamma}_B \boldsymbol{\Phi}_B^\dagger \boldsymbol{X}^\dagger \boldsymbol{A}^\dagger \right) \right) \\ &= \log_2 \left(\det \left(\boldsymbol{I}_r + \frac{P}{t} \boldsymbol{A}\tilde{\boldsymbol{X}} \boldsymbol{\Gamma}_B \tilde{\boldsymbol{X}}^\dagger \boldsymbol{A}^\dagger \right) \right) \end{aligned} \tag{3}$$

where $\boldsymbol{\Phi}_B$ is the unitary matrix defined by the matrix decomposition $\boldsymbol{BB}^\dagger = \boldsymbol{\Phi}_B \boldsymbol{\Gamma}_B \boldsymbol{\Phi}_B^\dagger$ and $\boldsymbol{\Gamma}_B$ is the diagonal matrix of eigenvalues of $\boldsymbol{BB}^\dagger$. Since $\boldsymbol{\Phi}_B$ is unitary, the statistics of $\tilde{\boldsymbol{X}}$ are identical to those of $\boldsymbol{X}$ and we drop the ˜ superscript for convenience. For correlation at the transmitter only we have the result

$$C = \log_2 \left(\det \left(\boldsymbol{I}_r + \frac{P}{t} \boldsymbol{X}\, \boldsymbol{\Gamma}_B \boldsymbol{X}^\dagger \right) \right). \tag{4}$$

If the correlation is at the receiver end only then we write

$$\begin{aligned} C &= \log_2 \left(\det \left(\boldsymbol{I}_r + \frac{P}{t} \boldsymbol{AXX}^\dagger \boldsymbol{A}^\dagger \right) \right) \\ &= \log_2 \left(\det \left(\boldsymbol{I}_t + \frac{P}{t} \boldsymbol{X}^\dagger \boldsymbol{A}^\dagger \boldsymbol{AX} \right) \right) \\ &= \log_2 \left(\det \left(\boldsymbol{I}_t + \frac{P}{t} \tilde{\boldsymbol{X}}^\dagger \boldsymbol{\Gamma}_A \tilde{\boldsymbol{X}} \right) \right) \end{aligned} \tag{5}$$

where $\boldsymbol{A}^\dagger \boldsymbol{A} = \boldsymbol{\Phi}_A \boldsymbol{\Gamma}_A \boldsymbol{\Phi}_A^\dagger$ using the same approach as before. Hence, both types of one-sided correlation result in a capacity equation of the same form (see (4) and (5)).

III. Derivation

Note that all the cases discussed in Section II yield a capacity expression of the form

$$C = \log_2 \left(\det \left(\boldsymbol{I}_M + \boldsymbol{X}\, \boldsymbol{\Gamma}\, \boldsymbol{X}^\dagger \right) \right) \tag{6}$$

where, for example, $\boldsymbol{\Gamma} = \frac{P}{t} \boldsymbol{\Gamma}_A$ in the case of correlation at the receiver end and M can be r or t depending on the application. Our analysis depends on the assumption that $M \le N$ where M is the number of rows of $\boldsymbol{X}$ and N is the number of columns. This includes the most useful special case where $M = N$, i.e., $r = t$. The requirement that $M \le N$ stems from the methodology used (see Appendixes A and B) which is based on the distribution of $\boldsymbol{X}\, \boldsymbol{\Gamma}\, \boldsymbol{X}^\dagger$ given in [20] for this particular case.

Let $\boldsymbol{\Gamma} = \operatorname{diag}(\gamma_1, \ldots, \gamma_N)$ where $\gamma_1 > \gamma_2, \ldots, > \gamma_N$. The characteristic function of C is given by

$$\Phi(s) = E\left(e^{jsC}\right) = E\left(\exp\left(js\ \log_2 \prod_{i=1}^{M} (1 + x_i) \right) \right) \tag{7}$$

where $x_1, \ldots, x_M$ are the nonnegative (unordered) eigenvalues of $\boldsymbol{X\Gamma X}^\dagger$. Letting $t = \frac{s}{\log_e(2)}$ yields

$$\begin{aligned} \Phi(s) &= E\left(\exp\left(jt\ \log_e \left(\prod_{i=1}^{M} (1 + x_i) \right) \right) \right) \\ &= E\left(\prod_i (1 + x_i)^{jt} \right) = \phi_1(t). \end{aligned} \tag{8}$$

In Appendix A, we show that $\phi_1(t)$ can be written as a ratio of determinants

$$\phi_1(t) = \frac{|\Delta_{12} \boldsymbol{I}_0\, \boldsymbol{I}_1 \ldots \boldsymbol{I}_{M-1}|}{\prod_{k=1}^{M-1} k!\ \Delta_3} \tag{9}$$

where

$$\Delta_3 = \begin{vmatrix} 1 & \gamma_1 & . & \gamma_1^{N-1} \\ 1 & \gamma_2 & . & \gamma_2^{N-1} \\ . & . & . & . \\ 1 & \gamma_N & . & \gamma_N^{N-1} \end{vmatrix} \tag{10}$$

is a Vandermonde matrix and Δ_{12} is an $N \times N - M$ submatrix of Δ_3 consisting of the first $N - M$ columns. The remaining terms are defined by

$$\boldsymbol{I}_k = [I_{1k}\ I_{2k} \ldots I_{Nk}]^T$$

where

$$I_{ik} = \int_0^\infty x^k (1 + x)^{jt} \exp\left(-\frac{x}{\gamma_i} \right) \gamma_i^{N-M-1}\, dx$$

is a confluent hypergeometric function [21]. Note that Δ_3 has the alternative form $\Delta_3 = \prod_{i>k} (\gamma_i - \gamma_k)$ where we assume that $\gamma_i \neq \gamma_k\ \forall i \neq k$. If two γ_k values are equal then both the numerator and the denominator of $\phi_1(t)$ are zero and this form cannot be used. The assumption of unequal γ's can be relaxed but the analysis then requires taking the limit of (9) as, say, $\gamma_i \to \gamma_k$. This is possible but is perhaps of marginal interest. The important case where equal eigenvalues occur is the classical one where $\boldsymbol{\Gamma}$ is proportional to the identity matrix and this is well known. For semicorrelated channels, the γ_k values result from correlation measurements or models. If based on measurements, they are unequal with probability one. Also, all channel models known to the authors lead to unequal γ_k values, including those in standards models [18] and in the literature [8], [13], [17]. Hence, in this correspondence, we only consider the case of distinct eigenvalues. More problematic is the issue of numerical robustness of the exact results when eigenvalues become extremely close to each other. This is discussed in Section IV.

