An Upper Bound on the Ergodic Mutual Information in Rician Fading MIMO Channels

Jari Salo, Filip Mikas, and Pertti Vainikainen

***Abstract*— We consider Rician fading Multiple-Input Multiple-Output (MIMO) channels where the transmitted signal has complex Gaussian distribution, iid across the transmit antennas. Based on expected values of elementary symmetric functions of complex noncentral Wishart matrices, we derive an upper bound on the average (ergodic) mutual information for arbitrary SNR, arbitrary rank of the deterministic line-of-sight matrix, and arbitrary number of transmit/receive antennas. The Rayleigh fading signal component is allowed to have spatial correlation at one end of the link. Upper bounds for the cases of rank-1 line-of-sight component and pure Rayleigh fading emerge as special instances of the general result.**

***Index Terms*— MIMO, Rician fading, mutual information, elementary symmetric functions, noncentral Wishart.**

I. Introduction

We consider Multiple-Input Multiple-Output radio channels with n_t transmit and n_r receive antennas. For a given channel input signal the highest achievable information rate is given by mutual information, which, for a random channel, is itself a random variable. Throughout this paper we assume that the channel input is isotropically distributed Gaussian signal. One is typically interested finding the distribution or the mean value of the mutual information. The distribution is useful for determining the maximum achievable information rate (for a given outage probability) of a communication system employing signal encoding over a single channel realization. The mean (ergodic) mutual information gives the maximum achievable information rate when encoding over a large number of channel realizations.

For Rayleigh channels the computation of exact or approximate distribution of mutual information has been considered in e.g. [1], [2], [3], while various expressions for the mean mutual information can be found in [4], [5], [6], [7]. However, the more general case of Rician fading has not been as thoroughly analyzed. The outage properties of mutual information from antenna design point of view have been studied by simulations in [8] and [9]. In [10] lower and upper bounds for the mean of the mutual information are derived, [11] provides a high-SNR approximation for its density, while in [12] the exact density for dual MIMO systems is derived; all these papers concentrate on Rician MIMO channels with rank-1 deterministic component. It should also be mentioned that all the aforementioned references, [8], [9], [10], [11], [12], assume a Gaussian isotropic channel input signal, which does not, in general, achieve channel capacity; the capacity-achieving input distribution for Rician MIMO channels, assuming that the transmitter knows the channel probability distribution, is characterized in [13] and [14], see also [15][16]. The water-filling capacity, which assumes that the transmitter knows the instantaneous channel matrix, has been considered in [4][17]. To our knowledge, there are no studies of ergodic mutual information addressing the case of Rician fading with arbitrary-rank deterministic component and spatially correlated stochastic (Rayleigh) component. Added in proof: Recent results that have appeared after the acceptance of this paper include [18] and [19].

In this paper we derive a closed-form analytical upper bound for the mutual information of Rician fading MIMO channels whose random (non-line-of-sight) signal component is semi-correlated[1] Rayleigh fading, and the deterministic component has arbitrary rank. The bounds are asymptotically tight at low SNR's and their difference to the exact mutual information tends to a constant at high SNR's.

The paper is organized as follows. In Section II we give the system model. The main results are in Section III. In Section IV the tightness and computational complexity of the bound are discussed. Numerical examples are given in Section V. Section VI concludes the paper. Some derivations can be found in the appendix.

II. System Model

Throughout the paper we shall denote determinant, trace, Frobenius norm, conjugate transpose, and rank of a matrix $\mathbf{A}$ with $|\mathbf{A}|$, $\mathrm{tr}(\mathbf{A})$, $\|\mathbf{A}\|_F$, $\mathbf{A}^H$, and $\mathrm{rank}(\mathbf{A})$, respectively. The $n \times n$ identity matrix is denoted by $\mathbf{I}_n$. We shall also use the shorthand notations $K = \min(n_r, n_t)$ and $L = \max(n_r, n_t)$. Expected value of a random variable X is denoted by $E[X]$.

Assume that the transmitted signal has complex circularly symmetric zero-mean multivariate Gaussian distribution with correlation matrix $\frac{P}{n_t}\mathbf{I}_{n_t}$ (isotropic complex Gaussian channel input), where P is the total transmitted signal power. Then, if the receiver knows the given $n_r \times n_t$ channel matrix

[1]By "semicorrelated" we mean that correlation is allowed only at one end of the link [2].

Manuscript received March 25, 2004; revised January 23, 2005; accepted February 24, 2005. The associate editor coordinating the review of this paper and approving it for publication was Y-C. Liang. This paper has been published in part in Nordic Signal Processing Symposium, Espoo, Finland, June, 2004. This work was supported by the Academy of Finland, Nokia Foundation, TKK Foundation, Foundation of Commercial and Technical Sciences, and Grant Agency of Czech Republic by DG No. 102/03/H109.

J. Salo and P. Vainikainen are with the SMARAD Centre of Excellence, Radio Laboratory, Helsinki University of Technology, P.O. Box 3000, FI-02015-TKK, Finland (e-mail: jari.salo@tkk.fi, pertti.vainikainen@tkk.fi).

F. Mikas was with Czech Technical University of Prague. He is now with Ericsson CZ, Sokolovska 79, Prague 8, 186 00, Czech Republic (e-mail: filip.mikas@ericsson.com)

Digital Object Identifier 10.1109/TWC.2006.04188.

