

Ergodic Capacity of MIMO Channels with Statistical Channel State Information at the Transmitter

Mario Kiessling^{1,2}, Joachim Speidel¹, Markus Reinhardt²

¹Institute of Telecommunications, University of Stuttgart, Germany

²Siemens AG, Information and Communication Mobile, Ulm, Germany
mario.kiessling@gmx.net, phone +49 731 9533 2005

Abstract—It is well known that with the availability of statistical channel state information at the transmitter, the capacity-achieving transmission strategy is transmission on the long-term eigenmodes of the transmit correlation matrix with adequate power allocation. However, the optimum power allocation strategy is not known in general. Using recent analytical results on mean mutual information of MIMO channels with transmit as well as receive correlation, in this paper we study the behavior of the capacity-achieving power allocation strategy. To this end, we also make use of asymptotical results for a large number of transmit or receive antennas and the high as well as low SNR regime, which in certain cases allows for a closed-form analysis. Furthermore, we investigate two low complexity power allocation schemes, which are based on certain upper bounds on mean mutual information.

I. INTRODUCTION

Based on a moment generating function approach, exact closed-form expressions of ergodic MIMO capacity were recently presented by the authors in [13][14]. The expressions therein cover the case of fading correlation at transmitter as well as receiver and base on the assumption of an uninformed transmitter, i.e. there is no channel state information (CSI) available at the transmitter and it applies a uniform power allocation on the transmit antennas. In this paper, we extend these results to cover the case of statistical CSI at the transmitter, whereas the transmitter is assumed to be aware of the correlation properties of the channel. It is well known that the transmitter has to adapt the covariance matrix of the transmit signal vector according to the CSI such that it essentially transmits with a proper power allocation (PA) on the eigenmodes of the fading correlation matrix at the transmitter side [1][2][3]. However, in contrast to the case of instantaneous TX CSI [4], to the authors' best knowledge, so far the optimum capacity-achieving statistical power allocation strategy is known only for the special case of 2 transmit antennas and uncorrelated receive antennas [5]. Based on the exact mean mutual information (MMI) expressions given in [14], in this paper we fill this gap and extend the results of [5] to an arbitrary number of transmit antennas and arbitrary fading correlation at the receive antenna array. To this end, we numerically optimize the transmit signal covariance matrix using the novel exact MMI expressions. For gaining further insights into the problem, we also study ergodic capacity (EC) asymptotics and the corresponding power allocation strategies in the low and high SNR regime, whereas we make use of asymptotical MMI expressions derived by the authors in [16]. Similar studies are presented for the limit of a

large number of transmit as well as receive antennas, respectively. Furthermore, we compare the results with some lower complexity power allocation schemes based on certain bounds on MMI [6][15]. Specifically, we demonstrate that a power allocation based on the tight bound derived by the authors in [15] is hardly distinguishable from the exact capacity-achieving strategy for arbitrary system parameters. On the other hand, we show that a power allocation scheme based on a loose bound on MMI introduced in [6] is asymptotically optimal for a large number of receive antennas due to the so-called 'channel hardening' effect (see also [7]).

Various simulation results confirm the effectiveness of the proposed power allocation strategies.

II. SIGNAL AND CHANNEL MODEL

We consider a flat fading MIMO link modeled by

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}, \quad (1)$$

where \mathbf{s} is the $T \times 1$ TX symbol vector, \mathbf{H} is the $R \times T$ MIMO channel matrix with correlated Rayleigh fading elements, \mathbf{n} is the $R \times 1$ noise vector, and \mathbf{y} is the $R \times 1$ receive vector. By R we denote the number of RX antennas and T is the number of TX antennas. In the following we assume additive Gaussian noise, where the noise covariance matrix is given by $\mathbf{R}_{nn} = N_0 \cdot \tilde{\mathbf{R}}_{nn}$. The signal covariance matrix is given by the analogue expression $\mathbf{R}_{ss} = E_s \cdot \tilde{\mathbf{R}}_{ss}$.

In this paper, \mathbf{I}_n denotes an $n \times n$ identity matrix (the index can be omitted, if the size is clear from the context), $\text{tr}(\mathbf{X})$ denotes the trace of matrix \mathbf{X} , $\text{rk}(\mathbf{X})$ denotes the rank (i.e. number of nonzero eigenvalues), $\text{eig}(\mathbf{X})$ returns a diagonal matrix of eigenvalues of \mathbf{X} , $\text{diag}(\mathbf{X})$ a diagonal matrix of nonzero eigenvalues of \mathbf{X} , \mathbf{x}^H denotes Hermitian (conjugate transpose), and \equiv means 'is statistically equivalent'.

Using a widely accepted channel model [8][9], the correlated MIMO channel can be described by the matrix product

$$\mathbf{H} = \mathbf{A}^H \mathbf{H}_w \mathbf{B}, \quad (2)$$

where \mathbf{H}_w is a $R \times T$ matrix of complex i.i.d. Gaussian variables of unity variance and

$$\mathbf{A}\mathbf{A}^H = \mathbf{R}_{RX} \quad \mathbf{B}\mathbf{B}^H = \mathbf{R}_{TX} = \mathbf{V}_{TX} \mathbf{\Lambda}_{TX} \mathbf{V}_{TX}^H, \quad (3)$$

where \mathbf{R}_{RX} and \mathbf{R}_{TX} is the long-term stable normalized

$$\text{tr}(\mathbf{R}_{RX}) = R \quad \text{tr}(\mathbf{R}_{TX}) = T \quad (4)$$

receive and transmit correlation matrix, respectively. Moreover, in (3) we have introduced an eigenvalue decomposition of the transmit correlation matrix with diagonal matrix of eigenvalues $\mathbf{\Lambda}_{TX}$, whereas the eigenvalues are assumed to be arranged in decreasing order.

III. MIMO MUTUAL INFORMATION

In this section we present a unifying notation for mutual information. Based on that, we give a closed-form expression for mean mutual information (MMI) (i.e. mutual information averaged over the channel statistics). Furthermore, we outline asymptotical MMI formulas for the low and high SNR regime.