The mean and variance is now readily obtained from the characteristic function as follows:

$$E[C] = -j\Phi'(0) = \frac{1}{\log_e 2} \left(-j\phi_1'(0) \right) \tag{11}$$

$$Var.[C] = \frac{1}{(\log_e 2)^2} \left(\phi_1'(0)^2 - \phi_1''(0) \right). \tag{12}$$

In Appendix B, we show how the characteristic function can be differentiated to give $\phi_1'(0)$ and $\phi_1''(0)$. Substituting these in (11), (12) gives the following results:

$$E[C] = \frac{1}{\log_e 2} \sum_{i=0}^{M-1} \sum_{k=1}^{N} c_{ik}\, I(i,\, \gamma_k) \tag{13}$$

where

$$c_{ik} = \frac{(-1)^{(M-i-1)}\gamma_k^{(N-M-1)}}{i!} \left(\prod_{h\neq k}(\gamma_k - \gamma_h)\right)^{-1} \cdot \sum_{1\leq i_1<\ldots<i_{M-i-1}\leq N}^{i_r\neq k} \gamma_{i_1} \cdots \gamma_{i_{M-i-1}} \tag{14}$$

and

$$I(i,\, \gamma_k) = \int_0^\infty \log_e(1+x)\; x^i\, e^{-\frac{x}{\gamma_k}}\, dx. \tag{15}$$

Evaluation of (15) can be done numerically, or from the expression

$$I(i,\,\gamma) = \sum_{r=0}^{i} \frac{i!(-1)^{i-r}}{(i-1)!}\, \gamma^{r+1}\, E_1\left(\frac{1}{\gamma}\right) + \sum_{r=1}^{i}\sum_{k=0}^{r-1}\sum_{h=0}^{r-k-1} \frac{i!(-1)^{i-r}\gamma^{h+k+2}}{(i-r)!(r-k-1-h)!(r-k)} \tag{16}$$

which follows using integration by parts and the definition of the exponential integral $E_1(\cdot)$. The coefficient c_{ik} may be more easily evaluated using its representation as a determinant (see (36) in Appendix A). Note that the form of (13) is similar to Telatar's result [5] for the mean capacity in the uncorrelated case. Both are finite sums of integrals which can be expressed in terms of the exponential integral. Hence, although the correlated case adds N extra parameters, their effect is simply to make the coefficients more complicated. The variance is given by (17) at the bottom of the page, where c, β_{ik}, and δ_{ikrs} are constants and I_{ik} is given above. These terms are defined as follows:

$$c = \left(\prod_{k=1}^{M-1} k!\right)^{-1} \Delta_3^{-1} \tag{18}$$

$$I'_{ik}(0) = j\gamma_k^{N-M-1} I(i,\,\gamma_k) \tag{19}$$

$$I''_{ik}(0) = -\gamma_k^{N-M-1} \int_0^\infty [\log_e(1+x)]^2 x^i \exp\left(-\frac{x}{\gamma_k}\right) dx \tag{20}$$

$$\beta_{ik} = \left(\prod_{r=1}^{M-1} r!\right) \frac{\alpha_{ik}}{i!} \tag{21}$$

where α_{ik} is the determinant of Δ_3 with column i and row k removed

$$\delta_{ikrs} = \left(\prod_{n=1}^{M-1} n!\right) \frac{\alpha_{ikrs}}{i!\,k!} \tag{22}$$

where α_{ikrs} is the determinant of Δ_3 with columns $i,\, k$ and rows $r,\, s$ removed. Again, the result is similar to the uncorrelated case derived in [15] with the basic form of the expression remaining the same but the extra parameters causing an increase in the complexity of the coefficients.

IV. Results

We first focus on numerical verification of the results on the exact mean and variance. Here, we consider a simple scenario with one-sided correlation in a $4 \to 4$ MIMO system. The correlation has a simple exponential decay form where $\mathrm{Corr}(\boldsymbol{X}_{ik}\boldsymbol{X}_{hk}) = \rho^{|i-h|}$ for $k \in \{1,\, 2,\, 3,\, 4\}$ and $0 \leq \rho \leq 1$. We refer to ρ as the exponential correlation parameter. This gives correlation at the receiver (or transmitter) which drops off exponentially with antenna separation. For example, an equally spaced linear array with antennas sequentially numbered has $|i-h|$ proportional to the distance between antenna i and h. For this model, we simulate 5000 capacity values for a variety of signal-to-noise ratio (SNR) and ρ values. In Figs. 1 and 2, we plot the mean capacity and capacity variance versus ρ for SNR $\in \{3\text{ dB},\, 9\text{ dB},\, 15\text{ dB}\}$. Good agreement is observed between simulated values and the exact calculations. Also shown in this correlation scenario is the expected drop in mean capacity and capacity variance as correlation increases. In the extreme case of perfect correlation, there is no diversity at one end of the link and so the capacity will become that of a SIMO or a multiple-input single-output (MISO) system. In addition, we note that the variance appears to be more sensitive than the mean to changes in correlation. As mentioned in Section III, when γ values become very close to each other, the presence of difference terms, especially in the denominator can be a problem in terms of numerical computation. Producing algorithms which are robust to all parameter values is beyond the scope of the correspondence. Our own implementation of the results begins to encounter difficulties around $N = 10$. Beyond this point, special care needs to be taken in the handling of the computations.

