Fading Correlations in Wireless MIMO Communication Systems

Michel T. Ivrlač, Wolfgang Utschick, and Josef A. Nossek, *Fellow, IEEE*

***Abstract*—We investigate the effects of fading correlations on wireless communication systems employing multiple antennas at both the receiver and the transmitter side of the link, so called multiple-input multiple-output (MIMO) systems. It turns out that the amount of transmitter sided channel knowledge plays an important part when dealing with fading correlations. Furthermore, the possible availability of time diversity in a time-selective channel can have essential influence on performance. To study the influence of time-selectivity, the concept of *sample-mean outage* is introduced and applied to information theoretic measures, like capacity or cutoff rate. It will be shown, that in some cases correlated fading may offer better performance than uncorrelated fading permits, which is due to exploitable antenna gain, that will also be defined in a general form for MIMO systems.**

***Index Terms*—Eigenbeamforming, fading correlation, long-term channel knowledge, MIMO capacity, multiple-input multiple-output (MIMO) antenna gain, sample-mean outage cutoff rate.**

I. Introduction

MULTIPLE-INPUT multiple-output (MIMO) communication systems recently have drawn considerable attention in the area of wireless communications as they promise huge capacity increase. If the fading between pairs of transmit and receive antennas is independently Rayleigh distributed, it is well known [4], [14], [16], that in the high transmit power region the average capacity increases linearly with the minimum number of transmit and receive antennas, even if the transmitter has no knowledge of the channel. However, in a real world scenario the fades are usually not independent, but will exhibit certain fading correlations. It has been observed [3], [15], that channel capacity degrades significantly in the presence of fading correlations. However, these observations were built on the assumption of having zero transmitter channel knowledge and no other source of diversity, like time or frequency, available. In this paper, we like to point out that, allowing the transmitter to know the channel *on average*, correlated fading can be used in advantage and actually may lead to higher channel capacity than uncorrelated fading would permit, the more so, if time or frequency diversity are available to some degree. This is a more general conjecture than in [10], where optimality of beamforming in the low signal-to-noise ratio (SNR) region was proved. After introducing the system model, we will define and discuss the impact of fading correlations, channel time-selectivity and transmitter channel knowledge on information theoretic measures like channel capacity and cutoff rate. A scheme will be proposed that makes efficient use of present fading correlations and turns their existence from curse into blessing. We will also consider the effects of linear modulation schemes on system performance by cutoff-rate analysis.

Manuscript received May 1, 2002; revised October 31, 2002.

The authors are with the Institute for Circuit Theory and Signal Processing, Munich University of Technology, Munich, Germany (e-mail: ivrlac@nws.ei.tum.de; utschick@nws.ei.tum.de; nossek@nws.ei.tum.de).

Digital Object Identifier 10.1109/JSAC.2003.810348

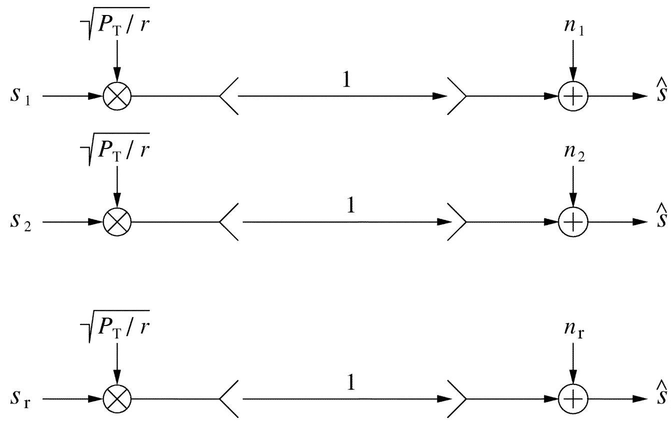


Fig. 1. A simple example of multistream transmission, where a number of r independent data streams share the total transmit power P_T and get transmitted over r constant channels with unity transmission gain and perturbed with Gaussian noise of variance σ_n^2.