$\mathbf{H}$ perfectly, the mutual information is given by $I_{\mathbf{H}} = \log_2 |\mathbf{I}_K + \mathbf{W}|$, where the $K \times K$ matrix

$$\mathbf{W} = \begin{cases} \frac{\rho}{n_t}\mathbf{H}\mathbf{H}^H, & \text{if } n_r \le n_t \\ \frac{\rho}{n_t}\mathbf{H}^H\mathbf{H}, & \text{if } n_r > n_t, \end{cases}$$

and ρ is the average SNR at the output of each of the n_r receiving antennas [4]. For Rician fading a channel realization can be written $\mathbf{H} = a\mathbf{H}_d + b\mathbf{H}_s$, where

$$a = \sqrt{\frac{K_r}{K_r+1}}, \quad b = \sqrt{\frac{1}{K_r+1}}.$$

Here $K_r = \frac{a^2}{b^2}$ is the Rician K-factor. The line-of-sight component $\mathbf{H}_d$ is modelled as a deterministic constant. For $n_r \le n_t$ ($n_r > n_t$) the columns (rows) of the stochastic component $\mathbf{H}_s$ are iid complex multivariate zero-mean Gaussian random variables with $n_r \times n_r$ ($n_t \times n_t$) correlation matrix $\mathbf{\Sigma}_s = n_t^{-1}E[\mathbf{H}_s\mathbf{H}_s^H]$ ($\mathbf{\Sigma}_s = n_r^{-1}E[\mathbf{H}_s^H\mathbf{H}_s]$). Following the convention from [2] we call this semicorrelated fading. We adopt the usual power normalization $E[\|\mathbf{H}\|_F^2] = n_r n_t$. Complex isotropic Gaussian noise is assumed throughout the paper.

In what follows, we derive an upper bound for $E[I_{\mathbf{H}}]$ when $\mathbf{H}$ is Rician fading. The result allows estimation of the highest achievable information rates with uniform transmit power allocation, which is typically used in open loop systems where transmitter has no channel state information. Note that, for the Rician channel the Gaussian isotropic input does not, in general, achieve channel capacity. However, it has been shown that it is the asymptotically optimal channel input for high SNR's [15]. Furthermore, in [13] it is demonstrated that the channel state feedback delay must be fairly short with respect to the channel coherence time for the capacity to be significantly higher than the information rates achieved with uniform power allocation. However, assuming that the instantaneous channel state information is available at the transmitter side, the uniform power allocation (for $K_r \gg \rho$) may be inferior to the optimal water-filling strategy [17]. Nevertheless, the results in the next section provide a convenient means for estimating channel capacity in many cases of practical interest.

III. Main Results

A. The Basic Idea of Derivation

The derivation of the upper bound is essentially based on two key components: Jensen's inequality and expected values of elementary symmetric functions of a complex noncentral Wishart matrix. This becomes apparent from the following development:

$$\begin{aligned} E[I_{\mathbf{H}}] &= E\left[\log_2 |\mathbf{I}_K + \mathbf{W}|\right] \\ &\le \log_2 \left[E\left[\prod_{p=1}^{K} (1+\lambda_p) \right] \right] \\ &= \log_2 \left[\sum_{p=0}^{K} E[\mathrm{tr}_p(\mathbf{W})] \right], \end{aligned} \tag{1}$$

where Jensen's inequality and the concavity of $\log_2(\cdot)$ was used in the second line, and $\mathrm{tr}_p(\mathbf{W})$, $p = 0, \ldots, K$, is the pth *elementary symmetric function*[2] of $\mathbf{W}$ [21]. The function $\mathrm{tr}_p(\mathbf{W})$ depends only on the eigenvalues of $\mathbf{W}$, which are denoted with $\lambda_1, \ldots, \lambda_K$. For instance, $\mathrm{tr}_0(\mathbf{W}) = 1$, $\mathrm{tr}_1(\mathbf{W}) = \sum_{p=1}^{K} \lambda_p = \mathrm{tr}(\mathbf{W})$ and $\mathrm{tr}_K \mathbf{W} = |\mathbf{W}|$. In general, $\mathrm{tr}_p(\mathbf{W}) = \sum \lambda_{i_1} \cdots \lambda_{i_p}$, where the sum is over all $\binom{K}{p}$ combinations of the p indices with $i_1 < \ldots < i_p$. Interestingly, the expected values of elementary symmetric functions are known for many different statistics of $\mathbf{W}$ [22]. However, for *complex-valued* distributions results are considerably more scarce. In the case of Rician fading the distribution of $\mathbf{W}$ is *complex noncentral Wishart*[23]. Therefore, in order to compute the bound (1), we need to evaluate the expected values of the elementary symmetric functions of $\mathbf{W}$ over this distribution. Deferring the details of the derivation to appendix, we now jump to the main result of the paper.

B. The Upper Bound in the General Case

Denote $\mathbf{T} = \mathbf{H}_d\mathbf{H}_d^H$ (or $\mathbf{T} = \mathbf{H}_d^H\mathbf{H}_d$ if $n_r > n_t$), and with $\mathbf{A}^{i,p}$ the ith $p \times p$ principal submatrix [21] of the $K \times K$ matrix $\mathbf{A}$, $i = 1, \ldots, \binom{K}{p}$. The Pochhammer symbol is denoted $(a)_p = a(a+1)\cdots(a+p-1)$. We have

Theorem 1: Let $\mathbf{H}$ be Rician fading. Denote $K = \min(n_r, n_t)$, $L = \max(n_r, n_t)$, $K_r = \frac{a^2}{b^2}$, $a^2 = \frac{K_r}{K_r+1}$, and $b^2 = \frac{1}{K_r+1}$. The average mutual information can be upper bounded as

$$\begin{aligned} E[I_{\mathbf{H}}] \le \log_2 \Bigg[& 1 + \sum_{p=1}^{K} \left(\frac{\rho b^2}{n_t}\right)^p (L-p+1)_p \, \mathrm{tr}_p(\mathbf{\Sigma}_s) \\ & + \sum_{p=1}^{K} \left(\frac{\rho b^2}{n_t}\right)^p \sum_{j=1}^{p} K_r^j \, (L-p+1)_{(p-j)} \\ & \times \sum_{i=1}^{\binom{K}{p}} |\mathbf{\Sigma}_s^{i,p}| \, \mathrm{tr}_j[(\mathbf{\Sigma}_s^{i,p})^{-1}\mathbf{T}^{i,p}] \Bigg]. \end{aligned} \tag{2}$$

Proof: See appendix. ■

The first sum can be seen to be the contribution of the stochastic Rayleigh fading channel component (see Corollary 3), while the latter summation entails the effect of the LOS component.