Next we investigate a Gaussian approximation to the capacity distribution. Here, we consider three physical models for the correlated channel in a $4 \to 4$ MIMO system. In particular, we use the correlation model in [17] specified by the average angle of arrival or departure (AOA), the angle spread (AS), the distribution of the angles, and equally spaced antennas with separation d. Numerical values for the parameters are given in Table I. These parameters were then used to compute the elements of a corresponding channel correlation matrix for the three scenarios. We compute the eigenvalues for these models which form the diagonal entries of the $\boldsymbol{\Gamma}$ matrix in (6). These eigenvalues are plotted in Fig. 3, normalized so that the maximum eigenvalue is 1 for ease of comparison. Note that there is a wide spread of correlation here; the low correlation eigenvalues decay slowly whereas the highly correlated case has essentially only one eigenvalue of any size.

For the three scenarios in Table I we plot, in Fig. 4, the empirical distribution function of 10 000 simulated capacity values. Superimposed on each curve is the Gaussian approximation fitted from the exact mean and variance values. As for the uncorrelated case, good agreement is observed between the empirical distribution and the Gaussian approximation even down to lower tail probabilities of the order of 10^{-3}. This observation holds even for the high correlation scenario where there is only one eigenvalue of any size. Here the Gaussian approximation is likely to be most suspect. In [22], it was also shown that temporal sequences of capacity variables can be approximated by a Gaussian process in the presence of both spatial and temporal correlation.

A more thorough investigation of the Gaussian approximation is given in Fig. 5. There are seven curves in Fig. 5 corresponding to two levels of correlation ($\rho = 0.1,\, 0.9$) by three sets of antenna numbers ($r = t/2$, $r = t$, $r = 2 \times t$). The seventh curve (bold) corresponds to a true Gaussian random variable. For each scenario, a Kolmogorov–Smirnov goodness of fit test for a Gaussian distribution was

$$\mathrm{Var}[C] = \frac{-1}{(\log_e 2)^2} \left(c \sum_{i=0}^{M-1}\sum_{k=1}^{N} \beta_{ik} I''_{ik}(0) + 2c \sum_{0\leq i<k\leq M-1} \sum_{r=1}^{N} \sum_{s=1}^{N(s\neq r)} \delta_{ikrs} I'_{ir}(0) I'_{ks}(0) \right) - E[C]^2 \tag{17}$$

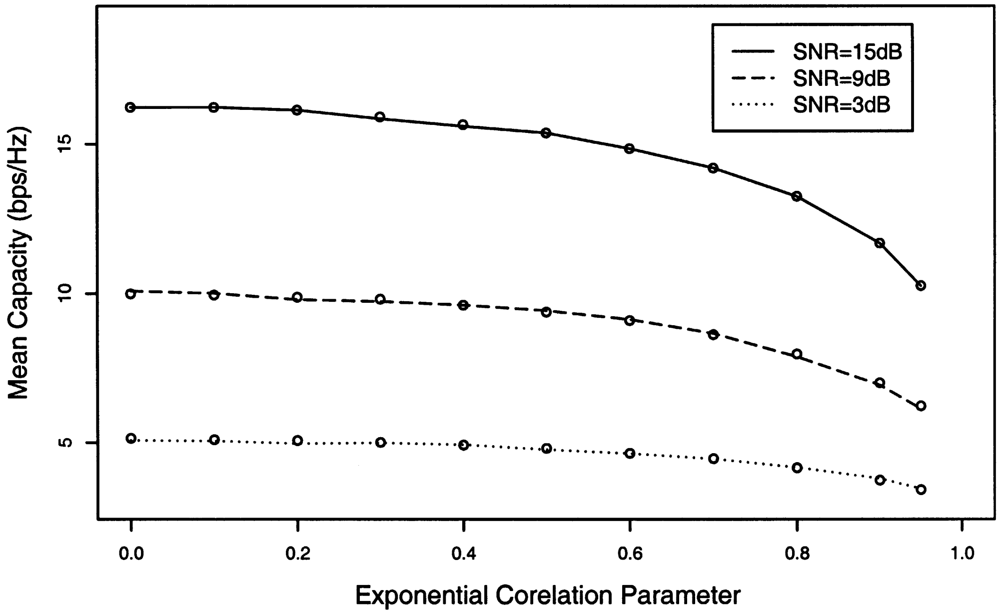

Fig. 1. Plot of mean capacity versus exponential correlation parameter for a $4 \rightarrow 4$ MIMO system. Simulated results are given by the lines and analytical results are given by the points.

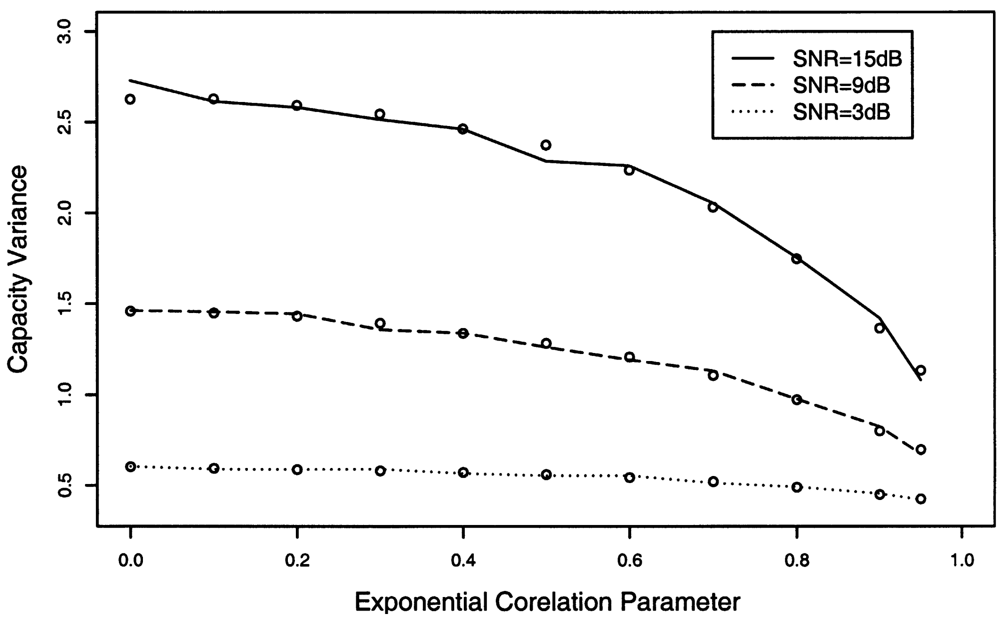

Fig. 2. Plot of capacity variance versus exponential correlation parameter for a $4 \rightarrow 4$ MIMO system. Simulated results are given by the lines and analytical results are given by the points.