II. Motivation

The fundamental concept underlying MIMO information theory is the idea of splitting a data stream into several streams transmitted in parallel over individual subchannels. Fig. 1 shows a simple example of a communication link comprised of r independent and equal additive white Gaussian noise (AWGN) subchannels. Because of symmetry, it is optimum to divide the available transmit power P_T equally between the r subchannels. The capacity C_r of this system is simply the sum of the individual subchannel capacities and, therefore, reads as

$$C_r = r \cdot \log_2 \left(1 + \frac{1}{r} \cdot \frac{P_T}{\sigma_n^2}\right) \tag{1}$$

where σ_n^2 is the variance of noise each subchannel perturbs the signal with. The relationship

$$\begin{aligned} &\log_2 \left(1 + \frac{P_T}{\sigma_n^2}\right) \\ &\quad = C_1 < C_2 < \cdots C_r < C_{r+1} < \cdots < C_\infty \\ &\quad = \frac{1}{\ln 2} \cdot \frac{P_T}{\sigma_n^2} \end{aligned} \tag{2}$$

clearly shows the capacity improving potential MIMO systems offer. Asymptotically, the capacity even becomes a linear function of transmit power. In this simple example, the number of data streams equals the number of transmit and receive antennas and all subchannels are independent. Such a MIMO channel is usually referred to as being diagonal. However, in practice MIMO channels rarely are diagonal, nor is the number of data streams equal to the number of transmit or receive antennas, as the algebraic rank r of the MIMO channel dictates the number of data streams which are supported for simultaneous transmission. When we refer to *rank deficient* channels, we see that relation (2) does *not* hold in general. Rank deficiency may in fact lead to larger capacity than permitted by a full rank channel, depending on the applied transmit power. However, before going into details, let us first define the system model we will be using throughout the text.

III. System Model

We will assume a frequency flat and possibly correlated Rayleigh-fading wireless channel, that is accessed by N transmit and M receive antennas to transmit L independent data streams, leading to the system model

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{T}\boldsymbol{P}^{1/2}\boldsymbol{s} + \boldsymbol{n} \tag{3}$$

where $\boldsymbol{s} \in \mathcal{C}^L$ is the L-dimensional data stream vector with zero mean and unity covariance matrix, while $\boldsymbol{P} \in \mathcal{R}_+^{L\times L}$ is a positive definite diagonal matrix used to set the transmit power for each data stream with total transmit power given by $P_T = \text{tr}\, \boldsymbol{P}$ and finally the matrix $\boldsymbol{T} \in \mathcal{C}^{N\times L}$ performs the mapping from L data streams onto N transmit antennas and is composed of unity norm column vectors. This mapping can be viewed as beamforming. The channel is modeled by the matrix $\boldsymbol{H} \in \mathcal{C}^{M\times N}$ with possibly correlated complex zero-mean Gaussian entries. The receive signal vector $\boldsymbol{y} \in \mathcal{C}^M$ is corrupted by additive zero-mean white Gaussian noise $\boldsymbol{n} \in \mathcal{C}^M$ with power σ_n^2 per receive antenna.

IV. Constant Channels

If the channel matrix $\boldsymbol{H}$ is constant at all times, we may assume it known to both the receiver and the transmitter. In such a scenario, it is straightforward to arrive at independent subchannels by using singular value decomposition

$$\begin{aligned}\boldsymbol{H} &= \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H,\ \boldsymbol{H}^H\boldsymbol{H} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^H\\ \boldsymbol{\Lambda} &= \boldsymbol{\Sigma}^T\boldsymbol{\Sigma} = \text{diag}\,\{\lambda_i\}_{i=1}^N,\ \lambda_i \ge \lambda_{i+1}.\end{aligned} \tag{4}$$

Setting $\boldsymbol{T} = \boldsymbol{V}$ and applying the one-to-one transformation $\hat{\boldsymbol{s}} = \boldsymbol{U}^H\boldsymbol{y}$ we arrive at independent subchannels: $\hat{s}_j = \sqrt{\lambda_j \cdot P_j} \cdot s_j + n'_j$, where n'_j is zero-mean Gaussian noise with variance σ_n^2, which is independent for different streams. The stream index ranges over $1 \le j \le r$, where $r = \text{rank}\, \boldsymbol{H} \le \min(M, N)$, as $\lambda_j = 0$ for $j > r$. Using the waterfilling [5] power distribution $P_j = \max(0, \mu - \sigma_n^2/\lambda_j)$, the channel capacity reads as [16]

$$C_r = \sum_{j=1}^{r} \log_2 \max\left(1, \mu \cdot \frac{\lambda_j}{\sigma_n^2}\right) \tag{5}$$

where μ is a positive constant chosen to fulfill the power constraint $\text{tr}\, \boldsymbol{P} = P_T$.

