Characterizing the Statistical Properties of Mutual Information in MIMO Channels: Insights into Diversity-Multiplexing Tradeoff

Özgür Oyman[1)], Rohit U. Nabar[2)], Helmut Bölcskei[2)], and Arogyaswami J. Paulraj[1)]

[1)]Information Systems Laboratory, Stanford University
228 Packard, 350 Serra Mall, Stanford, CA 94305
Email: {oyman,apaulraj}@stanford.edu

[2)] Communication Technology Laboratory, ETH Zürich
ETF E122, Sternwartstrasse 7, 8092 Zürich, Switzerland
Email: {nabar,boelcskei}@nari.ee.ethz.ch

Abstract

We consider gaussian multiple-input multiple-output (MIMO) fading channels assuming that the channel is unknown at the transmitter and perfectly known at the receiver. Taking into account spatial fading correlation both at the transmitter and the receiver, we provide tight closed-form lower-bounds for ergodic capacity and derive extremely accurate closed-form approximations of the variance of mutual information over such channels. Based on these results, we investigate the impact of the number of antennas and transmit and receive correlation on ergodic capacity and on the variance of mutual information and draw insights into the tradeoff between diversity gain and spatial multiplexing gain.

I. Introduction

The use of multiple antennas at both the transmitter and the receiver of a wireless system enables significant capacity gains through a technique known as *spatial multiplexing* [1]-[4]. Analytical expressions for the resulting gains are in general very difficult to obtain.

Contributions. In this paper, we extend the results reported in [5] to provide extremely accurate closed-form approximations of the first and second order statistics of the mutual information of multiple-input multiple-output (MIMO) Rayleigh fading channels. We consider the case where the channel is unknown at the transmitter and perfectly known at the receiver, and restrict our analysis to the high signal-to-noise ratio (SNR) regime. Furthermore, we investigate the impact of spatial fading correlation both at the transmitter and the receiver and the number of transmit and receive antennas on the statistics of mutual information. Based on these results, we provide insights into the tradeoff between diversity gain and multiplexing gain in MIMO channels.

Organization of the paper. The rest of this paper is organized as follows. In Sec. II, we introduce the channel model and state our assumptions. In Sec. III, we review some of the results from [5] and derive a lower-bound on ergodic capacity for the case of joint transmit-receive correlation. In Sec. IV, we derive the variance of mutual information of a Rayleigh flat-fading MIMO channel in the high SNR regime. Based on these results, we provide insights into the diversity-multiplexing tradeoff in MIMO channels in Sec. V. In Sec. VI, we present numerical results. We conclude in Sec. VII.

II. The Channel Model

In the following M_T and M_R denote the number of transmit and receive antennas, respectively. We restrict ourselves to a purely Rayleigh flat-fading scenario, where the elements of the $M_R \times M_T$ channel matrix $[\mathbf{H}]_{m,n} = h_{m,n}\,(m = 0, 1, ..., M_R - 1, n = 0, 1, ..., M_T - 1)$ are (possibly correlated) circularly symmetric[1] zero mean complex gaussian random variables. We employ the block-fading model used in [6] wherein the channel remains constant over at least M_T symbol periods and

Ö. Oyman's work was supported by the Benchmark Stanford Graduate Fellowship.

[1]A circularly symmetric complex gaussian random variable is a random variable $z = (x + jy) \sim \mathcal{CN}(0, \sigma^2)$, where x and y are i.i.d. $\mathcal{N}(0, \sigma^2/2)$.

then changes in an independent fashion to a new realization. We furthermore assume that spatial fading correlation occurs both at the transmitter and the receiver and that the distance between transmitter and receiver is small or the scattering radii at the transmitter and the receiver are sufficiently large. Under these assumptions the MIMO channel model recently proposed in [7] reduces to