A. General derivation

It is well known [10] that the mutual information $I(s, y)$ between input vector s and output y of the MIMO link according to (1) is given by

$$I(s, y) = \log_2 \left| \mathbf{I} + \gamma \tilde{\mathbf{R}}_{ss} \mathbf{H}^H \tilde{\mathbf{R}}_{nn}^{-1} \mathbf{H} \right| \quad (5)$$

with the standard mean SNR per transmit symbol definition $\gamma = E_s/N_0$. Plugging the channel model (2) with Kronecker product covariance structure in (5), we find

$$I(s, y) = \log_2 \left| \mathbf{I} + \gamma \tilde{\mathbf{R}}_{ss} \mathbf{B}^H \mathbf{H}_w^H \mathbf{A} \tilde{\mathbf{R}}_{nn}^{-1} \mathbf{A}^H \mathbf{H}_w \mathbf{B} \right|. \quad (6)$$

In the following, we reduce (6) to a concise equivalent formulation that allows for a unified analysis of correlated MIMO systems. At this point, we emphasize that we assume full rank channel correlation and signal and noise covariance matrices in this paper. An extension is straightforward but would unnecessarily complicate notation, thus detracting from the main problems. By noticing that the distribution of the i.i.d. complex Gaussian distributed $R \times T$ matrix \mathbf{H}_w is invariant to left- or right multiplications with unitary matrices \mathbf{U} and \mathbf{V} , i.e.

$$\mathbf{U} \mathbf{H}_w \mathbf{V} \cong \mathbf{H}_w, \quad (7)$$

assuming capacity-achieving transmission with channel distribution information at the transmitter (CDIT) on the eigenmodes of \mathbf{R}_{TX} [1][2][3] according to (3), such that

$$\tilde{\mathbf{R}}_{ss} = \mathbf{V}_{TX} \cdot \mathbf{\Phi}^2 \cdot \mathbf{V}_{TX}^H \quad (8)$$

with diagonal $T \times T$ power allocation (PA) or alternatively waterfilling (WF) matrix $\mathbf{\Phi} = \text{diag}(\phi_1, \dots, \phi_T)$, we find after some simplifications

$$I(s, y, \mathbf{\Phi}) \cong \log_2 \left| \mathbf{I} + \gamma \cdot \mathbf{S}(\mathbf{\Phi}) \tilde{\mathbf{H}}_w^H \mathbf{O} \tilde{\mathbf{H}}_w \right|. \quad (9)$$

We let L denote the number of nonzero diagonal elements of the PA matrix $\mathbf{\Phi}$ (equivalently, in a practical system L is the number of independent subchannels that is transmitted over the MIMO link). In (9) we have introduced the $R \times L$ matrix of i.i.d. complex Gaussian elements $\tilde{\mathbf{H}}_w$ and the $R \times R$ diagonal matrix of eigenvalues associated to the receive side

$$\mathbf{O} \equiv \text{eig}(\tilde{\mathbf{R}}_{nn}^{-1} \mathbf{R}_{RX}) = \text{diag}(o_1, \dots, o_R), \quad (10)$$

which comprises the effects of receive fading correlation and colored additive Gaussian noise. Finally, $\mathbf{S}(\mathbf{\Phi})$ is a $L \times L$ diag-

onal matrix, which reads for the optimum transmission strategy in (8)

$$\begin{aligned} \mathbf{S}(\mathbf{\Phi}) &= \text{eig}_0(\tilde{\mathbf{R}}_{ss} \mathbf{R}_{TX}) \\ &= \text{eig}_0(\mathbf{\Lambda}_{TX} \cdot \mathbf{\Phi}^2) = \text{diag}(s_1(\mathbf{\Phi}), \dots, s_L(\mathbf{\Phi})) \end{aligned}, \quad (11)$$

which takes into account PA via $\mathbf{\Phi}$ (on the strongest L eigenmodes) and fading correlation at the transmitter. Note that we can alternatively formulate

$$I(s, y, \mathbf{\Phi}) \cong \log_2 \left| \mathbf{I} + \gamma \cdot \mathbf{S}(\mathbf{\Phi}) \tilde{\mathbf{H}}_w^H \mathbf{O} \tilde{\mathbf{H}}_w \right| = \log_2 \left| \mathbf{I} + \gamma \cdot \mathbf{O} \tilde{\mathbf{H}}_w \mathbf{S}(\mathbf{\Phi}) \tilde{\mathbf{H}}_w^H \right|. \quad (12)$$

We rewrite (12) such that the matrix argument $\mathbf{S}(\mathbf{\Phi}) \tilde{\mathbf{H}}_w^H \mathbf{O} \tilde{\mathbf{H}}_w$ or $\mathbf{O} \tilde{\mathbf{H}}_w \mathbf{S}(\mathbf{\Phi}) \tilde{\mathbf{H}}_w^H$, respectively, of the determinant is of full rank, thereby simplifying the subsequent analysis. To this end, we define

$$\mu \equiv \min(R, L) \quad \nu \equiv \max(R, L), \quad (13)$$

the $\mu \times \mu$ diagonal matrix

$$\mathbf{\Sigma}(\mathbf{\Phi}) \equiv \begin{cases} \mathbf{S}(\mathbf{\Phi}) & R \geq L \\ \mathbf{O} & L > R \end{cases} \quad \mathbf{\Sigma}(\mathbf{\Phi}) = \text{diag}(\sigma_1, \dots, \sigma_\mu), \quad (14)$$

the $\nu \times \nu$ diagonal matrix

$$\mathbf{\Omega}(\mathbf{\Phi}) \equiv \begin{cases} \mathbf{O} & R \geq L \\ \mathbf{S}(\mathbf{\Phi}) & L > R \end{cases} \quad \mathbf{\Omega}(\mathbf{\Phi}) = \text{diag}(\omega_1, \dots, \omega_\nu), \quad (15)$$

and the $\nu \times \mu$ matrix of i.i.d. complex Gaussian entries \mathbf{G} . With above definitions we can introduce a unifying expression for MIMO mutual information, which will serve as a basis for all following derivations

$$I(s, y, \mathbf{\Phi}) \cong \log_2 \left| \mathbf{I} + \gamma \cdot \mathbf{\Sigma}(\mathbf{\Phi}) \mathbf{G}^H \mathbf{\Omega}(\mathbf{\Phi}) \mathbf{G} \right|. \quad (16)$$

B. Exact mean mutual information

The authors have calculated the exact mean mutual information in [13] and [14] via a novel moment generating function approach.