TABLE I
PARAMETER VALUES FOR THE THREE CORRELATION SCENARIOS

Scenario	$AOA°$	$AS°$	d	Distribution
Low correlation	0	72	0.5λ	Von Mises
Moderate correlation	0	18	0.5λ	Von Mises
High correlation	90	10	0.3λ	Von Mises

performed. Large values of the test statistic correspond to large deviations from Gaussianity. The following conclusions can be drawn from Fig. 5. For $m \geq 3$, all scenarios are similar and compare very well to a Gaussian distribution. For $m < 3$, correlation reduces the goodness of fit. For fixed m, $r = t$ has the worst fit and $r = 2 \times t$ is the best fitting.

V. CONCLUSION

We have derived a closed-form characteristic function for the capacity of a MIMO system in semicorrelated flat-fading channels. This provides closed-form expressions for the mean and variance and thus enables an investigation of a Gaussian approximation to the capacity distribution. We have shown that for moderate numbers of antennas ($m \geq 3$) the Gaussian approximation is adequate over a wide range of correlations.

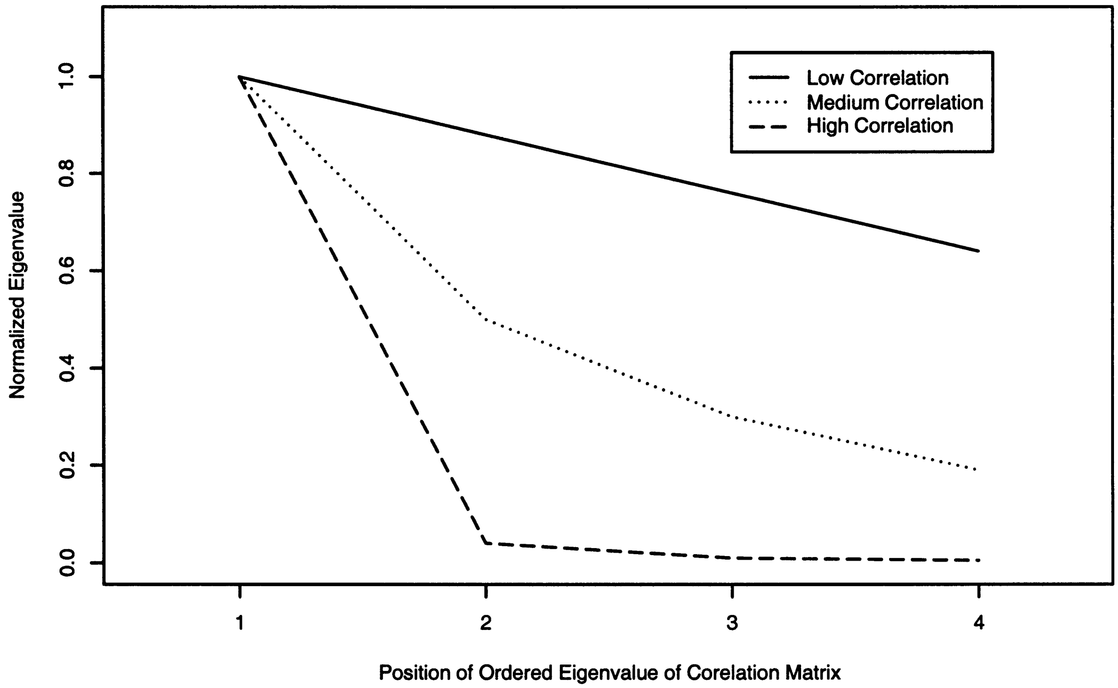

Fig. 3. Plot of the normalized eigenvalues for the correlation scenarios versus eigenvalue position for a $4 \rightarrow 4$ MIMO system.