$$\mathbf{H} = \mathbf{R}^{1/2}\mathbf{H}_w\mathbf{S}^{1/2}, \tag{1}$$

where $\mathbf{H}_w$ is an $M_R \times M_T$ i.i.d. complex gaussian matrix with zero mean unit variance entries, and $\mathbf{S} = \mathbf{S}^{1/2}\mathbf{S}^{1/2}$ and $\mathbf{R} = \mathbf{R}^{1/2}\mathbf{R}^{1/2}$ are the transmit and the receive correlation matrices, respectively. We use the eigendecompositions[2] $\mathbf{R} = \mathbf{U_R}\mathbf{\Lambda_R}\mathbf{U_R}^H$ and $\mathbf{S} = \mathbf{U_S}\mathbf{\Lambda_S}\mathbf{U_S}^H$, respectively, where $\mathbf{\Lambda_R} = \mathrm{diag}\{\lambda_i(\mathbf{R})\}_{i=1}^{M_R}$ and $\mathbf{\Lambda_S} = \mathrm{diag}\{\lambda_i(\mathbf{S})\}_{i=1}^{M_T}$. We note that the channel model in (1) has also been used in [8] to analyze the asymptotic (in the number of antennas) capacity behavior of MIMO channels in the presence of correlated fading.

III. Ergodic Capacity Bound

Assuming a circularly symmetric complex gaussian i.i.d. input signal vector, the mutual information of the MIMO system introduced in the previous section is given by[3] [2], [3]

$$I \;=\; \log_2 \det\left(\mathbf{I}_{M_R} + \frac{\rho}{M_T}\mathbf{H}\mathbf{H}^H\right) \;\; \text{bps/Hz}, \tag{2}$$

where ρ is the average SNR at each of the receive antennas. Assuming that the fading process is ergodic, a Shannon capacity or ergodic capacity exists and is given by[4] $C = \mathcal{E}\{I\}$.

The i.i.d. case. For the sake of completeness, we shall next briefly review some of the results reported in [5]. The ergodic capacity of an $M_R \times M_T$ i.i.d. Rayleigh fading MIMO channel can be lower-bounded as

$$C \geq L\log_2\left(1 + \frac{\rho}{M_T}\exp\left(\frac{1}{L}\sum_{j=1}^{L}\sum_{p=1}^{K-j}\frac{1}{p} - \gamma\right)\right), \tag{3}$$

where $L = \min(M_T, M_R)$, $K = \max(M_T, M_R)$, and $\gamma \approx 0.57721566$ is Euler's constant. In the high SNR regime($\rho \gg 1$), the ergodic capacity can be approximated as

$$C \approx L\log_2\left(\frac{\rho}{M_T}\right) + \frac{1}{\ln 2}\left(\sum_{j=1}^{L}\sum_{p=1}^{K-j}\frac{1}{p} - \gamma L\right). \tag{4}$$

Moreover, in the large array limit, where $M_T \to \infty$, $M_R \to \infty$ keeping $\beta = M_T/M_R$ fixed, the lower bound on the ergodic capacity per receive antenna C_{pr} converges to

$$C_{pr} \geq \begin{cases} \beta\log_2\left(1 + \frac{\rho}{\beta}e^{-1}\left(\frac{1}{1-\beta}\right)^{\frac{1}{\beta}-1}\right) & \beta \leq 1 \\ \log_2\left(1 + \rho e^{-1}\left(\frac{\beta}{\beta-1}\right)^{\beta-1}\right) & \beta > 1 \end{cases}, \tag{5}$$

where β is the *system load.* These results are intuitively appealing since they show explicitly that the ergodic capacity grows linearly with $\min(M_T, M_R)$ (linear increase with respect to β for $\beta < 1$). More specifically, C increases by $\min(M_T, M_R)$ for every 3 dB increase in SNR. Thus, the number of spatial data pipes that can be opened up between the transmitter and the receiver is constrained by the minimum of the number of antennas at the transmitter and the receiver. This is a well-known fact first observed in [9] without providing an explicit analytical expression for C. For high loading levels (i.e. $\beta > 1$), the additional degrees of freedom lead to an increase in capacity that is logarithmic in β. As $\beta \to \infty$, we observe that $C_{pr} \to \log_2(1+\rho)$, which corresponds to the orthogonal channel capacity.