Theorem 1. The mean mutual information $M(\mathbf{\Phi}) = E[I(s, y, \mathbf{\Phi})]$ with $I(s, y, \mathbf{\Phi})$ according to (16) of a correlated MIMO system with transmit as well as receive correlation is given by

$$M(\mathbf{\Phi}) = \frac{\Gamma_\nu(\nu) \cdot (-1)^{\frac{(\nu-\mu) \cdot (\nu-\mu-1)}{2}}}{\ln 2 \cdot \alpha_\mu(\mathbf{\Sigma}(\mathbf{\Phi})) \cdot \alpha_\nu(\mathbf{\Omega}(\mathbf{\Phi})) \cdot \gamma^{\frac{\nu \cdot (\nu-1)}{2}}} \cdot \sum_{l=1}^{\mu} \left| \frac{\mathbf{\Xi}(l, \mathbf{\Phi})}{\mathbf{\Psi}(\mathbf{\Phi})} \right| \quad (17)$$

with the $\mu \times \nu$ matrices (row index i runs from 1 to μ and column index j from 1 to ν)

$$\mathbf{\Xi}(l, \mathbf{\Phi}) = \begin{cases} \Gamma(\nu) \cdot \sigma_i^{\mu-1} \cdot (\gamma \omega_j)^{\nu-1} \cdot e^{\frac{1}{\gamma \sigma_i \omega_j}} \cdot E_i\left(\frac{1}{\gamma \sigma_i \omega_j}\right) & i = l \\ \sum_{k=\nu-\mu}^{\nu-1} (-1)^k \cdot (\sigma_i)^{k-(\nu-\mu)} \cdot (\gamma \omega_j)^k \cdot [1-\nu]_k & i \neq l \end{cases} \quad (18)$$

and $(\nu-\mu) \times \nu$ matrix $\mathbf{\Psi}(\mathbf{\Phi})$ defined by (row index i' runs from 1 to $\nu-\mu$ and column index j' from 1 to ν)

$$\mathbf{\Psi}(\mathbf{\Phi}) = \left[(-\gamma \omega_{j'})^{i'-1} \cdot [-\nu+1]_{i'-1} \right]. \quad (19)$$

We note that the σ_i and ω_j are implicitly a function of Φ via (14) and (15). $E_1(z)$ is the exponential integral (see [11]), Pochhammer's symbol denotes

$$[a]_k = a \cdot (a+1) \cdot \dots \cdot (a+k-1) \quad [a]_0 = 1, \quad (20)$$

the definition

$$\Gamma_m(r) = \prod_{i=1}^m \Gamma(r-i+1), \quad (21)$$

and the Vandermonde determinant of a diagonal $m \times m$ matrix $\mathbf{X} = \text{diag}(x_1, \dots, x_m)$ can be expressed as

$$\alpha_m(\mathbf{X}) = \prod_{i < j} (x_i - x_j). \quad (22)$$

Proof: See [14].

C. High SNR asymptotics of mean mutual information

Via certain series expansions we can derive from Theorem 1 the following asymptotical result

Theorem 2. The MMI asymptotics $\bar{M}(\Phi)$ of a channel with correlation at TX and RX at high SNR with arbitrary number of transmit and receive antennas are given by

$$\bar{M}(\Phi) = \frac{-\mu \cdot E + \ln|\Sigma(\Phi)| + \zeta_1(\Omega(\Phi)) - \zeta_2(\mu, \nu)}{\ln 2} + \mu \cdot \log_2 \gamma. \quad (23)$$

where $E \approx 0.5772156649$ is Euler's constant, $\zeta_1(\Omega)$ is an auxiliary function depending on $\Omega(\Phi)$ (in the $\nu \times \nu$ determinants, i and j run from 1 to ν)

$$\zeta_1(\Omega(\Phi)) = \frac{1}{\alpha_\nu(\Omega(\Phi))} \cdot \sum_{l=1}^{\mu} \left| \begin{array}{cc} \omega_j^{\nu-i} \cdot \ln \omega_j & i = l \\ \omega_j^{\nu-i} & i \neq l \end{array} \right| \quad (24)$$

and the auxiliary function

$$\zeta_2(\mu, \nu) = \sum_{l=\nu-\mu}^{\nu-2} a_{E, \nu-1-l} \cdot \Gamma(\nu-l) \quad (25)$$

with

$$a_{E, k} = \sum_{l=1}^k \frac{(-1)^l}{l \cdot l!} \cdot \frac{1}{(k-l)!}. \quad (26)$$

Proof: See [16].

We now consider some special cases of Theorem 2.

Corollary 1. The MMI high SNR asymptotics of a MIMO system with semi-correlated channel and $\Omega = \mathbf{I}$ are given by

$$\bar{M}(\Phi) = \frac{1}{\ln 2} \cdot \left(\ln(|\Sigma(\Phi)|) + \sum_{k=0}^{\mu-1} \psi(\nu-k) \right) + \mu \cdot \log_2 \gamma, \quad (27)$$

where $\psi(x)$ is the Digamma function [11].

Proof: See [16].

D. Low SNR asymptotics of mean mutual information

In the low SNR regime the mean mutual information expression simplifies considerably and we get

Theorem 3. The MMI $\underline{M}(\Phi)$ of an arbitrarily correlated MIMO channel in the low SNR region is given by

$$\underline{M}(\Phi) = \frac{\text{tr}(\Sigma(\Phi)) \cdot \text{tr}(\Omega(\Phi))}{\ln 2} \cdot \gamma. \quad (28)$$

Proof: See [16].

IV. ERGODIC CAPACITY WITH CDIT

Based on the mean mutual information expressions given in the last section, we now study the behavior of the capacity achieving power allocation strategy, whereas it turns out that closed-form PA formulas can be given only for special cases.

A. General problem

The ergodic capacity $C_{\text{erg}}^{\text{CDIT}}(\gamma)$ with channel distribution information at the transmitter (CDIT) in a correlated Rayleigh fading environment is given by ('s.t.' stands for 'subject to')

$$C_{\text{erg}}^{\text{CDIT}}(\gamma) = \max_{\Phi} M(\Phi) \quad \text{s.t. } \text{tr}(\Phi\Phi^H) = \rho \quad (29)$$

with transmit power constraint ρ , i.e. we have to maximize mean mutual information with respect to the power allocation matrix. It is clear that due to the complexity of the MMI expression in Theorem 1, a solution to the constrained optimization problem in (29) in general cannot be given in closed form. Therefore, we have to resort to numerical optimization algorithms and results of this optimization process will be presented in the numerical result section below.