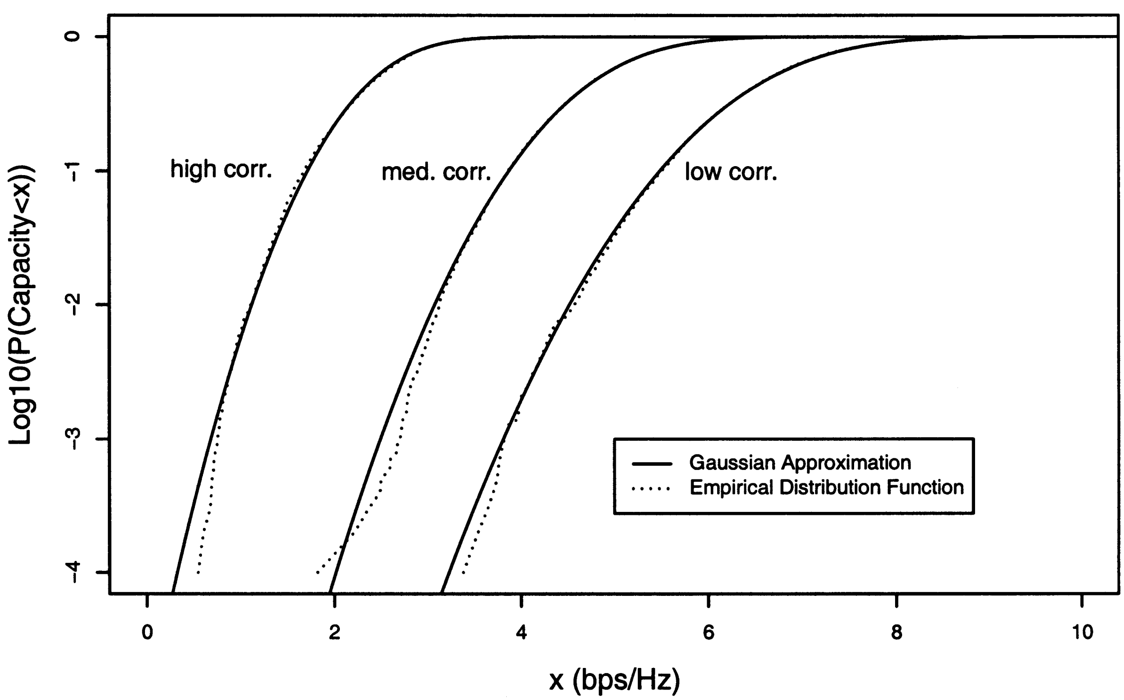

Fig. 4. Capacity distribution functions for a $4 \rightarrow 4$ system under low, medium, and high correlation scenarios (SNR $= 15$ dB).

Appendix A

Consider the distribution of the *ordered* eigenvalues $x_{(1)} \leq \cdots \leq x_{(M)}$ of the $M \times M$ random Hermitian matrix $\boldsymbol{Z}$ given by

$$\boldsymbol{Z} = \boldsymbol{V} \, \Gamma \, \boldsymbol{V}^{\dagger}$$

where $\boldsymbol{V}$ is $M \times N$ normal with i.i.d. $\mathcal{CN}(0, 1)$ entries, $M \leq N$, and

$$\boldsymbol{\Gamma} = \mathrm{diag}[\gamma_1, \ldots, \gamma_N]$$

with $\gamma_i > 0, \forall i$. Collecting the $\frac{M(M+1)}{2}$ distinct elements of $\boldsymbol{Z}$ in the vector $\boldsymbol{z}$, the results of [20] state that the joint probability density function (pdf) of $\boldsymbol{Z}$ is given by

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = \pi^{-M(M-1)/2} \frac{\Delta_1(\boldsymbol{z})}{\Delta_2(\boldsymbol{z}) \Delta_3} \tag{23}$$

where $\Delta_1, \Delta_2, \Delta_3$ are various determinants. Now, it is known from [23] that the joint pdf of the ordered eigenvalues $\boldsymbol{x}_o = [x_{(1)}, \ldots, x_{(M)}]$ of $\boldsymbol{Z}$ is related to that of $\boldsymbol{z}$ via

$$f_{\boldsymbol{X}_o}(\boldsymbol{x}_o) = \frac{\pi^{M(M-1)/2}}{\prod_{k=1}^{M-1} k!} \Delta_2^2(\boldsymbol{x}_o) \, f_{\boldsymbol{Z}}(\boldsymbol{x}_o) \tag{24}$$

From (23) and (24) we have

$$f_{\boldsymbol{X}_o}(\boldsymbol{x}_o) = \frac{1}{\prod_{k=1}^{M-1} k!} \frac{\Delta_1(\boldsymbol{x}_o)\Delta_2(\boldsymbol{x}_o)}{\Delta_3} \quad x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(M)} \tag{25}$$

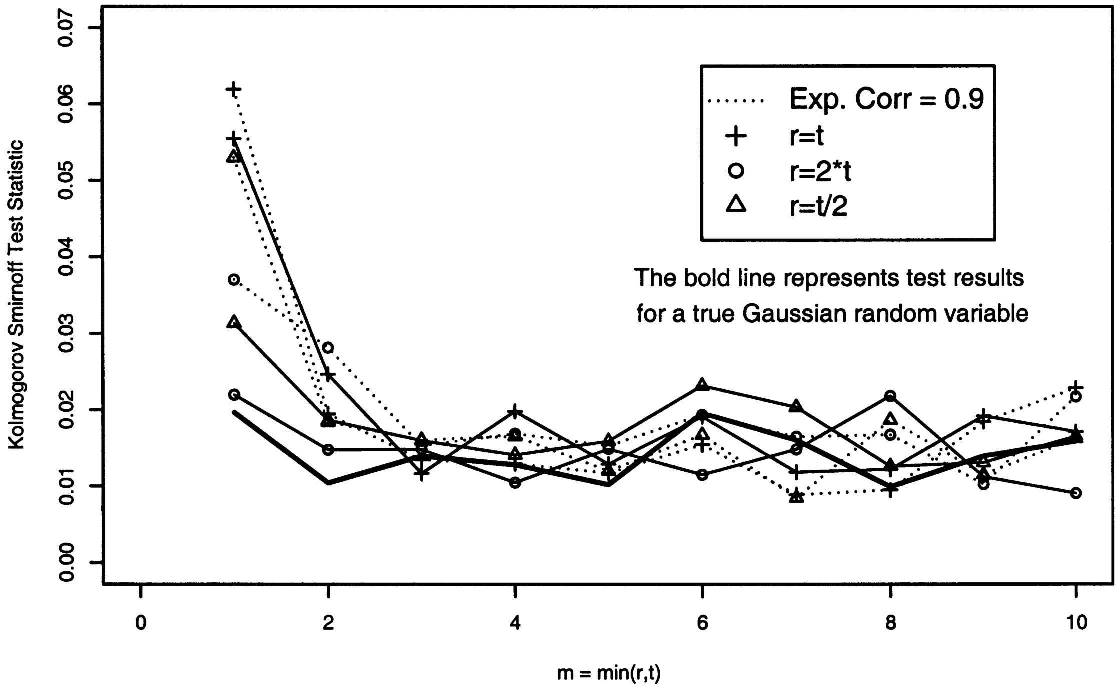

Fig. 5. Gaussian goodness of fit results for MIMO systems at various levels of correlation: $\rho = 0.1$ and $\rho = 0.9$. Antenna numbers are set at $r = t$, $r = t/2$, and $r = 2 \times t$ (SNR = 3 dB).