C. The Upper Bound in Special Cases

1) Rayleigh iid $\mathbf{H}_s$: In this case $\mathbf{\Sigma}_s = \mathbf{I}_K$ and Theorem 1 simplifies to the following form.

Corollary 1: Let $\mathbf{H}$ be Rician fading with $\mathbf{\Sigma}_s = \mathbf{I}_K$. Then Theorem 1 becomes

$$\begin{aligned} E[I_{\mathbf{H}}] \le \log_2 \Bigg[& \sum_{p=0}^{K} \left(\frac{\rho b^2}{n_t}\right)^p \sum_{j=0}^{p} K_r^j \, (L-p+1)_{(p-j)} \\ & \times \binom{K-j}{p-j} \mathrm{tr}_j(\mathbf{T}) \Bigg]. \end{aligned} \tag{3}$$

Proof: See appendix. ■

Note that, for $\mathbf{\Sigma}_s = \mathbf{I}_K$, the upper bound depends only on $\mathbf{T} = \mathbf{H}_d\mathbf{H}_d^H$ through the elementary symmetric polynomials of its eigenvalues. This is in agreement with the findings in e.g. [16].

[2]We remark that the generating function of the elementary symmetric functions of $\mathbf{W}$ is $\sum_{p=0}^{K} \mathrm{tr}_p(\mathbf{W})t^p = \prod_{p=1}^{K}(1+\lambda_p t)$ [20]. This coincides with $|\mathbf{I}_K + t\mathbf{W}|$ and provides an alternative view on the capacity determinant.

2) Rayleigh iid $\mathbf{H}_s$ and rank-1 $\mathbf{H}_d$: In practice, the deterministic component $\mathbf{H}_d$ often has rank one, since in a line-of-sight scenario with linear transmit and receive arrays it is an outer product of two array response vectors [17], see also [9]. In this case we can further simplify Corollary 1 to yield

Corollary 2: Let $\mathbf{H}$ be Rician fading with $\mathbf{\Sigma}_s = \mathbf{I}_K$ and let $\mathrm{rank}(\mathbf{H}_d) = 1$. Then the upper bound (2) reduces to

$$E[I_{\mathbf{H}}] \leq \log_2\left[1 + \sum_{p=1}^{K}\sum_{j=0}^{1}\left(\frac{\rho b^2}{n_t}\right)^p (K_r K L)^j \times (L-p+1)_{(p-j)}\binom{K-j}{p-j}\right]. \quad (4)$$

Proof: See appendix. ∎

Note that in this case the average mutual information does not depend on the actual entries of the line-of-sight matrix $\mathbf{H}_d$, since, due to the power normalization, the only nonzero eigenvalue of $\mathbf{T}$ equals KL; hence $\mathrm{tr}_1(\mathbf{T}) = \mathrm{tr}(\mathbf{T}) = KL$ and $\mathrm{tr}_j(\mathbf{T}) = 0$ for $j > 1$.

3) Semicorrelated Rayleigh Fading: This special case results by setting $K_r = 0$ in Theorem 1.

Corollary 3: Let $\mathbf{H}$ be Rayleigh fading with correlation matrix $\mathbf{\Sigma}_s$. Theorem 1 becomes

$$E[I_{\mathbf{H}}] \leq \log_2\left[\sum_{p=0}^{K}\left(\frac{\rho}{n_t}\right)^p (L-p+1)_p \, \mathrm{tr}_p(\mathbf{\Sigma}_s)\right]. \quad (5)$$

The case of Rayleigh iid fading results by replacing $\mathrm{tr}_p(\mathbf{\Sigma}_s)$ with $\binom{K}{p}$ (since all eigenvalues are one). This results in an alternative form of the bounds [6, Eq. (22)][24, Eq. (8)] (see also [24, Eq. (28)]).

IV. Discussion

A. Tightness of the Bound

Denoting the random variable $X = |\mathbf{I}_K + \mathbf{W}|$ and its mean $m_X = E[X]$, the exact ergodic mutual information can written as

$$E[I_{\mathbf{H}}] = \log_2(m_X) - \epsilon, \quad (6)$$

where ϵ is the positive bounding error that depends on the spread of X; roughly speaking, the larger the variance, the larger the bounding error. In the sequel, we first consider the case with $K_r = 0$ and $\mathbf{\Sigma}_s = \mathbf{I}_K$ (Corollary 3). We show that at high SNR's the relative error of the bound tends to zero, and at low SNR's the absolute error tends to zero. We then argue heuristically that the bounding error ϵ cannot increase if $\mathbf{\Sigma}_s \neq \mathbf{I}_K$ *or* $K_r > 0$. The case $\mathbf{\Sigma}_s \neq \mathbf{I}_K$ *and* $K_r > 0$ will be examined by Monte Carlo simulations.

1) Relative error of the upper bound tends to zero as $\rho \to \infty$: Assume that $\rho \geq n_t$. Then we can further upper bound (5) as

$$\begin{aligned} E[I_{\mathbf{H}}] &\leq \log_2\left[\left(\frac{\rho}{n_t}\right)^K \sum_{p=0}^{K}(L-p+1)_p\binom{K}{p}\right] \\ &= K\log_2\left(\frac{\rho}{n_t}\right) \\ &\quad + \log_2\left[\sum_{p=0}^{K}(L-p+1)_p\binom{K}{p}\right]. \end{aligned} \quad (7)$$

Now consider the mutual information *lower* bound [24]

$$E[I_{\mathbf{H}}] \geq K\log_2\left(\frac{\rho}{n_t}\right) + \frac{1}{\ln 2}\sum_{i=0}^{K-1}\psi(L-i), \quad (8)$$

where $\psi(x)$ is the digamma function. By subtracting (8) from (7) we can upper bound the error as

$$\epsilon \leq \log_2\left[\sum_{p=0}^{K}(L-p+1)_p\binom{K}{p}\right] - \frac{1}{\ln 2}\sum_{i=0}^{K-1}\psi(L-i), \quad (9)$$

which is a constant (for fixed n_r and n_t) not depending on ρ. Therefore, for $\rho \geq n_t$, there is a constant upper bound on ϵ in (6). This means that as $\rho \to \infty$, the *relative* error of the upper bound (2) tends to zero, i.e., the ratio of upper bound (5) and the exact mutual information tends to one. We remark that for fixed L, (9) is maximized for $K = L$, and becomes smaller as $\frac{K}{L} \to 0$. For example, (9) gives 4.79 and 1.25 bits/s/Hz for 4×4 and 4×2 systems, respectively. This is obviously related to the fact that variance of mutual information decreases as the MIMO system becomes asymmetric [25], hence making the error term in (6) smaller.