B. High SNR regime

From the asymptotic formulas in Theorem 2 and Corollary 1, one can notice that MMI is proportional to $\mu = \min(L, R)$ for high SNR. Therefore, in case of $T \leq R$ the transmit prefilter should exploit all spatial dimensions at the transmitter side with $L=T$, such that Φ should be chosen to be of full rank $\text{rk}(\Phi) = T$. With a full rank Φ , from Theorem 2 together with the optimal transmission strategy (8) it is clear that the ergodic capacity with CDIT in the high SNR regime with $T \leq R$ is given by

$$\bar{C}_{\text{erg}, T \leq R}^{\text{CDIT}} = \bar{\zeta}_{C, T \leq R} + \max_{\Phi} \log_2(|\Phi|^2) \quad \text{s.t. } \text{tr}(\Phi\Phi^H) = \rho, \quad (30)$$

with a constant $\bar{\zeta}_{C, T \leq R}$, which is independent of Φ . However, the determinant $|\Phi|^2$ is a Schur-concave function [12] of the squared power allocation coefficients ϕ_k^2 , such that in the high SNR regime we have the optimal power allocation strategy $\bar{\Phi}_{\text{opt}, T \leq R} = \mathbf{I}_T$, i.e. essentially no waterfilling takes place and therefore the absolute capacity gain $\bar{\Delta C}_{T \leq R}$ with CDIT due to waterfilling is zero

$$\bar{\Delta C}_{T \leq R}^{\text{WF}} = \bar{C}_{\text{erg}, T \leq R}^{\text{CDIT}} - \bar{C}_{\text{erg}, T \leq R} = 0, \quad (31)$$

where $\bar{C}_{erg, T \leq R}$ is the capacity with uninformed transmitter for $T \leq R$ in the high SNR regime.

On the other hand, for $T > R$ the slope of the high SNR asymptotics in Theorem 2 are now determined by μ and ν . Let $L \geq R$, and in general we have to resort to numerical optimization for finding the optimal PA matrix Φ that achieves capacity. Specifically, from Theorem 2 together with (11) it can be seen that in the high SNR regime the optimum $\Phi_{opt, T > R}$ is given by

$$\bar{\Phi}_{opt, T > R} = \arg \max_{\Phi} \left[\zeta_1(\text{eig}_0(\Lambda_{TX} \cdot \Phi^2)) - \zeta_2(R, rk(\Phi)) \right], \quad (32)$$

s.t. $\text{tr}(\Phi^2) = \rho$

with the definitions of the auxiliary functions $\zeta_1(\Omega(\Phi))$ and $\zeta_2(\mu, \nu)$ in (24) and (25). The absolute gain due to statistical waterfilling with $\bar{\Phi}_{opt, T > R}$ follows from Theorem 2

$$\frac{\bar{C}_{T > R}^{WF}}{\Delta C_{T > R}^{WF}} = \frac{1}{\ln 2} \cdot \left[\zeta_1(\text{eig}_0(\Lambda_{TX} \cdot (\bar{\Phi}_{opt, T > R})^2)) - \zeta_2(R, rk(\bar{\Phi}_{opt, T > R})) \right] - \left[\zeta_1(\Lambda_{TX}) - \zeta_2(R, T) \right]. \quad (33)$$

Again, optimization problem (32) is in general not solvable in closed form and we therefore resort to numerical optimization algorithms. Results of the optimization will be presented at the end of this paper.

C. Low SNR regime

With the optimal transmission strategy in (8), from Theorem 3 the ergodic capacity with channel state information at the transmitter in the low SNR region is given by

$$\underline{C}_{erg}^{CDIT}(\gamma) = \max_{\Phi} \frac{\text{tr}(\Lambda_{TX} \cdot \Phi^2) \cdot \text{tr}(\mathbf{O})}{\ln 2} \cdot \gamma \quad \text{s.t. } \text{tr}(\Phi \Phi^H) = \rho. \quad (34)$$

It is straightforward to see that $\text{tr}(\Lambda_{TX} \cdot \Phi^2)$ is Schur-concave in the power allocation coefficients ϕ_k^2 and therefore the optimal capacity-achieving transmission strategy in the low SNR regime is beamforming, where all transmit power is put on the strongest long-term eigenmode of the channel via $\Phi_{opt} = \text{diag}(\sqrt{\rho}, 0, \dots, 0)$. With the normalization $\text{tr}(\mathbf{R}_{TX}) = \text{tr}(\Lambda_{TX}) = \rho$ and the power constraint $\rho = T$, the relative waterfilling gain \underline{G}_C^{WF} can now be calculated from (34)

$$\underline{G}_C^{WF} = \frac{\bar{C}_{erg}^{CDIT}}{\bar{C}_{erg}} = \frac{\lambda_{TX,1} \cdot \rho}{\text{tr}(\Lambda_{TX})} = \frac{\lambda_{TX,1} \cdot T}{T} = \lambda_{TX,1}, \quad (35)$$

where $\lambda_{TX,1}$ is the maximum eigenvalue of the transmit correlation matrix.

V. LOW COMPLEXITY PA SCHEMES

In this section, we study two PA schemes that are derived from upper bounds on mean mutual information.

A. Power allocation based on tight bound

By exploiting Jensen's inequality, it can be shown that mean mutual information can be tightly upper bounded by

$$M(\Phi) \leq M^B(\Phi) = \log_2 E[\mathbf{I} + \gamma \cdot \Sigma(\Phi) \mathbf{G}^H \Omega(\Phi) \mathbf{G}]. \quad (36)$$

The expected value in (36) can be calculated in closed form by exploiting certain expected values with respect to the Wishart distribution.

Theorem 4. A tight upper bound on MMI can be given in terms of the elementary symmetric functions

$$M^B(\Phi) = \log_2 \left(\sum_{k=0}^{\mu} k! \cdot \gamma^k \cdot \text{tr}_k(S(\Phi)) \cdot \text{tr}_k(\mathbf{O}) \right), \quad (37)$$

where $\text{tr}_k(\mathbf{X})$ denotes the k th order elementary symmetric function in the eigenvalues of the matrix \mathbf{X} [12].

Proof: See [15].

Based on the tight bound on MMI given in Theorem 4, a reduced-complexity PA scheme can be found, where the PA matrix is analogously to (29) defined by the constrained optimization problem

$$\Phi_B = \arg \max_{\Phi} \log_2 \left(\sum_{k=0}^{\mu} k! \cdot \gamma^k \cdot \text{tr}_k(S(\Phi)) \cdot \text{tr}_k(\mathbf{O}) \right) \quad (38)$$

s.t. $\text{tr}(\Phi \Phi^H) = \rho$

We first shall restrict ourselves to the $2 \times R$ antenna case to allow for a concise analytical solution of (38). Omitting details, the diagonal elements of Φ_B^2 obtained from a Lagrange optimization process are

$$\phi_{B,1}^2 = \frac{\gamma^2 \cdot |\mathbf{R}_{TX}| \cdot \text{tr}_2(\mathbf{O}) \cdot \rho + \frac{\gamma}{2} \cdot (\lambda_{TX,1} - \lambda_{TX,2}) \cdot \text{tr}(\mathbf{O})}{2 \cdot \gamma^2 \cdot |\mathbf{R}_{TX}| \cdot \text{tr}_2(\mathbf{O})}, \quad (39)$$