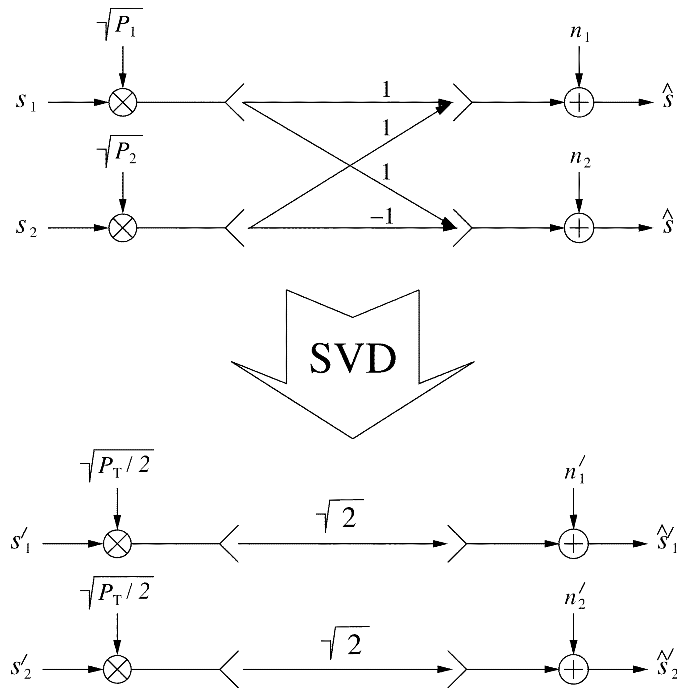


Fig. 2. A constant *full rank* MIMO channel (top) and its diagonalization (bottom).

V. Rank Deficient Channels

From (5) and (4) it is clear, that channel capacity depends on the eigenvalue profile of the Gramian matrix $\boldsymbol{H}^H\boldsymbol{H}$. It may be enlightning to study the performance of systems with flat eigenvalue profile, i.e., $\lambda_1 = \lambda_2 = \cdots = \lambda_r = \text{tr}\, \boldsymbol{\Lambda}/r$, which commands $\mu = P_T/r + r \cdot \sigma_n^2/\text{tr}\, \boldsymbol{\Lambda}$ and yields capacity

$$C_r = r \cdot \log_2\left(1 + \frac{P_T}{\sigma_n^2} \cdot \frac{\text{tr}\, \boldsymbol{\Lambda}}{r^2}\right). \tag{6}$$

Now, let us fix the dimensions of the channel matrix $\boldsymbol{H}$ and vary its rank r. For a fair comparison, we keep the sum of all eigenvalues λ_j, i.e., $\text{tr}\, \boldsymbol{\Lambda}$ constant. We therefore merely change the channel rank, while we keep the total channel power constant, as $\text{tr}\, \boldsymbol{\Lambda} = \sum_{i,j} |H_{i,j}|^2$. Under these conditions, it is easy to show that

$$(C_1 \ge C_r) \leftrightarrow \left(\frac{P_T}{\sigma_n^2} \le \gamma_r\right) \tag{7}$$

with equality holding for $P_T/\sigma_n^2 = \gamma_r$. The transmit power threshold γ_r and the capacity $C_{\text{th},r}$ at the threshold can be lower bounded by

$$\gamma_r > \frac{r-1}{\text{tr}\, \boldsymbol{\Lambda}} \quad \text{and} \quad C_{\text{th},r} > \log_2 r. \tag{8}$$

If the benefit of a MIMO system is to supply a given channel capacity with lower transmit power than a single-input single-output (SISO) system, this result shows, that for desired capacities lower than $C_{\text{th},r}$, the rank-one channel will demand less transmit power than the full-rank one. Asymptotically, for $r \to \infty$, the rank-one channel will *always* get the job done with lower transmit power, while for finite r, it will operate favorably at low transmit power up to the threshold γ_r (see also [10]). To illustrate this point, Figs. 2 and 3 show two very similar MIMO