The correlated case. In the case of spatial fading correlation only at the receiver ($\mathbf{S} = \mathbf{I}_{M_T}$), assuming that $r(\mathbf{R}) \leq M_T$, it can be shown that the ergodic capacity for the channel described by (1) is lower-bounded as

$$\begin{aligned} C \;\geq\; & r(\mathbf{R})\log_2\Bigg(1 + \frac{\rho}{M_T}\left(\det\left(\mathbf{\Lambda_R}\right)\right)^{1/r(\mathbf{R})} \\ & \exp\left(\frac{1}{r(\mathbf{R})}\sum_{j=1}^{r(\mathbf{R})}\sum_{p=1}^{M_T-j}\frac{1}{p} - \gamma\right)\Bigg). \end{aligned} \tag{6}$$

Similarly, in the case of spatial fading correlation only at the transmitter ($\mathbf{R} = \mathbf{I}_{M_R}$), assuming that $r(\mathbf{S}) \leq M_R$, we obtain

$$\begin{aligned} C \;\geq\; & r(\mathbf{S})\log_2\Bigg(1 + \frac{\rho}{M_T}\left(\det\left(\mathbf{\Lambda_S}\right)\right)^{1/r(\mathbf{S})} \\ & \exp\left(\frac{1}{r(\mathbf{S})}\sum_{j=1}^{r(\mathbf{S})}\sum_{p=1}^{M_R-j}\frac{1}{p} - \gamma\right)\Bigg). \end{aligned} \tag{7}$$

[2] The superscript ^{H} stands for conjugate transposition. $\lambda_i(\mathbf{R})$ is the i-th nonzero eigenvalue of $\mathbf{R}$. $r(\mathbf{R})$ denotes the rank of $\mathbf{R}$.

[3] $\mathbf{I}_m$ denotes the $m \times m$ identity matrix.

[4] $\mathcal{E}$ stands for the expectation operator.

In the case of spatial fading correlation both at the transmitter and the receiver, assuming that $\mathbf{S}$ and $\mathbf{R}$ have equal rank with $r(\mathbf{S}) = r(\mathbf{R}) = N$, we obtain

$$\begin{aligned} C \quad \geq \quad & N \log_2 \Big(1 + \frac{\rho}{N} \left(\det\left(\mathbf{\Lambda_S}\right) \det\left(\mathbf{\Lambda_S}\right)\right)^{1/N} \\ & \exp\left(\frac{1}{N} \sum_{j=1}^{N} \sum_{p=1}^{N-j} \frac{1}{p} - \gamma \right) \Big) . \end{aligned} \tag{8}$$

From (6)-(8), it is clear that there is a loss in ergodic capacity due to spatial fading correlation. For the MIMO channel with joint transmit/receive correlation, at high SNR, assuming that $\mathbf{S}$ and $\mathbf{R}$ have full rank, this loss is quantified by $\log_2(\det(\mathbf{\Lambda_S})) + \log_2(\det(\mathbf{\Lambda_R}))$ bps/Hz. This loss is $\log_2(\det(\mathbf{\Lambda_S}))$ and $\log_2(\det(\mathbf{\Lambda_R}))$ bps/Hz, respectively, for the cases of transmit only and receive only correlation. Furthermore, we note that the number of spatial data pipes opened up between transmitter and receiver is constrained by the rank of the correlation matrices. Numerical results (obtained through Monte Carlo methods) in Sec. VI reveal (6)-(8) to be tight lower-bounds on ergodic capacity at any SNR and extremely accurate approximations in the high SNR regime.

IV. Variance of Mutual Information of MIMO Channels

The i.i.d case. In the high SNR case, the variance of mutual information can be approximated as[5]

$$\sigma_I^2 \approx \operatorname{var}\left\{ \log_2 \det\left(\frac{\rho}{M_T} \mathbf{H}\mathbf{H}^H \right) \right\}. \tag{9}$$

From [10], we can infer that

$$\sigma_I^2 \approx \left(\frac{1}{\ln 2} \right)^2 \sum_{j=1}^{L} \operatorname{var}(\ln(Y_j)), \tag{10}$$

where Y_j $(j = 1, 2, ..., L)$ are independent chi-square distributed random variables with $2(K - j + 1)$ degrees of freedom and K and L are as defined in (3). Using results from [11] and [12], a general closed-form expression for σ_I^2 is obtained as

$$\sigma_I^2 \approx \left(\frac{1}{\ln 2} \right)^2 \sum_{j=1}^{L} \sum_{p=1}^{\infty} \frac{1}{(p + K - j)^2}. \tag{11}$$

Comparing (4) and (11), we can see that unlike ergodic capacity, the variance of I displays symmetry with respect to M_T and M_R in the sense that σ_I^2 for an $m \times n$ system equals σ_I^2 for an $n \times m$ system. Simulations in Sec. 5 reveal (11) to be an accurate approximation in the high SNR regime.