$$\phi_{B,2}^2 = \frac{\gamma^2 \cdot |\mathbf{R}_{TX}| \cdot \text{tr}_2(\mathbf{O}) \cdot \rho + \frac{\gamma}{2} \cdot (\lambda_{TX,2} - \lambda_{TX,1}) \cdot \text{tr}(\mathbf{O})}{2 \cdot \gamma^2 \cdot |\mathbf{R}_{TX}| \cdot \text{tr}_2(\mathbf{O})},$$

where we have to assure that $\phi_{B,1/2} > 0$, otherwise the coefficient $\phi_{B,2}$ is set to 0 and all available transmit power is given to $\phi_{B,1}^2 = \rho$. Note that in agreement with results stated e.g. in [169], the total transmit power is equally distributed on both subchannels, i.e. $\phi_{B,1}^2 = \phi_{B,2}^2 = \rho/2$, if no transmit correlation is present, i.e. $\mathbf{R}_{TX} = \mathbf{I}$, no matter what receive correlation is prevailing. This result also holds for the general case with an arbitrary number of transmit antennas. To this end, note that for the case of receive correlation only we find from (38) the optimization problem in terms of the elementary symmetric functions

$$\Phi_{B,rx} = \arg \max_{\Phi} \log_2 \left[\sum_{k=0}^{\mu} k! \cdot \gamma^k \cdot \text{tr}_k(\Phi^2) \cdot \text{tr}_k(\mathbf{O}) \right] \quad \text{s.t. } \text{tr}(\Phi^2) = \rho. \quad (40)$$

The objective function in (40) can be shown to be Schur-concave in the ϕ_k^2 with $\Phi = \text{diag}(\phi_1, \dots, \phi_T)$. The optimal power allocation strategy is thus uniform. This is in accordance with known results in literature [10] and intuition: if there is no transmit correlation present, there are no prominent directions and the transmitter equally distributes power.

For the general case of arbitrary array sizes and arbitrary transmit and receive correlation, a closed form solution of the problem in (38) can not be given. Numerical methods have to be applied to find the optimal PA coefficients. The results of

such numerical optimization processes will be shown in the simulations below.

B. Power allocation based on two loose bounds

In this paragraph, we study two approximate low-complexity PA schemes based on other bounds on MMI that can also be derived via Jensen's inequality and the concavity of the log det function. The first loose bound takes into account only the transmit correlation matrix \mathbf{R}_{TX} and reads [76]

$$\tilde{M}^{B,1}(\Phi) = \log_2 \left| \mathbf{I} + \gamma \cdot \tilde{\mathbf{R}}_{ss}(\Phi) \cdot E[\mathbf{H}^H \cdot \tilde{\mathbf{R}}_{nn}^{-1} \cdot \mathbf{H}] \right|, \quad (41)$$

which can be written as

$$\tilde{M}^{B,1}(\Phi) = \log_2 \left| \mathbf{I} + \gamma \cdot \text{tr}(\tilde{\mathbf{R}}_{nn}^{-1} \mathbf{R}_{RX}) \cdot \Phi^2 \mathbf{\Lambda}_{TX} \right|. \quad (42)$$

Based on this loose bound, again a low-complexity PA scheme can be derived

$$\tilde{\Phi}_{B,1} = \arg \underbrace{\max}_{\Phi} \log_2 \left| \mathbf{I} + \gamma \cdot \text{tr}(\tilde{\mathbf{R}}_{nn}^{-1} \mathbf{R}_{RX}) \cdot \Phi^2 \mathbf{\Lambda}_{TX} \right| \quad (43)$$

s.t. $\text{tr}(\Phi \Phi^H) = \rho$

In contrast to the statistical power allocation schemes presented above, problem (43) can be solved in closed form by the standard waterfilling algorithm (e.g. [4]) for an arbitrary number of transmit and receive antenna elements. We emphasize that the effects of the receive correlation matrix \mathbf{R}_{RX} are basically not captured by this approximate scheme and numerical results underpin this observation. However, the bound in (41) is exact for a large number of antenna elements R due to the channel hardening effect [7], which means that the stochastic matrix $\mathbf{H}^H \cdot \tilde{\mathbf{R}}_{nn}^{-1} \cdot \mathbf{H}$ approaches its expected value by the law of large numbers.

On the other hand, paralleling (41), a second loose bound on MMI is given by

$$\tilde{M}^{B,2}(\Phi) = \log_2 \left| \mathbf{I} + \gamma \cdot \tilde{\mathbf{R}}_{nn}^{-1} \cdot E[\mathbf{H} \cdot \tilde{\mathbf{R}}_{ss}(\Phi) \cdot \mathbf{H}^H] \right|, \quad (44)$$

which can be written as

$$\tilde{M}^{B,2}(\Phi) = \log_2 \left| \mathbf{I} + \gamma \cdot \text{tr}(\Phi^2 \mathbf{\Lambda}_{TX}) \cdot \tilde{\mathbf{R}}_{nn}^{-1} \mathbf{R}_{RX} \right|. \quad (45)$$

Similar to the optimization problem in (34), it is clear that the optimum PA scheme in the sense of maximizing the loose MMI bound $\tilde{M}^{B,2}(\Phi)$, namely

$$\tilde{\Phi}_{B,2} = \arg \underbrace{\max}_{\Phi} \log_2 \left| \mathbf{I} + \gamma \cdot \text{tr}(\Phi^2 \mathbf{\Lambda}_{TX}) \cdot \tilde{\mathbf{R}}_{nn}^{-1} \mathbf{R}_{RX} \right|, \quad (46)$$

s.t. $\text{tr}(\Phi \Phi^H) = \rho$

is again given by a beamforming solution $\tilde{\Phi}_{B,2} = \text{diag}(\sqrt{\rho}, 0, \dots, 0)$. The bound on MMI in (45) becomes tight and the corresponding PA policy becomes exact by the channel hardening effect for a large number of transmit antennas T .

A. Simulation setup

Without loss of generality, we study systems with white input signals of power E_s and additive white Gaussian noise of variance N_0 (other signal and noise covariances can easily be absorbed in an equivalent channel), i.e.

$$\mathbf{R}_{ss} = E_s \cdot \mathbf{I} \quad \mathbf{R}_{nn} = N_0 \cdot \mathbf{I}. \quad (47)$$

Furthermore, in the following, we consider exponential correlation matrices at the receiver and the transmitter with

$$[\mathbf{R}_{RX/TX}]_{ij} = (r_{RX/TX})^{|i-j|}, \quad (48)$$

i.e. r_{RX} is the correlation coefficient between two neighboring receive antennas and r_{TX} models the correlation between two transmit antennas. We note that with the given channel model, the correlation between two antennas decreases exponentially with their distance. Finally, the SNR in dB is defined by

$$\gamma_{dB} \equiv 10 \cdot \log_{10} \frac{\rho \cdot E_s}{N_0} \quad [dB], \quad (49)$$

whereas the transmit power constraint is set to $\rho = T$.