2) Absolute error tends to zero as $\rho \to 0$: Assume that $\rho \ll n_t$. We can upper bound (5) as

$$\begin{aligned} E[I_{\mathbf{H}}] &\leq \log_2\left(1 + \frac{\rho}{n_t}KL\right) \\ &\leq \frac{\rho KL}{n_t \ln 2} \end{aligned} \quad (10)$$

by using $\ln(1+x) \leq x$. Using the lower bound from [25] and $\ln(1+x) \geq x - \frac{1}{2}x^2$ we can write

$$\begin{aligned} E[I_{\mathbf{H}}] &\geq E\left[\log_2\left(1 + \frac{\rho}{n_t}\|\mathbf{H}\|_F^2\right)\right] \\ &\geq \frac{\rho}{n_t \ln 2}E\left[\|\mathbf{H}\|_F^2\right] - \frac{1}{2\ln 2}\left(\frac{\rho}{n_t}\right)^2 E\left[\|\mathbf{H}\|_F^4\right] \\ &= \frac{\rho}{n_t}\frac{KL}{\ln 2} - \left(\frac{\rho}{n_t}\right)^2\frac{KL(KL+1)}{2\ln 2}, \end{aligned} \quad (11)$$

since $E[\|\mathbf{H}\|_F^4] = KL(KL+1)$. Subtracting (11) from (10) shows that the error of the mutual information upper bound is at most $\left(\frac{\rho}{n_t}\right)^2\frac{KL(KL+1)}{2\ln 2}$. Hence, as $\rho \to 0$, the bounding error tends to zero. The mutual information itself decreases linearly with ρ as $\rho \to 0$.

3) The Upper Bound Becomes Sharper as $K_r > 0$ or $\mathbf{\Sigma}_s \neq \mathbf{I}_K$: Assume first that $K_r = 0$. The results in [2] and [25] indicate[3] that the variance of the mutual information does not increase for $\mathbf{\Sigma}_s \neq \mathbf{I}_K$. Hence, the bounding error ϵ cannot increase as the level of spatial correlation increases. Finally, we remark that as K_r increases (for a fixed $\mathbf{H}_d$ and $\mathbf{\Sigma}_s = \mathbf{I}_K$) the variance of the mutual information decreases, since the stochastic component $\mathbf{H}_s$ vanishes [11]. For $K_r = \infty$, the term inside $\log_2(\cdot)$ becomes deterministic and the mutual information upper bound holds trivially with equality.

[3]Since [25, Eqs. (35)-(37)] are based on truncated sum of random variables (see [25, Eq. (13)]) they provide upper bounds on the variance of mutual information for the spatially semicorrelated Rayleigh fading case.

B. Computational Complexity

It can be shown that, for K large, the computational complexity of the general bound (2) scales as $O(2^K K^3)$ with $K = \min(n_r, n_t)$. This means that the computation time more than doubles when K is increased by one. Hence, using the general result becomes quickly impractical for $K > 10$. The complexity arises from the inner summation where a hermitian eigenvalue problem with cubic complexity needs to be solved $2^K - 1$ times. The complexity decreases dramatically for the Corollaries 1-2, where it is assumed that $\mathbf{\Sigma}_s = \mathbf{I}_K$. For (3), the complexity increases as $O(K^3)$, since it is enough to compute the eigenvalues of $\mathbf{T} = \mathbf{H}_d \mathbf{H}_d^H$ only once. Computational complexity of Corollary 2 scales as $O(K)$, while that of Corollary 3 scales as $O(K^3)$ for $\mathbf{\Sigma}_s \neq \mathbf{I}_K$. However, none of Corollaries 1– 3 should present any computational complexity problems for any practical value of K.

Note that the elementary symmetric polynomials can be computed recursively with low cost as follows. Let $\{\lambda_p\}_{p=1}^K$ denote the eigenvalues of the $K \times K$ matrix $\mathbf{A}$. Then the power sums $S_n = \sum_{p=1}^K \lambda_p^n$ and the elementary symmetric polynomials of $\mathbf{A}$ are related by the Newton formula as [20] $\mathrm{tr}_p(\mathbf{A}) = \frac{1}{p} \sum_{n=1}^p (-1)^{n-1} S_n \, \mathrm{tr}_{p-n}(\mathbf{A})$. Hence the summation over the $\binom{K}{p}$ monomials in the definition of $\mathrm{tr}_p(\mathbf{A})$ is avoided.

V. Numerical Examples

In the examples we consider a 4×4 MIMO system. We choose two 4×4 LOS matrices: $\mathbf{H}_{d1} = 2\mathbf{I}_4$, and $\mathbf{H}_{d2} = \mathbf{a}(\theta_r)\mathbf{a}(\theta_t)^T$, where $\mathbf{a}(\theta) = [1 \; e^{j\pi \sin(\theta)} \ldots e^{j\pi 3 \sin(\theta)}]^T$, $\theta_r = 30°$, and $\theta_t = 0°$. Note that $\mathbf{H}_{d1}$ corresponds to four parallel equal-gain AWGN channels[4], whereas $\mathbf{H}_{d2}$ is a rank-1 LOS matrix arising from the outer product of the responses of linear transmit and receive arrays with zero degree angle of departure and 30 degree angle of arrival. In all cases, the exact values for mutual information have been estimated from Monte Carlo simulation with 10^5 realizations. Note that we have plotted the worst case results in the sense that the bound is tighter for $L > K$, as can be seen from (9) and from numerical experiments (not shown due to space limitations).