The correlated case. In the presence of spatial fading correlation, assuming that transmit and receive correlation matrices $\mathbf{S}$ and $\mathbf{R}$ have equal rank with $r(\mathbf{S}) = r(\mathbf{R}) = N$, the variance of mutual information at high SNR may be approximated as

$$\sigma_I^2 \approx \left(\frac{\det\left(\mathbf{\Lambda_S}\right) \det\left(\mathbf{\Lambda_R}\right)}{\ln 2} \right)^2 \sum_{j=1}^{N} \sum_{p=1}^{\infty} \frac{1}{(p + N - j)^2}. \tag{12}$$

From (12), we infer that the variance of mutual information decreases by a factor of $(\det(\mathbf{\Lambda_S}\mathbf{\Lambda_R}))^2$ compared to that of the i.i.d. channel.

V. Insights into the Tradeoff between Diversity and Spatial Multiplexing

MIMO systems offer two different benefits: 1) spatial diversity gain to combat channel fading, and 2) an increase in spectral efficiency through spatial multiplexing. These benefits are in general conflicting. It is therefore necessary to understand the tradeoff between multiplexing gain and diversity gain in designing MIMO systems with certain performance requirements. The variance analysis in the previous section can provide useful insights into this tradeoff.

We shall first relate the variance of mutual information to the diversity order provided by the channel. To see this note that the pdf of mutual information approaches a Dirac function as the number of degrees of freedom approaches infinity (and the fading channel approaches an AWGN channel) by letting the number of antennas on one side (transmit or receive) of the link increase constantly keeping the number of antennas on the opposite side fixed. Hence, the variance of I, $\sigma_I^2 = \mathcal{E}\{I^2\} - C^2$, is inversely proportional to the number of degrees of freedom provided by the channel and can therefore be seen as a measure of the diversity gain supported by the channel. This statement has been made rigorous in [13], [14]. In particular, applying Chebyshev's inequality, we have

$$P\left(|I - C| \geq A\right) \leq \frac{\sigma_I^2}{A^2}.$$

Therefore, σ_I^2 can be seen as a measure of the diversity gain supported by the MIMO channel. The smaller the variance the higher the diversity order and vice versa. We begin by considering the case of an i.i.d. channel.

[5] $\operatorname{var}(X)$ stands for the variance of the random variable X.

The i.i.d. case. In this subsection, we shall show that σ_I^2 is maximized if $L = K$ and hence $M_T = M_R$. Starting from (11), fixing K and considering L variable (to reflect this we use the notation $\sigma_I^2(L)$) we obtain

$$\sigma_I^2(L) = \sigma_I^2(L-1) + \left(\frac{1}{\ln 2}\right)^2 \sum_{p=1}^{\infty} \frac{1}{(p+K-L)^2}.$$

This implies that $\sigma_I^2(L)$ is a strictly increasing function of L which is maximum for $L = K$ and hence $M_T = M_R$. Similarly, if we fix L in (11), we observe that σ_I^2 strictly decreases with increasing K, once again establishing that the variance of mutual information is maximum when $M_T = M_R$. This result is somewhat surprising since it says that fixing the number of receive antennas $M_R > M_T$ and increasing M_T leads to an increase in the variance of I and hence reduced diversity order. (Recall that the diversity order is inversely proportional to the variance of mutual information). There is, however, a physically appealing interpretation of this phenomenon. As we increase the number of transmit antennas (for fixed $M_R > M_T$) we also increase the number of parallel spatial data pipes, i.e., the number of independent data streams that are spatially multiplexed. Thus, the additional degrees of freedom (obtained by increasing the number of transmit antennas) are exploited to increase the throughput by multiplexing a higher number of independent data streams rather than exploiting them to increase the diversity order. Once the number of transmit antennas becomes larger than the number of receive antennas the number of parallel spatial data pipes that can be opened up is constrained by the number of receive antennas (cf. (4)) and the transmit antennas in excess of this number are used to increase the diversity order experienced by the independent data streams thus reducing σ_I^2. This diversity gain amounts to the logarithmic increase in ergodic capacity in the high loading regime. We have thus exhibited a fundamental tradeoff between multiplexing gain and diversity gain. We emphasize that our conclusions are a consequence of the very nature of the transmit signal vector's statistics (i.i.d. complex gaussian) which implies a pure multiplexing mode. We finally note that the diversity-multiplexing tradeoff in i.i.d. MIMO channels has been investigated in [15] using a different framework.