B. Exact ergodic capacity with CDIT

Plots of the ergodic capacity (EC) for a system with $T=4$ transmit and $R=6$ receive antennas are given in Fig. 1. The channel exhibits strong fading correlation at the transmitter with $r_{TX}=0.97$, while the fading at the receiver is only weakly correlated with $r_{RX}=0.3$.

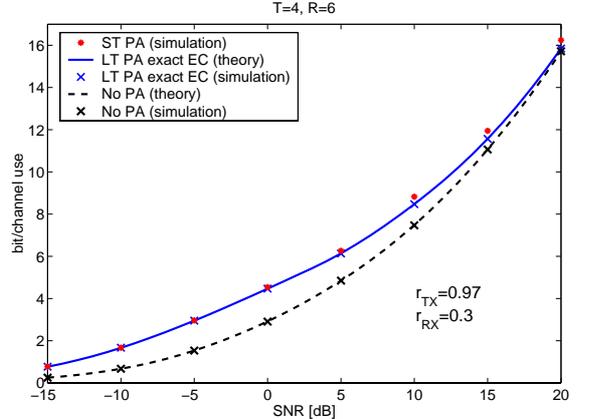


Fig. 1: EC with PA, $T=4$, $R=6$, $r_{TX}=0.97$, $r_{RX}=0.3$

We compare three different transmission strategies. Without channel state information (CSI) at the transmitter, the optimal transmission strategy is to transmit without a special power allocation (i.e. power is distributed uniformly over the transmit antennas), which yields the lowest capacity. With the availability of CDIT, the transmitter can apply long-term (LT) statistical power allocation (PA) according to optimization problem in (29), which can be solved numerically. For this heavily correlated scenario, one can observe a significant increase in ergodic

capacity (EC) due to CDIT especially in the lower SNR region. Finally, with the availability of instantaneous CSI at the transmitter we can apply optimum short-term (ST) power allocation (waterfilling) [35], which as expected achieves the highest ergodic capacity. However, we note that for this highly correlated MIMO channel, the difference between ST PA and LT PA is only minor due to the fact that only a small number (in the limit only one) of eigenmodes can be effectively used for data transmission in the low SNR regime. We note that in the high SNR regime all transmission schemes in general yield the same ergodic capacities. Furthermore, in Fig. 1 a close agreement between theoretical curves according to Theorem 1 and Monte-Carlo simulation results can be observed.

With less fading correlation ($r_{TX}=0.7$) at the transmitter, the ergodic capacity increase due to adequate power allocation at the transmitter is reduced (Fig. 2). Moreover, it can be seen that the SNR range, where CSI at the transmitter is beneficial, is limited to lower SNR values. At the same time, there is now a noticeable difference between the case of ST CSI and CDIT. Due to the reduced correlation, it is now obviously more likely that a higher number of channel eigenmodes is used for data transmission and this can be exploited, if the transmitter is aware of the instantaneous channel state.

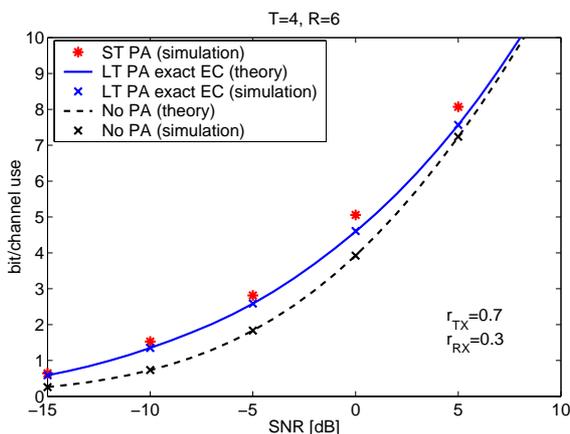


Fig. 2: EC with PA, $T=4$, $R=6$, $r_{TX}=0.7$, $r_{RX}=0.3$

C. Low complexity approximate power allocation

The effectiveness of the various statistical PA strategies that were presented above is studied in Fig. 3 for a system with $T=R=4$. On the transmitter side, we assume strong fading correlation with $r_{TX}=0.9$ and the receiver side is correlated with $r_{RX}=0.7$.

For validating our theoretical results, we have plotted Monte-Carlo simulation results and theoretical curves according to the analysis presented in Theorem 1. Again, there is a perfect agreement between theory and simulation. As expected, again the PA schema based on short-term CSI performs best, while the performance of the long-term PA schemes with CDIT comes close to the optimum at low SNR. Furthermore, we can observe that LT PA based on the exact EC analysis according to (29) is superior than the scheme based on the

loose MMI bound according to (43) in the medium SNR range from 10 to 20 dB. We emphasize that the EC curve for the PA scheme based on the tight bound in (40) cannot be distinguished from the exact CDIT based LT PA scheme and we have therefore removed it from Fig. 3 for clarity.

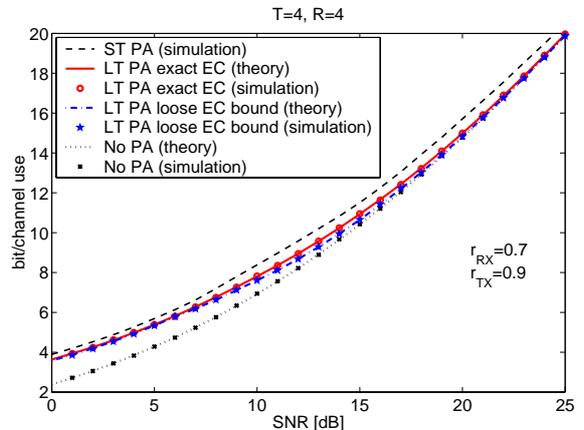


Fig. 3: EC with PA, $T=R=4$, $r_{RX}=0.7$, $r_{TX}=0.9$

More insight can be obtained by looking at the power allocation coefficients of the matrix $\Phi = \text{diag}(\phi_1, \dots, \phi_T)$, which are depicted in Fig. 4. Obviously, the PA coefficients resulting from the PA allocation algorithm based on the loose bound in (43) deviate significantly from the other two schemes, which are very close together. Specifically, additional eigenmodes are activated already at lower SNR, thus leading to a degradation of mean mutual information.