A. Effect of SNR and K-factor with Spatially Uncorrelated Fading ($\mathbf{\Sigma}_s = \mathbf{I}_4$)

In Fig. 1, we plot the exact and bounded ergodic mutual information (right), and the estimated bounding error (left). For large SNR's, the error of the upper bound tends to a constant; for a 4×4 system (9) predicts this constant to be at most 4.79 bits/s/Hz, whereas the true error is below 1.8 bits/s/Hz in all examined cases. The bounding error becomes quite small for $\rho < K_r$. The bound is loosest for $K_r = -\infty$ dB as is to be expected, since the variance of mutual information in (9) is largest in this case. We also plot the result due to Hösli and Lapidoth [15], which upper bounds the rate loss between Gaussian isotropic inputs and the optimal power allocation, assuming that the transmitter knows only the channel distribution[5] (instead of the instantaneous $\mathbf{H}$ as in the water-filling scheme). This illustrates the optimality of uniform power allocation for $\rho \gg K_r$. In comparison to the result in [11, Th. 2] the error of the bounds in the present paper are smaller at low SNR's, whereas the lower bound therein is tighter for $\rho \gg K_r$.

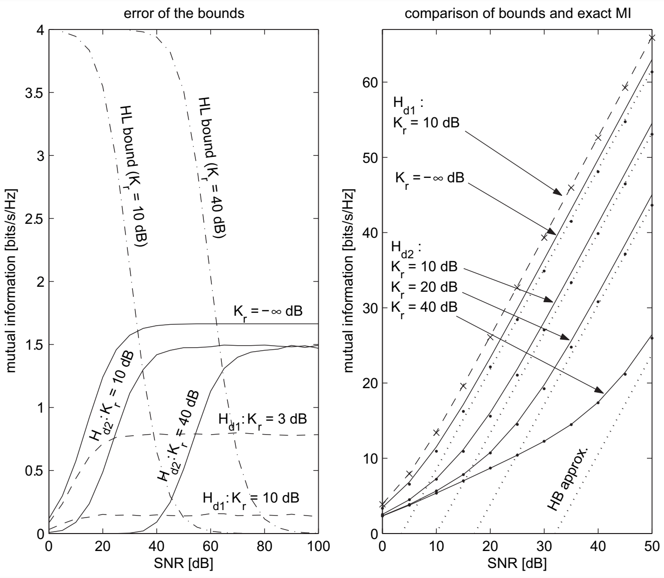

Fig. 1. Comparison of bounds to exact MI (Monte Carlo simulated) for 4×4 MIMO system. Dashed line: $\mathbf{H}_{d1} = 2\mathbf{I}_4$ computed with bound (3). Solid line: $\mathbf{H}_{d2}$ (rank-1) computed with bound (4). *Left:* Error of the upper bounds in bits/s/Hz. Dash-dot line: Hösli-Lapidoth bound [15, Th. 1]. *Right:* Comparison of bounds and exact MI. Crosses: Exact MI for $\mathbf{H}_{d1}$. Dots: Exact MI for $\mathbf{H}_{d2}$. Dotted line: Hansen-Bölcskei approximation [11, Th. 2].

B. Effect of Spatial Correlation and K-factor

We consider the simple exponential correlation introduced in [26] and subsequently used as a test case in a number of diversity and capacity studies, e.g. [2] [7]. The entries of $\mathbf{\Sigma}_s$ are $\sigma_{ij} = r^{|i-j|}$ with $r \in [0, 1)$. In Fig. 2, the effect of correlation is plotted for varying K_r and $\rho = 30$ dB. The results are computed with the general bound (2). It is clear that the bound becomes tighter as spatial correlation increases. The results also confirm the intuition that high K factor provides robustness against spatial correlation.

VI. Conclusion

We have derived a general upper bound for the ergodic mutual information in Rician fading MIMO channels with uniform power allocation at the transmitter. The bound was shown to be asymptotically tight for low SNR's with a vanishing relative error at high SNR's. It provides practical means for estimating achievable information rates in many practical MIMO scenarios. The cases of rank-1 line-of-sight matrix and spatially correlated Rayleigh fading are special instances of the general result.

[4]For this LOS matrix the uniform power allocation achieves capacity, when the transmitter knows only the probability distribution of the MIMO channel [13].

[5]Due to different power normalization, comparison to this result requires replacing ρ in [15, Eq. (3)] with $\rho_{eff} = \frac{\rho}{n_t(1+K_r)}$, where in this case $K_r = \frac{\|\mathbf{H}_d\|_F^2}{n_r n_t}$. Otherwise, the upper bound on rate loss is independent of the Rician K-factor.

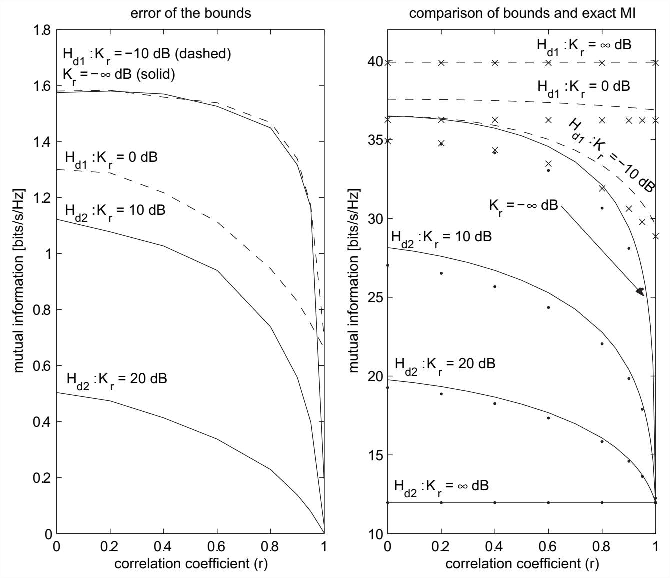

Fig. 2. Comparison of the general bound (2) to exact MI (Monte Carlo simulated) for 4 × 4 MIMO system for different values of correlation coefficient r, $\rho = 30$ dB. Dashed line: $\mathbf{H}_{d1} = 2\mathbf{I}_4$. Solid line: $\mathbf{H}_{d2}$ (rank-1). *Left:* Error of the upper bounds in bits/s/Hz. *Right:* Comparison of bounds and exact MI. Crosses: Exact MI for $\mathbf{H}_{d1}$. Dots: Exact MI for $\mathbf{H}_{d2}$.