The correlated case. In Sec. III, we observed a loss in ergodic capacity due to spatial fading correlation. Interestingly however, in the presence of spatial fading correlation, assuming that $r(\mathbf{S}) = r(\mathbf{R}) = N$, in the high SNR regime, we concluded that the variance of mutual information decreases compared to that of the i.i.d. channel. Again, this observation has a physically appealing interpretation. Assume for example that $M_T = M_R$ and that only transmit correlation is present. High spatial fading correlation amounts to a reduced number of effective antennas. As the spatial fading correlation decreases the effective number of transmit antennas increases which means that the effective number of parallel spatial data pipes that can be opened up is increased. Thus, reducing the spatial fading correlation is like increasing the number of transmit antennas. Now, from the previous discussion we know that σ_I^2 increases until the effective number of transmit antennas equals the effective number of receive antennas, which in our scenario amounts to fully uncorrelated spatial fading.

VI. Numerical Results

In this section, we demonstrate the accuracy of our analytical expressions derived in earlier sections.

A. Variance of Mutual Information

In Fig. 1, we compare the empirically obtained variance of I (through Monte Carlo methods) and its analytical estimate (11) (summed over p ranging from 1 to 500) for a system with $M_R = 10$ and varying M_T at high SNR. The empirical and analytical results are in close agreement verifying the accuracy of our analysis. We can see that the results depicted in Fig. 1 are consistent with the results derived in the previous section, i.e., σ_I^2 is maximum for $M_T = M_R$.

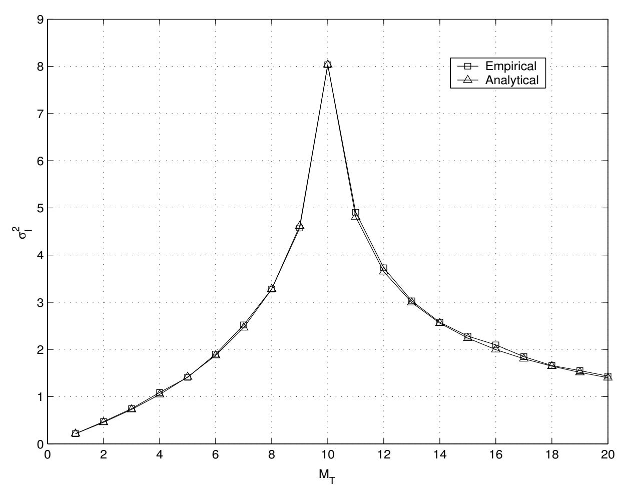

Fig. 1. Comparison of empirically determined σ_I^2 and the approximation (11) for various values of M_T with fixed $M_R = 10$ at high SNR.

B. Ergodic Capacity

In this example, we investigate the ergodic capacity loss due to spatial fading correlation for a Rayleigh flat-fading MIMO channel with $M_T = M_R = 2$. We use the channel model specified in (1) with

$$\mathbf{S} = \begin{bmatrix} 1 & s \\ s^* & 1 \end{bmatrix} \qquad \text{and} \qquad \mathbf{R} = \begin{bmatrix} 1 & r \\ r^* & 1 \end{bmatrix}, \tag{13}$$

where s and r are the complex correlation coefficients between the two transmit and the two receive antennas, respectively. In Fig. 2, we compare the lower bound (8) with the empirically obtained ergodic capacity (through Monte Carlo methods) for three different levels of transmit and receive correlation, namely $s = r = 0$ (i.i.d. channel), $s = r = 0.4$ (low transmit/receive correlation), and $s = 0.95, r = 0.4$ (high transmit, low receive correlation). As predicted by the analytical estimate $\log_2(\det(\mathbf{S})) + \log_2(\det(\mathbf{R}))$, we observe a very small ergodic capacity loss for the case of low transmit/receive correlation. In the case of high correlation at the transmitter, we observe an ergodic capacity loss of about 3.6 bps/Hz, which is consistent with the loss of 3.61 bps/Hz, predicted by the analytical estimate.