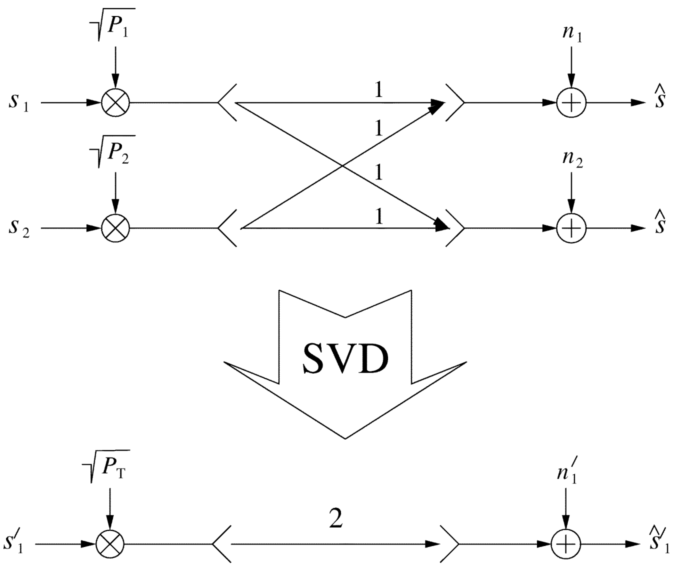


Fig. 3. A constant *rank deficient* MIMO channel (top) and its diagonalization (bottom).

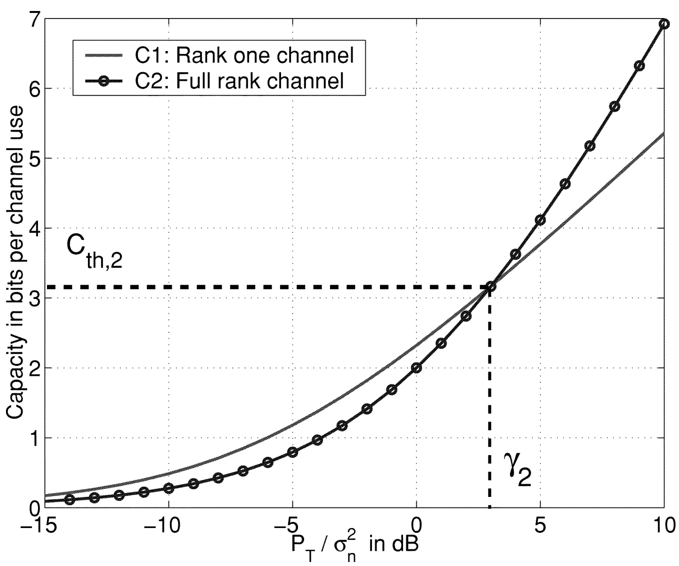


Fig. 4. Channel capacity of a full rank and a rank deficient 2×2 MIMO system with instantaneous channel knowledge at the transmitter. At low transmit power the rank deficient system yields higher capacity due to antenna gain, while the full rank system takes the lead at high transmit power, due to multistream transmission. In this example, the full rank channel behaves favorably only if the desired capacity is larger than $\log_2 9 \approx 3.17$ bits per channel use.

systems using $N = 2$ transmit and $M = 2$ receive antennas. The channel matrices are given as

$$\boldsymbol{H}_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad \boldsymbol{H}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

where $\boldsymbol{H}_2$ has full rankand $\boldsymbol{H}_1$ exhibits rank deficiency. Fig. 4 compares the capacities

$$C_1 = \log_2 \left(1 + 4 \cdot \frac{P_T}{\sigma_n^2}\right) \text{ and } C_2 = 2 \cdot \log_2 \left(1 + \frac{P_T}{\sigma_n^2}\right)$$

of both systems and shows that the rank deficient channel yields higher capacity for low transmit powers up to $\gamma_2 = 8/\text{tr}\, \boldsymbol{\Lambda} = 2$.

TABLE I
DEFINITION OF TYPES OF FADING CORRELATION

receiver side	transmitter side	type of fading
uncorrelated	uncorrelated	uncorrelated
uncorrelated	correlated	semi-correlated
correlated	uncorrelated	semi-correlated type 2
correlated	correlated	fully-correlated

Alternatively, desired capacities lower than $C_{\text{th},2} = \log_2 9 \approx 3.17$, are supplied demanding less transmit power.

In the sequel, we will deal with stochastic channels, where $\boldsymbol{H}$ is a zero-mean Gaussian random variable, modeling a Rayleigh frequency-flat fading wireless MIMO channel. However, we will allow correlations between entries of the channel matrix to be present, leading us to *correlated fading*. Such channels may exhibit a *stochastic* rank deficiency, meaning that the correlation matrices have zero, or very small eigenvalues. Similar to the deterministic case above, stochastic rank deficiency may lead to higher channel capacity as uncorrelated fading permits. Let us now define the correlation model.