APPENDIX I

A. Notation

$(a)_p$	$a(a+1)\dots(a+p-1)\,, \quad (a)_0 = 1$
$\mathrm{etr}(\mathbf{X})$	$\exp[\mathrm{tr}(\mathbf{X})]$
$\Gamma(t)$	$\int_0^\infty x^{t-1}e^{-x}dx$
$\Gamma_K(L)$	$\pi^{K(K-1)/2}\prod_{p=1}^K \Gamma(L-p+1)$

B. Partitions and Complex Zonal Polynomials

Let $\kappa = (k_1, k_2, \dots, k_Q)$ be a Q-partition[6] of positive integer p such that $k_1 \geq k_2 \geq \dots \geq k_Q \geq 0$ and $\sum_{i=1}^Q k_i = p$. The *generalized complex hypergeometric coefficient* associated with the Q-partition κ is defined as [23]

$$[a]_\kappa = \prod_{i=1}^{Q}(a-i+1)_{k_i}\,. \tag{12}$$

The *complex zonal polynomial* of a $K \times K$ hermitian positive definite matrix $\mathbf{A}$ can be defined as $C_\kappa(\mathbf{A}) = \chi_\kappa(1)\chi_\kappa(\mathbf{A})$ [27] where the scaling constant is given by

$$\chi_\kappa(1) = \frac{p!\left[\prod_{m<n}^{K}(k_m - k_n - m + n)\right]}{\prod_{m=1}^{K}(k_m + K - m)!}\,, \tag{13}$$

and the polynomial in the eigenvalues of $\mathbf{A}$ is

$$\chi_\kappa(\mathbf{A}) = \frac{\begin{vmatrix} \lambda_1^{k_1+K-1} & \lambda_1^{k_2+K-2} & \cdots & \lambda_1^{k_K} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_K^{k_1+K-1} & \lambda_K^{k_2+K-2} & \cdots & \lambda_K^{k_K} \end{vmatrix}}{\begin{vmatrix} \lambda_1^{K-1} & \lambda_1^{K-2} & \cdots & \lambda_1^{0} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_K^{K-1} & \lambda_K^{K-2} & \cdots & \lambda_K^{0} \end{vmatrix}}\,.$$

[6]For instance, the 2-partitions of $p = 2$ are $\kappa = (2, 0)$ and $\kappa = (1, 1)$.

Consider a partition κ_1 with $k_i = 1, i = 1, \dots, p$, and $k_i = 0, i = p+1, \dots, K$. To conclude this section we outline the derivation of the fact that $C_{\kappa_1}(\mathbf{A}) = \mathrm{tr}_p(\mathbf{A})$, i.e., the pth elementary symmetric function of $\mathbf{A}$. The result differs from the case of real-valued zonal polynomials and will be needed in the sequel. To compute the scaling constant we note that[7] where some of the details can be verified by the reader. From a result on Vandermonde determinants [28, Eq. (4.9)] it readily follows that $\chi_{\kappa_1}(\mathbf{A}) = \mathrm{tr}_p(\mathbf{A})$, therefore establishing the desired result.

C. Proofs

We next outline the derivation of $E[\mathrm{tr}_p(\mathbf{W})]$. For the real-valued noncentral Wishart $\mathbf{W}$ the result has been proven in [29]. The corresponding proof for the complex-valued $\mathbf{W}$ appears to be unavailable, so we outline the main steps here.

Let $\mathbf{S}$ be complex noncentral Wishart matrix with parameters $(L, \mathbf{\Sigma}_s, \mathbf{\Theta})$, where $\mathbf{\Theta} = \mathbf{\Sigma}_s^{-1}\mathbf{H}_d\mathbf{H}_d^H$ is the noncentrality parameter. Using the pdf of $\mathbf{S}$ from [23] the expected value of its determinant can be written as

$$\begin{aligned} E[|\mathbf{S}|] &= \int_{\mathbf{S}>0} |\mathbf{S}| f(\mathbf{S}) d\mathbf{S} \\ &= \frac{\mathrm{etr}[-\mathbf{\Theta}]}{\Gamma_K(L)|\mathbf{\Sigma}_s|^L} \int_{\mathbf{S}>0} |\mathbf{S}|^{L-K+1}\,\mathrm{etr}[-\mathbf{\Sigma}_s^{-1}\mathbf{S}] \\ &\quad \times {}_0F_1\left(L; \mathbf{\Theta}\mathbf{\Sigma}_s^{-1}\mathbf{S}\right) d\mathbf{S} \\ &= \frac{\mathrm{etr}[-\mathbf{\Theta}]\,|\mathbf{\Sigma}_s|\,\Gamma_K(L+1)}{\Gamma_K(L)}\,{}_1F_1\left(L+1; L; \mathbf{\Theta}\right)\,, \end{aligned}$$

where ${}_0F_1(b; \mathbf{A})$ and ${}_1F_1(a; b; \mathbf{A})$ are hypergeometric functions of complex matrix argument and we also used a definite integral[8] from [27]. From the Kummer relation [30, Eq. (2.8)] ${}_1F_1(a; b; \mathbf{A}) = \mathrm{etr}(\mathbf{A})\,{}_1F_1(b-a; b; -\mathbf{A})$. Furthermore, since $\frac{\Gamma_K(L+1)}{\Gamma_K(L)} = (L-K+1)_K$ we can write $E[|\mathbf{S}|] = (L-K+1)_K\,|\mathbf{\Sigma}_s|\,{}_1F_1\left(-1; L; -\mathbf{\Theta}\right)$. By definition [27, p. 369]

$${}_1F_1\left(-1; L; -\mathbf{\Theta}\right) = \sum_{p=0}^{\infty}\sum_{\text{all }\kappa} \frac{[-1]_\kappa}{[L]_\kappa}\frac{C_\kappa(-\mathbf{\Theta})}{p!}\,. \tag{14}$$