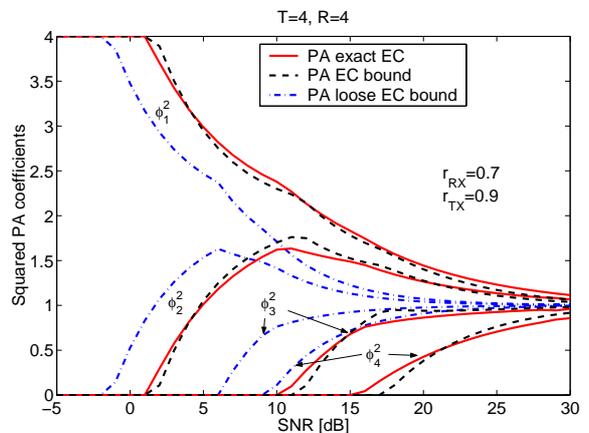


Fig. 4: PA coefficients, $T=R=4$, $r_{RX}=0.7$, $r_{TX}=0.9$

At high SNR, however, all three schemes lead to a uniform power allocation, which has been predicted in the derivation of (31), such that all eigenmodes are equally used for transmission and thus the statistical waterfilling gain vanishes at higher SNR. Equivalently, in the sense of maximizing EC, CDIT is useless in the high SNR regime for this particular system, where $T=R=4$. However, in contrast to that, we note that below we present results for a system with $T>R$, where a capacity

gain can be achieved even in the high SNR region (see also (33)).

D. High and low SNR regime

A more detailed view on the relative gain of statistical water-filling (compared to a system without CSI at the transmitter) is given in Fig. 5 for a $T=R=4$ system with correlation at the receiver $r_{RX}=0.7$ and varying correlation at the transmitter $r_{TX}=\{0.5, 0.7, 0.9, 0.97\}$.

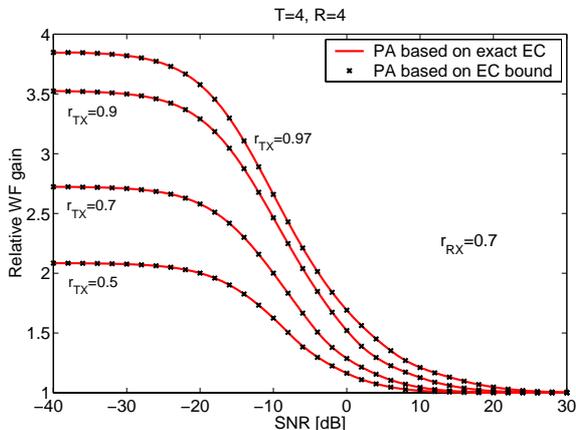


Fig. 5: Relative WF gain, $T=R=4$, $r_{RX}=0.7$, $r_{TX}=\{0.5, 0.7, 0.9, 0.97\}$

A close agreement between the WF scheme based on the exact MMI analysis according to (29) and the WF scheme based on the tight MMI bound according to (40) can be observed. Furthermore, in the low SNR regime the theoretical result in (35) is confirmed, i.e. statistical waterfiling achieves a relative gain equal to the largest eigenvalue of the transmit correlation matrix \mathbf{R}_{TX} . With increasing transmit correlation (increasing r_{TX}), the largest eigenvalue of the transmit correlation matrix increases and thus the relative WF gain increases. The maximum eigenvalues of the transmit correlation matrix for the different correlation coefficients are

$$\lambda_{TX,1} = \{2.08, 2.72, 3.53, 3.85\}. \quad (50)$$

At high SNR, we note that the relative capacity gain converges to 1, which agrees with (31). A closer look on the performance of a WF scheme based on the loose capacity bound according to (43) is taken in Fig. 6 for the same system parameters as in Fig. 5. While the asymptotical performance in the low and high SNR regime is the same as for the WF scheme based on the exact MMI analysis, it is obviously inferior in the medium SNR region. The loss in capacity is in the range of 3 to 4 percent for the given scenario. We note that this loss is higher for strong fading correlation at the receiver, as the loose capacity bound in (42) does not take into account correlation at the receive antenna array.

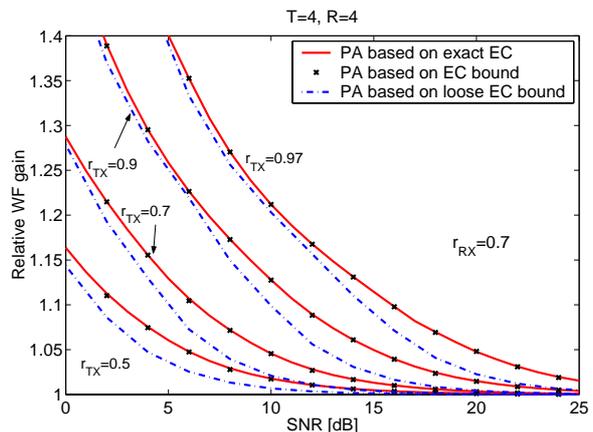


Fig. 6: Relative WF gain, $T=R=4$, $r_{RX}=0.7$, $r_{TX}=\{0.5, 0.7, 0.9, 0.97\}$

The absolute gain in bit per channel use due to statistical waterfiling based on the exact MMI analysis in the high SNR regime (see also (3.116)) is depicted in Fig. 7 for a system with $R=2$ receive antennas and $r_{RX}=0.7$ (we compare the ergodic capacity with a system with uninformed transmitter). As expected, the gain increases with a higher number of transmit antennas and with stronger correlation at the transmit side. Basically, in the presence of CDIT, the transmitter can make use of the beamforming capabilities of the transmit antenna array for improving the capacity of the link.

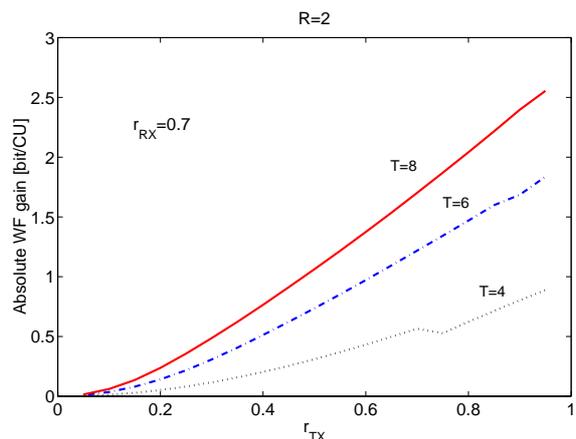


Fig. 7: Absolute WF gain, high SNR, $T=\{4, 6, 8\}$, $R=2$, $r_{RX}=0.7$

In Fig. 8, the squared PA coefficients are depicted for a system with $T=8$ transmit and $R=2$ receive antenna elements for $r_{RX}=0.7$ in the high SNR regime.