VI. Spatial Fading Correlations

In general, we model fading correlations by writing the channel matrix as

$$\boldsymbol{H} = \frac{1}{\sqrt{\text{tr}\, \boldsymbol{R}_T}} \cdot \boldsymbol{R}_R^{1/2}\, \boldsymbol{G}\, \boldsymbol{R}_T^{1/2} \tag{9}$$

where $\boldsymbol{R}_R = \mathrm{E}\{\boldsymbol{H}\boldsymbol{H}^H\}$ is the $M \times M$ receive correlation matrix and $\boldsymbol{R}_T = \mathrm{E}\{\boldsymbol{H}^H\boldsymbol{H}\}$ is the $N \times N$ transmit correlation matrix, while $\boldsymbol{G} \in \mathcal{C}^{M \times N}$ is a random matrix with independent, *zero-mean unity variance* complex entries. In contrast to [6], where multiplicative normal distributions were used to describe key-hole channels, we will stick to the assumption, that $\boldsymbol{G}$ is drawn from a complex Gaussian distribution leading to correlated Rayleigh fading. Note that

$$\text{tr}\, \boldsymbol{R}_T = \text{tr}\, \boldsymbol{R}_R = \mathrm{E}\left\{\|\boldsymbol{H}\|_F^2\right\} = \sum_{m=1}^{M}\sum_{n=1}^{N} \mathrm{E}\left\{|H_{m,n}|^2\right\} \tag{10}$$

which is the channel's total power amplification. We can distinguish four fundamental cases of fading correlation: uncorrelated, semicorrelated, semicorrelated type 2, and fully correlated as defined in Table I. For brevity, we will restrict the discussion to the first two cases, furthermore, the semicorrelated model is valid for urban mobile radio channels, as has been shown by a recent measurement campaign taken in downtown Helsinki [12].

A. Uncorrelated Rayleigh Channel

Such a channel may arise if both transmitter and receiver live in a rich scattering environment (see Fig. 6). The result will be independent Rayleigh fading from each transmit to each receive antenna. We have

$$\boldsymbol{R}_T = M \cdot \boldsymbol{I}_N \qquad \boldsymbol{R}_R = N \cdot \boldsymbol{I}_M \tag{11}$$

producing $\boldsymbol{H} = \boldsymbol{G}$ and $\text{tr}\, \boldsymbol{R}_R = \text{tr}\, \boldsymbol{R}_T = M \cdot N$. Note, that $\boldsymbol{I}_n$ is the $n \times n$ identity matrix.

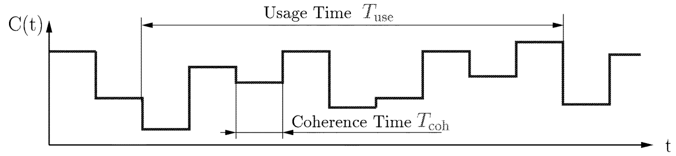


Fig. 5. Example of the temporal evolution of instantaneous channel capacity of a block-fading channel with constant transmit power.

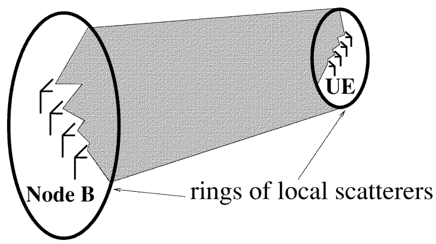


Fig. 6. Geometrical interpretation of an uncorrelated MIMO channel. Both receiver and transmitter are located in a rich scattering environment, with no line of sight connection.

B. Semicorrelated Rayleigh Channel

Imagine the *transmitter* is now removed from its rich scattering environment. From the transmitter's point of view the spatial structure of the channel now is governed by remote scattering objects and will most likely result in a highly spatially correlated scenario, for usually there will only be a small number K of dominant remote scattering objects (see Fig. 8). Hence, we can write

$$\boldsymbol{H} = \sqrt{\frac{N}{\text{tr } \boldsymbol{A}\boldsymbol{A}^H}} \cdot \boldsymbol{Z}\boldsymbol{A}^T$$

where $\boldsymbol{A} \in \mathcal{C}^{N \times K}$ is an array steering matrix containing K array response vectors of the transmitting antenna array corresponding to K directions of departure (DoD) and $\boldsymbol{Z} \in \mathcal{N}_{\mathcal{C}}^{M \times K}(0,1)$ has zero-mean independent and identically distributed (i.i.d.) Gaussian random entries. Angle spread is easily modeled by a high enough number of discrete DoDs. Applied to the statistical model defined in (9), we have

$$\boldsymbol{R}_T = \frac{M \cdot N}{\text{tr } \boldsymbol{A}^* \boldsymbol{A}^T} \cdot \boldsymbol{A}^* \boldsymbol{A}^T , \quad \boldsymbol{R}_R = N \cdot \boldsymbol{I}_M. \tag{12}$$

Note, that the total power amplification of this channel is normalized to tr $\boldsymbol{R}_R =$ tr $\boldsymbol{R}_T = M \cdot N$, which is the same as in the uncorrelated case.