where

$$\Delta_2(\boldsymbol{x}_o) = \begin{vmatrix} 1 & x_{(1)} & . & . & x_{(1)}^{M-1} \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & x_{(M)} & . & . & x_{(M)}^{M-1} \end{vmatrix} \tag{26}$$

$$\Delta_3 = \begin{vmatrix} 1 & \gamma_1 & . & . & \gamma_1^{N-1} \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & \gamma_N & . & . & \gamma_N^{N-1} \end{vmatrix} \tag{27}$$

and in (28) at the bottom of the page. Now the *unordered* eigenvalues $\boldsymbol{X}_u$ have a joint pdf given by

$$f_{\boldsymbol{X}_u}(\boldsymbol{x}) = \frac{f_{\boldsymbol{X}_o}(\boldsymbol{x})}{M!} = \frac{1}{\prod_{k=1}^{M} k!} \frac{\Delta_1(\boldsymbol{x})\Delta_2(\boldsymbol{x})}{\Delta_3}, \quad x_1 \geq 0, \; x_2 \geq 0, \ldots, x_M \geq 0. \tag{29}$$

Using the unordered eigenvalues we have

$$\begin{aligned} \phi_1(t) &= E\left(\prod_{i=1}^{M}(1+x_i)^{jt}\right) \\ &= \int_0^{\infty} \cdots \int_0^{\infty} \prod_{i=1}^{M}(1+x_i)^{jt}\, \Delta_1\, \Delta_2\, \Delta_3^{-1} \left(\prod_{k=1}^{M} k!\right)^{-1} dx_1 \ldots dx_M. \end{aligned} \tag{30}$$

The above can be simplified by noting that

$$\prod_{i=1}^{M}(1+x_i)^{jt}\, \Delta_1 = |\Delta_{12}\; \boldsymbol{h}_1\; \boldsymbol{h}_2 \ldots \boldsymbol{h}_M|$$

where Δ_{12} is an $N \times N-M$ matrix consisting of the first $N-M$ columns of Δ_1 and the N-vectors

$$\boldsymbol{h}_k = (1+x_k)^{jt} \left[\gamma_1^{N-M-1} e^{-\frac{x_k}{\gamma_1}}, \ldots, \gamma_N^{N-M-1}\, e^{-\frac{x_k}{\gamma_N}}\right]^T.$$

Now it can be seen from the expression for Δ_2 that

$$\Delta_2 = \sum (-1)^{\alpha} \prod_{k=1}^{M} x_k^{i_k}$$

where $i_k \in 0, 1, \ldots M-1$ and α is the sign of the permutation. Using this in (30) gives

$$\begin{aligned} \phi_1(t) &= \Delta_3^{-1} \left(\prod_{i=1}^{M} k!\right)^{-1} \int_0^{\infty} \cdots \\ &\quad \int_0^{\infty} \sum (-1)^{\alpha} \left|\Delta_{12} x_1^{i_1} \boldsymbol{h}_1, \ldots, x_M^{i_M}\, \boldsymbol{h}_M\right| dx_1, \ldots, dx_M \\ &= \Delta_3^{-1} \left(\prod_{k=1}^{M} k!\right)^{-1} \sum (-1)^{\alpha} |\Delta_{12} I_{i_1}, \ldots, I_{i_M}| \end{aligned} \tag{31}$$

$$\Delta_1(\boldsymbol{x}_o) = \begin{vmatrix} 1 & \gamma_1 & . & . & \gamma_1^{N-M-1} & \gamma_1^{N-M-1} e^{-x_{(1)}/\gamma_1} & . & . & \gamma_1^{N-M-1}\, e^{-x_{(M)}/\gamma_1} \\ . & . & . & . & . & & & & \\ . & . & . & . & . & & & & \\ 1 & \gamma_N & . & . & \gamma_N^{N-M-1} & \gamma_N^{N-M-1} e^{-x_{(1)}/\gamma_N} & . & . & \gamma_N^{N-M-1}\, e^{-x_{(M)}/\gamma_N} \end{vmatrix}. \tag{28}$$

where

$$I_{i_k} = \int_0^\infty x^{i_k} \left(\gamma_1^{(N-M-1)} e^{-\frac{x}{\gamma_1}} \cdots \gamma_N^{(N-M-1)} e^{-\frac{x}{\gamma_N}}\right)^T (1+x)^{jt} dx \quad (32)$$

Since α is the number of column swaps required in the ordering $i_1 \dots i_M$, all the terms in the sum above are identical and hence

$$\phi_1(t) = \Delta_3^{-1} \left(\prod_{k=1}^{M-1} k!\right)^{-1} |\Delta_{12} I_0 \dots I_{M-1}|. \quad (33)$$

APPENDIX B
DERIVATION OF $\phi_1'(0)$ AND $\phi_1''(0)$

Here we give a brief derivation of the first and second derivatives of the characteristic function which enables the computation of the mean and variance. Defining

$$c = \Delta_3^{-1} \left(\prod_{k=1}^{M-1} k!\right)^{-1}$$

and differentiating (33) with respect to t gives

$$\phi_1'(t) = c\left|\Delta_{12} I_0' \dots I_{M-1}\right| + \dots + c\left|\Delta_{12} I_0 \dots I_{M-1}'\right| \quad (34)$$

With I_i defined in (32) it is straightforward to show that

$$I_i(0) = i![\gamma_1^{N-M+i}, \dots, \gamma_N^{N-M+i}]^T$$

and hence,

$$\phi_1'(0) = c\left(\prod_{k=1}^{M-1} k!\right) \sum_{i=0}^{M-1} \frac{1}{i!} \left|\Delta_{1,\,N-M+i}\; I_i'(0)\; \Delta_{N-M+i+2,\,N}\right| \quad (35)$$

where $\Delta_{r,\,s}$ represents the submatrix of Δ_3 containing columns r through s inclusive. Expanding the determinant in (35) by column $N - M + i + 1$ gives

$$\phi_1'(0) = c\left(\prod_{k=1}^{M-1} k!\right) \sum_{i=0}^{M-1} \frac{1}{i!} \sum_{k=1}^{N} \alpha_{ik} I_{ik}'(0) \quad (36)$$

where α_{ik} is the ikth cofactor of Δ_3. From (32) we have $I_{ik}'(0) = j\gamma_k^{N-M-1} I(i, \gamma_k)$ and from [24] we have

$$\alpha_{ik} = (-1)^{N-M+1+i+k} \prod_{r>s}^{r,\,s\neq k} (\gamma_r - \gamma_s) \cdot \sum_{1\le i_1<\dots<i_{M-i-1}\le N}^{i_r \neq k} \gamma_{i_1}\gamma_{i_2}\dots\gamma_{i_{M-i-1}}. \quad (37)$$

Substituting for $I_{ik}'(0)$ and α_{ik} in (36) gives the result in (13).