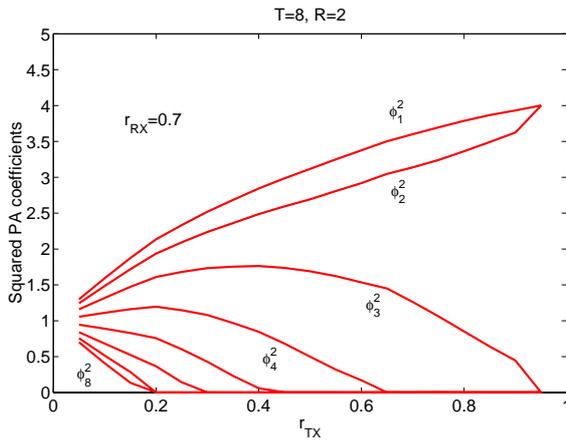


Fig. 8: PA coefficients, high SNR, $T=8$, $R=2$, $r_{RX}=0.7$, various r_{TX}

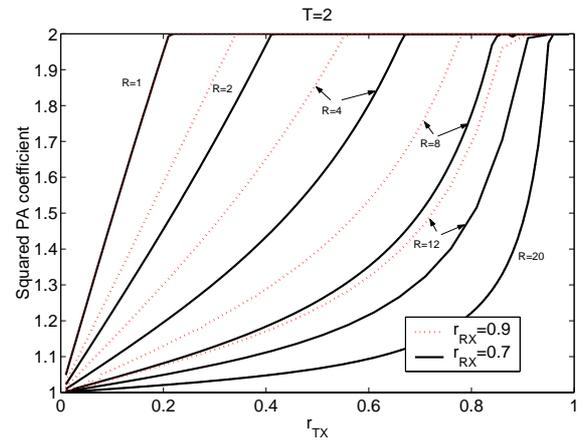


Fig. 9: Squared PA coefficients, $T=2$, $SNR=0$ dB, $r_{RX}=\{0.7, 0.9\}$

While for vanishing transmit correlation ($r_{TX}=0$) power is equally distributed on all 8 transmit eigenmodes of the MIMO channel, with strong fading correlation at the transmitter ($r_{TX} \rightarrow 1$) the overall transmit power $\rho = 8$ is assigned to a decreasing number of the strongest eigenmodes.

E. Large number of receive antennas

Due to the channel hardening effect, for a large number of receive antennas $R \gg T$ the MMI bound in (42) becomes exact and MMI essentially becomes independent of fading correlation at the receive antenna array. However, this means that also the power allocation coefficients become independent of correlation at the receiver. This is confirmed by the numerical results in Fig. 9, where we depict the power allocation coefficient ϕ_1^2 for a system with $T=2$ transmit antennas and a varying number of receive antennas resulting from a waterfilling strategy based on the exact EC analysis according to (29). We note that the second coefficient is implicitly given by $\phi_2^2 = \rho - \phi_1^2 = 2 - \phi_1^2$ due to the power constraint $\rho = T = 2$. While there is a significant discrepancy of the PA coefficient ϕ_1^2 for $R=8$ and $R=12$ receive antennas for a receive correlation coefficient of $r_{RX}=0.7$ and $r_{RX}=0.9$, the statistical PA algorithm essentially yields the same PA coefficients for $R=20$ receive antenna elements, independent of the receive fading correlation. Moreover, it can be seen that for a large number of receive antennas, a uniform PA is optimal for a wide range of transmit correlation values r_{TX} , i.e. CDIT has basically no effect on ergodic capacity.

REFERENCES

- [1] Visotsky E., Madhow U., "Space-time transmit precoding with imperfect feedback," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2632-2639, Sept. 2001
- [2] Jafar S., Goldsmith A., "Transmitter optimization and optimality of beamforming for multiple antenna systems with imperfect feedback", *IEEE Transactions on Wireless Communications*, 2003, *submitted*
- [3] Jorswieck E., Boche H., "Channel capacity and capacity-range of beamforming in MIMO wireless systems under correlated fading with covariance feedback," *IEEE Journal on Selected Areas in Communications*, 2003, *submitted*
- [4] Farokhi F. R., Foschini G. J., Lozano A., Valenzuela R. A., "Link-optimal space-time processing with multiple transmit and receive antennas," *IEEE Communications Letters*, vol. 5, no. 3, March 2001
- [5] Simon S. H., Moustakas A. L., "Optimizing MIMO antenna systems with channel covariance feedback," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 3, pp. 406-417, April 2003
- [6] Ivrlac M. T., Kurpjuhn T. P., Brunner C., Utschick W., "Efficient use of fading correlations in MIMO systems," *VTC*, Oct. 2001
- [7] Hochwald B. M., Marzetta T. L., Tarokh V., "Multi-antenna channel-hardening and its implications for rate feedback and scheduling," *IEEE Transactions on Information Theory*, 2002, *submitted*, available at <http://mars.bell-labs.com>
- [8] Kermaol J. P., Schumacher L., Pedersen K. I., Mogensen P. E., Frederiksen F., "A stochastic MIMO radio channel model with experimental validation," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1211-1226, Aug. 2002
- [9] Paulraj A., Nabar R., Gore D., *Introduction to space-time wireless communications*, Cambridge University Press, 2003
- [10] Telatar I. E., "Capacity of multi-antenna Gaussian channels," *Bell Labs Technical Memorandum*, June 1995
- [11] Abramowitz M., Stegun I. A., *Handbook of mathematical functions*, Dover Publications Inc., New York 1964
- [12] Marshall A. W., Olkin I., *Inequalities: Theory of majorization and its applications*, Academic Press, 1979
- [13] Kiessling M., Speidel J., "Mutual information of MIMO channels in correlated Rayleigh fading environments - a general solution", *ICC*, June 2004
- [14] Kiessling M., Speidel J., "Exact ergodic capacity of MIMO channels in correlated Rayleigh fading environments," *International Zurich Seminar on Communications*, Feb. 2004
- [15] Kiessling M., Viering I., Reinhardt M., Speidel J., "A closed form bound on correlated MIMO channel capacity," *VTC*, Sept. 2002
- [16] Kiessling M., Speidel J., "Asymptotics of ergodic MIMO capacity in correlated Rayleigh fading environments," *VTC*, May 2004