The key observation is that $(-1)_p = 0$ for $p > 1$ and hence we can restrict to partitions $\kappa_1 = (1, 1, \dots, 1)$ with p ones. From (12) it follows that $[-1]_{\kappa_1} = (-1)^p\,p!$ and $[L]_{\kappa_1} = (L-p+1)_p$. Hence we have

$${}_1F_1\left(-1; L; -\mathbf{\Theta}\right) = \sum_{p=0}^{K} \frac{(-1)^p}{(L-p+1)_p} C_{\kappa_1}(-\mathbf{\Theta})\,. \tag{15}$$

In Appendix IB, it was shown that the zonal polynomial $C_{\kappa_1}(-\mathbf{\Theta}) = (-1)^p\,\mathrm{tr}_p(\mathbf{\Theta})$. Hence, after noting that

$$\frac{(L-K+1)_K}{(L-p+1)_p} = (L-K+1)_{(K-p)}\,,$$

we arrive at

$$E[|\mathbf{S}|] = |\mathbf{\Sigma}_s| \sum_{p=0}^{K}(L-K+1)_{(K-p)}\ \mathrm{tr}_p(\mathbf{\Theta})\,. \tag{16}$$

[7]The first equality follows after straightforward, but somewhat tedious, examination of (13) and the k_i:s.

[8]The integral can be evaluated by using the zonal polynomial definition of ${}_0F_1(b; \mathbf{A})$ [27, p. 369] and integrating the sum term by term using [27, Eq. (6.1.20)].

$$\chi_{\kappa_1}(1) = \frac{p! \prod_{i<j}^{p}(j-i) \prod_{i=1}^{K-p-1}(K-p-i)! \prod_{i=1}^{p}(i+1)_{(K-p)}}{\prod_{i=1}^{p}(K+1-i)! \prod_{j=p+1}^{K}(K-j)!}$$
$$= \frac{p! \prod_{i=1}^{p-1}(p-i)! \prod_{i=1}^{K-p-1}(K-p-i)! \prod_{i=1}^{p}(i+1)_{(K-p)}}{\prod_{i=1}^{p}(K+1-i)! \prod_{j=1}^{K-p-1}(K-p-j)!} = \frac{\prod_{i=1}^{p}(K+1-i)!}{\prod_{j=1}^{p}(K+1-j)!} = 1$$

Denote with $\mathbf{A}^{i,p}$ the ith $p \times p$ principal submatrix of a $K \times K$ matrix $\mathbf{A}$. The pth elementary symmetric function, $(1 \leq p \leq K)$, can be written [21]

$$\mathrm{tr}_p(\mathbf{A}) = \sum_{i=1}^{\binom{K}{p}} |\mathbf{A}^{i,p}| \, .$$

The principal submatrix $\mathbf{S}^{i,p}$ is distributed as complex noncentral Wishart with parameters $(L, \mathbf{\Sigma}_s^{i,p}, (\mathbf{\Sigma}_s^{i,p})^{-1}\mathbf{T}^{i,p})$, where $\mathbf{T} = \mathbf{H}_d\mathbf{H}_d^H$. We can use (16) and write

$$E[\mathrm{tr}_p(\mathbf{S})] = (L-p+1)_p \ \mathrm{tr}_p(\mathbf{\Sigma}_s) + \sum_{j=1}^{p}(L-p+1)_{(p-j)} \times \sum_{i=1}^{\binom{K}{p}} |\mathbf{\Sigma}_s^{i,p}| \, \mathrm{tr}_j[(\mathbf{\Sigma}_s^{i,p})^{-1}\mathbf{T}^{i,p}] \, . \quad (17)$$

Theorem 1 follows from (17).

The special case $\mathbf{\Sigma}_s = c\mathbf{I}_K$ *:* If $\mathbf{\Sigma}_s = c\mathbf{I}_K$, for some positive scalar c, this reduces to ($|\mathbf{\Sigma}_s^{i,p}| = c^p$)

$$E[\mathrm{tr}_p(\mathbf{S})] = \sum_{j=0}^{p}(L-p+1)_{(p-j)} \, c^{p-j} \sum_{i} \mathrm{tr}_j(\mathbf{T}^{i,p}) = \sum_{j=0}^{p}(L-p+1)_{(p-j)} \, c^{p-j} \binom{K-j}{p-j} \mathrm{tr}_j(\mathbf{T}) \, . \quad (18)$$

The second equality follows because $\sum_{i=1}^{\binom{K}{p}} \mathrm{tr}_j(\mathbf{T}^{i,p})$ is a sum of $\binom{K}{p}\binom{p}{j}$ $j \times j$ principal minors where each term appears $\binom{K-j}{p-j}$ times. Corollary 1 follows by noting that $\mathrm{tr}_p(c\mathbf{A}) = c^p \, \mathrm{tr}_p(\mathbf{A})$.

The special case $\mathbf{\Sigma}_s = c\mathbf{I}_K$ *and* $\mathbf{H}_d$ *has rank one :* Since $\mathrm{tr}_j(\mathbf{T}) = 0$ for $j > 1$ we have from (18)

$$E[\mathrm{tr}_p(\mathbf{S})] = \sum_{j=0}^{1}(L-p+1)_{p-j} \, c^{p-j} \binom{K-j}{p-j} (KL)^j$$

Corollary 2 follows by noting that $\mathrm{tr}_1(\mathbf{H}_d\mathbf{H}_d^H) = KL$, which, in turn, results from the power normalization $E[\|\mathbf{H}\|_F^2] = KL$.