The procedure for the second derivative is very similar and starts from the equation

$$\phi_1''(0) = c\sum_{i=0}^{M-1} \left|\Delta_{12} I_0(0), \dots, I_i''(0), \dots, I_{M-1}(0)\right| + 2c \sum_{0\le i<k\le M-1} \left|\Delta_{12} I_0(0), \dots, I_i'(0), \dots, I_k'(0), \dots, I_{M-1}(0)\right|. \quad (38)$$

We expand both determinants in (38) by column $N-M+i+1$. The second determinant is then also expanded by the column containing $I_k'(0)$. This gives

$$\phi_1''(0) = c\sum_{i=0}^{M-} \sum_{k=1}^{N} \beta_{ik} I_{ik}''(0) + 2c \sum_{0\le i<k\le M-1} \sum_{r=1}^{N} \sum_{s=1}^{N(s\neq r)} \delta_{ikrs} I_{ir}'(0) I_{ks}'(0) \quad (39)$$

where β_{ik} and δ_{ikrs} are the relevant cofactors in (38). Note that computationally fast and simple recursions can be derived for all the constants required (c_{ik}, β_{ik}, and δ_{ikrs}) but this is beyond the scope of this correspondence.

ACKNOWLEDGMENT

The authors wish to thank Andrea Giorgetti for his help in providing the parameters for the correlation scenarios.

REFERENCES

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communication in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.
[2] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multielement antennas," *Bell Labs Tech. J.*, vol. 1, pp. 41–59, Autumn 1996.
[3] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 139–157, Jan. 1999.
[4] J. S. J. Winters and R. D. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. Commun.*, vol. 42, pp. 1740–1750, Feb./Mar/Apr. 1994.
[5] E. Telatar, "Capacity of multi-antenna gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, pp. 2172–2178, 2000.
[6] J. B. Anderson, "Array gain and capacity for known channels with multiple element arrays at both ends," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 2172–2178, Nov. 2000.
[7] J. Evans and D. N. C. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2059–2078, Sept. 2000.
[8] D. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multi-element antenna systems," *IEEE Trans. Commun.*, vol. 48, pp. 502–513, Mar. 2000.
[9] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multi-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inform. Theory*, submitted for publication.
[10] C.-N. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inform. Theory*, vol. 48, pp. 637–650, Mar. 2002.
[11] J. L. D. Chizhik, F. Rashid-Rarrokhi, and A. Lozano, "Effect of antenna separation on the capacity of blast in correlated channels," *IEEE Commun. Lett.*, vol. 4, pp. 337–339, Nov. 2000.
[12] A. M. Sengupta and P. P. Mitra, "Capacity of multivariate channels with multiplicative noise: 1. random matrix techniques and large-n expansions for full transfer matrices," LANL arXiv: physics/0010087, 2000.
[13] M. T. Ivrlac, W. Utschick, and J. A. Nossek, "Fading correlations in wireless mimo communication systems," *IEEE J. Select. Aareas Commun.*, vol. 21, pp. 819–828, June 2003.
[14] J. Laurila, K. Kalliiola, M. Toeltsch, K. Hugel, P. Vainikainen, and E. Bonek, "Wideband 3-d characterization of mobile radio channels in urban environments," *IEEE Trans. Antennas Propagat.*, vol. 50, pp. 233–243, Feb. 2002.
[15] P J. Smith and M. Shafi, "On a Gaussian approximation to the capacity of wireless MIMO systems," in *Proc. 2002 IEEE Int. Conf. Communications*, New York, Apr. 28–May 2 2002, pp. 406–410.
[16] M. Chiani, "Evaluating the capacity distribution of mimo rayleigh fading channels," in *Proc. IEEE Int. Symp. Advances in Wireless Communications*, Vancouver, BC, Canada, 2002, pp. 3–4.
[17] A. Abdi and M. Kaveh, "A space-time correlation model for multielement antenna systems in mobile fading channels," *IEEE J. Select. Areas Commun.*, vol. 20, pp. 550–560, Apr. 2002.
[18] 3GPP, "A standardized set of mimo radio propagation channels," in *TSGR#23 RI-01-1179*, Jeju, Korea, Nov. 2001.
[19] E. Biglieri and G Tarico, "Large-system analysis of multiple-antenna system capacities," *J. Commun. and Networks*, vol. 5, no. 2, pp. 96–103, June 2003.
[20] H. Gao and P. J. Smith, "A determinant representation for the distribution of quadratic forms in complex normal vectors," *J. Mult. Anal.*, vol. 73, no. 2, pp. 155–165, 2000.
[21] I. Stegun and M. Abramozitz, Eds., *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. Washington, DC: Nat. Bur. Stand., 1972.

[22] A. Giorgetti, M. Chiani, M. Shafi, and P. J. Smith, "Level crossing rates and mimo capacity fades: Impacts of spatial/temporal channel correlation," in *Proc. IEEE Int. Conf. Communications*, Anchorage, AK, 2003, pp. 3046–3050.
[23] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Ann. Math Statist.*, vol. 35, pp. 475–501, 1964.
[24] H. Gao and P. J. Smith, "Exact sinr calculations for optimum linear combining in wireless systems," *Probl. Eng. Inform. Sci.*, vol. 12, pp. 261–281, 1998.