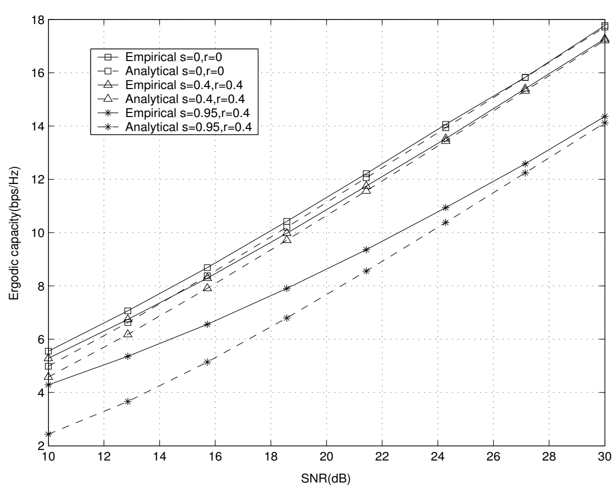

Fig. 2. Comparison of the empirically determined ergodic capacity and the analytical lower-bound (8) for $M_T = M_R = 2$ and various levels of spatial fading correlation.

VII. Conclusions

In this paper, we analyzed the first and second-order statistics of mutual information of a MIMO Rayleigh flat-fading channel with joint transmit-receive correlation. We provided a tight closed-form analytical lower-bound for ergodic capacity, derived an analytical expression for the variance of mutual information, and revealed an interesting tradeoff between multiplexing gain and diversity gain. Finally, we demonstrated the accuracy of our analytical expressions through comparison with numerical results.

References

[1] A. J. Paulraj and T. Kailath, "Increasing capacity in wireless broadcast systems using distributed transmission/directional reception," *U. S. Patent*, no. *5,345,599*, 1994.

[2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecomm.*, vol. 10, no. 6, pp. 585–595, Nov.-Dec. 1999.

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

[4] H. Bölcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Comm.*, vol. 50, no. 2, pp. 225–234, Feb. 2002.

[5] Ö. Oyman, R. U. Nabar, H. Bölcskei, and A. J. Paulraj, "Tight lower bounds on the ergodic capacity of Rayleigh fading MIMO channels," in *Proc. IEEE GLOBECOM 2002*, Taipei, Taiwan, to appear.

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

[7] D. Gesbert, H. Bölcskei, D. Gore, and A. J. Paulraj, "MIMO wireless channels: Capacity and performance prediction," in *Proc. IEEE GLOBECOM*, San Francisco, CA, vol. 2, pp. 1083–1088, Nov. 2000.

[8] C. Chuah, D. Tse, J. Kahn, and R. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 637–650, March 2002.

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

[10] N. R. Goodman, "The distribution of the determinant of a complex Wishart distributed matrix," *Ann. Math. Stat.*, vol. 34, no. 1, pp. 178–180, March 1963.

[11] P. M. Lee, *Bayesian Statistics: An Introduction*, Arnold/John Wiley, London/New York, 2nd edition, 1997.

[12] X. Gourdon and P. Sebah, "The Gamma function," available online at http://numbers.computation.free.fr, Sept. 2000.

[13] R. U. Nabar, H. Bölcskei, and A. J. Paulraj, "Outage performance of space-time block codes for generalized MIMO channels," *IEEE Trans. Inf. Theory*, submitted, March 2002.

[14] R. U. Nabar, H. Bölcskei, and A. J. Paulraj, "Influence of propagation conditions on the outage capacity of space-time block codes," in *Proc. EWC 2002*, Florence, Italy, Feb. 2002, pp. 629–634.

[15] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Trans. Inf. Theory*, 2002, to appear.