REFERENCES

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

[2] P. J. Smith, S. Roy, and M. Shafi, "Capacity of MIMO systems with semicorrelated flat fading," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2781 –2788, Oct. 2003.

[3] 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.

[4] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. on Telecomm.*, vol. 10, no. 6, pp. 585–596, Nov. 1999.

[5] T. Ratnarajah, R. Vaillancourt, and M. Alvo, "Complex random matrices and Rayleigh channel capacity," *Communications in Information and Systems*, vol. 3, no. 2, pp. 119–138, Oct. 2003.

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

[7] M. Kiessling and J. Speidel, "Mutual information of MIMO channels in correlated Rayleigh fading environments - a general solution," in *Proc. IEEE Int. Conf. Comm.*, vol. 2, Paris, France, 2004, pp. 814 – 818.

[8] P. F. Driessen and G. J. Foschini, "On the capacity formula for multiple input-multiple output wireless channels: a geometric interpretation," *IEEE Trans. Commun.*, vol. 47, no. 2, pp. 173–176, Feb. 1999.

[9] M. Khalighi, J.-M. Brossier, G. Jourdain, and K. Raoof, "On capacity of Rician MIMO channels," in *Proc. IEEE Int. Symp. Personal, Indoor and Mobile Radio Communications*, vol. 1, Sept. 2001, pp. 150–154.

[10] G. Lebrun, M. Faulkner, M. Shafi, and P. J. Smith, "MIMO Ricean channel capacity," in *Proc. IEEE Int. Conf. Comm.*, vol. 5, Paris, France, June 2004, pp. 2939–2943.

[11] J. Hansen and H. Bölcskei, "A geometrical investigation of the rank-1 Ricean MIMO channel at high SNR," in *Proc. IEEE Symp. Info. Theory.*, Chicago, USA, June 2004.

[12] 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.

[13] S. Venkatesan, S. H. Simon, and R. A. Valenzuela, "Capacity of a Gaussian MIMO channel with nonzero mean," in *Proc. IEEE Veh. Tech. Conf.*, Orlando, USA, Oct. 2003.

[14] D. Hösli and A. Lapidoth, "The capacity of a MIMO Ricean channel is monotonic in the singular values of the mean," in *Proc. 5th ITG Conf. Source and Channel Coding*, Jan. 2004.

[15] ——, "How good is an isotropic Gaussian input on a MIMO Ricean channel?" in *Proc. IEEE Int. Symp. Inf. Th.*, Chicago, USA, June 2004, p. 291.

[16] Y.-H. Kim and A. Lapidoth, "On the log determinant of non-central Wishart matrices," in *Proc. IEEE Int. Symp. Inf. Th.*, Yokohama, Japan, June 2003, p. 54.

[17] F. R. Farrokhi, G. J. Foschini, A. Lozano, and R. A. Valenzuela, "Link-optimal space-time processing with multiple transmit and receive antennas," *IEEE Commun. Lett.*, vol. 5, no. 3, pp. 85–87, Mar. 2001.

[18] M. Kang and M.-S. Alouini, "Capacity of MIMO Rician channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 112–122, Jan. 2006.

[19] 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.

[20] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, 2nd ed. Oxford University Press, 1995.

[21] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[22] A. K. Gupta and D. K. Nagar, *Matrix Variate Distributions*. Chapman and Hall/CRC, 2000.

[23] P. R. Krishnaiah, "Some recent developments on complex multivariate distributions," *Journal of Multivariate Analysis*, vol. 6, pp. 1–30, 1976.

[24] A. Grant, "Rayleigh fading multiple-antenna channels," *EURASIP Journal on Applied Signal Processing*, vol. 2002, no. 3, pp. 316–329, Mar. 2002.

[25] Ö. Oyman, R. Nabar, H. Bölcskei, and A. Paulraj, "Characterizing the statistical properties of mutual information in MIMO channels," *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2784 –2795, Nov. 2003.

[26] V. A. Aalo, "Performance of maximal-ratio diversity systems in a correlated Nakagami-fading environment," *IEEE Trans. Commun.*, vol. 43, no. 8, pp. 2360–2369, Aug. 1995.

[27] A. M. Mathai, *Jacobians of Matrix Transformations and Functions of Matrix Arguments*. World Scientific Publishing, 1997.

[28] F. Zhang, *Matrix Theory: Basic Results and Techniques*. Springer-Verlag, 1999.

[29] B. K. Shah and C. G. Khatri, "Proof of conjectures about the expected values of the elementary symmetric functions of a noncentral Wishart matrix," *The Annals of Statistics*, vol. 2, no. 4, pp. 833–836, July 1974.

[30] C. S. Herz, "Bessel functions of matrix argument," *The Annals of Mathematics*, vol. 61, no. 3, pp. 474–523, May 1955.

Jari Salo received the degrees of Master of Science in Technology and Licentiate of Science in Technology from Helsinki University of Technology (TKK) in 2000 and 2003, respectively. From 2001 to 2002 he was with Bitville Oy. Since November 2002, he has been with the Radio Laboratory of TKK where he is currently working toward the doctoral degree.

Filip Mikas was born in 1978 in Prague, Czech Republic. He received his M.S. and Ph.D. degrees in electrical engineering from The Czech Technical University of Prague in 2002 and 2005. Since 2005 he has been working for Ericsson Microwave Systems as customer support engineer in Prague.

Pertti Vainikainen (M'91) received the degree of Master of Science in Technology, Licentiate of Science in Technology and Doctor of Science in Technology from Helsinki University of Technology (TKK) in 1982, 1989 and 1991, respectively. From 1992 to 1993 he was Acting Professor of Radio Engineering, since 1993 Associate Professor of Radio Engineering and since 1998 Professor in Radio Engineering, all in the Radio Laboratory of TKK. In 1993-97 he was the director of the Institute of Radio Communications (IRC) of TKK and in 2000 a visiting professor at Aalborg University, Denmark. His main fields of interest are antennas and propagation in radio communications and industrial measurement applications of radio waves. He is the author or co-author of 6 books and about 240 refereed international journal or conference publications and the holder of 7 patents.