Upper Bounds of Rates of Complex Orthogonal Space–Time Block Codes

Haiquan Wang and Xiang-Gen Xia, *Senior Member, IEEE*

***Abstract*—In this correspondence, we derive some upper bounds of the rates of (generalized) complex orthogonal space–time block codes. We first present some new properties of complex orthogonal designs and then show that the rates of complex orthogonal space–time block codes for more than two transmit antennas are upper-bounded by $3/4$. We show that the rates of generalized complex orthogonal space–time block codes for more than two transmit antennas are upper-bounded by $4/5$, where the norms of column vectors may not be necessarily the same. We also present another upper bound under a certain condition.**

For a (generalized) complex orthogonal design, its variables are not restricted to any alphabet sets but are on the whole complex plane. In this correspondence, a (generalized) complex orthogonal design with variables over some alphabet sets on the complex plane is also considered. We obtain a condition on the alphabet sets such that a (generalized) complex orthogonal design with variables over these alphabet sets is also a conventional (generalized) complex orthogonal design and, therefore, the above upper bounds on its rate also hold. We show that commonly used quadrature amplitude modulation (QAM) constellations of sizes above 4 satisfy this condition.

***Index Terms*—Complex orthogonal designs, complex orthogonal space–time block codes, Hermitian compositions of quadratic forms, Hurwitz family, Hurwitz–Radon theory.**

I. Introduction

The first real/complex orthogonal space–time block code was proposed by Alamouti [1] for two transmit antennas. It was then generalized to real/complex orthogonal space–time block codes for more than two transmit antennas by Tarokh, Jafarkhani, and Calderbank [3]. There are two important properties of real/complex orthogonal space–time block codes: 1) they have fast maximum-likelihood (ML) decoding, namely, symbol-by-symbol decoding; 2) they have the full diversity. These two properties make real/complex orthogonal space–time block codes attractive in space–time code designs. By utilizing the Hurwitz–Radon theory [17]–[19], [23], [26], Tarokh, Jafarkhani, and Calderbank [3] provided a systematic method to construct *real* orthogonal space–time block codes of size $p \times n$ and rate 1 for k pulse-amplitude modulation (PAM) symbols, where n is the number of transmit antennas, p is the time delay (or block size), and $R = k/p$ is the code rate. They also provided a construction of rate $1/2$ *complex* orthogonal space–time block codes for phase-shift keying (PSK) and quadrature amplitude modulation (QAM) symbols using real orthogonal space–time block codes of rate 1. In order to maintain the fast ML decoding and the full diversity of a space–time block code, the orthonormality in the sense that the norms of all column vectors are the same can be relaxed to a general orthogonality where the norms of column vectors may not be necessarily the same [3]. A complex orthogonal space–time block code with the generalized orthonormality is called a generalized complex orthogonal space–time block code. In [2], [3], it has been shown that the rate $R \leq 1$ for both real and complex orthogonal space–time block codes for any number of transmit antennas. While the maximal rate 1, i.e., $R = 1$, is reachable for real orthogonal space–time block codes as we previously mentioned from the Hurwitz–Radon's constructive theory, it has been recently shown in [8] that $k \leq p - 1$ when $n > 2$, i.e., $R < 1$ and $R = 1$ is not reachable for (generalized) complex orthogonal space–time block codes no matter what the time delay p is unless the number of transmit antennas is two, i.e., the Alamouti's scheme. Notice that, if condition $p = n$ is required, i.e., *square* codes or *square* complex orthogonal designs, then $R < 1$ when $n > 2$ directly follows from the results on amicable designs [18], [21]–[23], [3], [5]–[7] that have small rates when $n \geq 8$. While both square and nonsquare *real* orthogonal designs (or compositions of quadratic forms) are well understood, not much is known for nonsquare *complex* orthogonal designs (or Hermitian compositions of quadratic forms [26]), [3], [26], [27].

In this correspondence, we derive some upper bounds on the rates R of (generalized) complex orthogonal space–time block codes (or complex orthogonal designs). We emphasize that the sizes of (generalized) complex orthogonal space–time block codes (or complex orthogonal designs) here are general and they may not be square, i.e., p may not be equal to n. We show that, when the number of transmit antennas is more than two, i.e., $n > 2$, the rates of complex orthogonal space–time block codes are upper-bounded by $3/4$, i.e.,

$$R \leq \frac{3}{4}$$

and the rates of generalized complex orthogonal space–time block codes are upper-bounded by $4/5$, i.e.,

$$R \leq \frac{4}{5}.$$

Note that rate–$3/4$ complex orthogonal space–time block codes for three and four transmit antennas have appeared in [3]–[6]. Therefore, the above upper bound tells us that these complex orthogonal space–time block codes have already reached the optimal rate. Also note that the above upper bound $3/4$ on the rates is not new for *square* complex orthogonal designs. In fact, it has been shown and reviewed from amicable designs in [18], [21]–[23], [3], [5]–[7]. However, this upper bound is *new* for nonsquare complex orthogonal designs. In the meantime, it is known that to generate orthogonal space–time codes, a square orthogonal design is not necessary [3].

In a conventional (generalized) complex orthogonal design, its variables may take any values in the complex plane. However, as we shall see later, to generate a space–time code, the variables only take values in some finite subsets, called alphabet sets, on the complex plane. The question then becomes whether it is helpful to produce more (generalized) complex orthogonal designs of high rates when their variables are restricted to some alphabet sets. This question has been partially

Manuscript received September 2, 2002; revised April 27, 2003. This work was supported in part by the Air Force Office of Scientific Research under Grant F49620-02-1-0157 and the National Science Foundation under Grants MIP-9703377, CCR-0097240, and CCR-0325180. The material in this correspondence was presented in part at the IEEE International Symposium on Information Theory, Lausanne, Switzerland, June/July 2002.

The authors are with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716 USA (e-mail: hwang@ee.udel.edu; xxia@ee.udel.edu).

Communicated by T. L. Marzetta, Guest Editor.
Digital Object Identifier 10.1109/TIT.2003.817830