VII. Sample-Mean Outage Capacity

The channel capacity represents an ultimate information theoretic upper bound on system performance. As the investigated channels are usually time-varying, they are represented by random processes. Maximization of mutual information with an average transmit power constraint was presented for time selective channels in [7]. Here, we will use a different approach, by using a constant transmit power constraint, independent of channel state. For simplicity, we look at a block-fading channel, which properties remain constant during the *coherence time* T_{coh} and afterwards change to a new, independent realization. The capacity of such a channel depends on the ratio between usage time T_{use} and coherence time T_{coh}, defined in Fig. 5.

- For $T_{\text{use}} \rightarrow \infty$, the appropriate characterization of capacity clearly is given by the temporal average $\bar{C}$, or assuming ergodicity by the expected value, the so called *ergodic capacity*: $C_{\text{erg}} = \text{E}\{C\}$.
- For $T_{\text{use}} < T_{\text{coh}}$, there is nothing to average over. An appropriate characterization of capacity in this case is the well known *outage capacity* $C_{\text{out}}(p)$ describing the capacity that can be guaranteed with probability equal to $1-p$, i.e., Prob $\{C < C_{\text{out}}(p)\} = p$
- For $T_{\text{use}} = m \cdot T_{\text{coh}}$, there are exactly m independent channel realizations during the usage time. By defining a new random variable $C' = 1/m \sum_{k=1}^{m} C_k$, where C_k are the instantaneous capacities corresponding to different channel realizations, the appropriate characterization of capacity is the proposed *sample-mean outage capacity* $C_{\text{soc}}(p,m)$: Prob $\{C' < C_{\text{soc}}(p,m)\} = p$.

The sample-mean outage capacity contains both the ergodic and the outage capacity as special cases, since

$$C_{\text{soc}}(p,1) = C_{\text{out}} \quad \text{and} \quad C_{\text{soc}}(p,\infty) = C_{\text{erg}}. \tag{13}$$

A system can approach sample-mean outage capacity in at least two different ways.

1) *Adaptive coding:* Change code rate and code according to the current channel quality, i.e., transmit at higher rate when the channel is good. This requires the transmitter to acquire the instantaneous capacity of the channel, which may involve establishment of a feedback link from the receiver. This method makes possible the achievement of the average capacity over different channel realizations.
2) *Interleaving:* Spread the codewords over m fading blocks and use a fixed rate code. This is simpler than adaptive coding and does not require knowledge of instantaneous channel capacity. However, this method makes possible the achievement of the capacity of the average channel only. It is in general a suboptimum, however, more practical approach than adaptive coding.

VIII. Instantaneous Channel Capacity

Assuming the system model defined in Section III and complete *receiver* side channel knowledge, the instantaneous mutual information $\mathcal{I}_{\boldsymbol{s}\rightarrow\boldsymbol{y}}$ from transmitted Gaussian signal $\boldsymbol{s}$ to received signal $\boldsymbol{y}$ is given as [16]

$$\mathcal{I}_{\boldsymbol{s}\rightarrow\boldsymbol{y}}\left(\boldsymbol{H}^H\boldsymbol{H}, \boldsymbol{T}, \boldsymbol{P}\right) = \log_2 \det\left(\boldsymbol{I}_L + \frac{1}{\sigma_n^2}\boldsymbol{T}^H\boldsymbol{H}^H\boldsymbol{H}\boldsymbol{T}\boldsymbol{P}\right). \tag{14}$$

The instantaneous channel capacity is the maximum mutual information

$$C\left(\boldsymbol{H}^H\boldsymbol{H}\right) = \max_{\boldsymbol{T},\boldsymbol{P}} \mathcal{I}_{\boldsymbol{s}\rightarrow\boldsymbol{y}}\left(\boldsymbol{H}^H\boldsymbol{H}, \boldsymbol{T}, \boldsymbol{P}\right), \text{s.t. tr } \boldsymbol{P} = P_T \tag{15}$$

where the maximization is done with respect to the spatial processing via $\boldsymbol{T}$ and power distribution via $\boldsymbol{P}$, with the constraint of having a given total transmit power P_T.

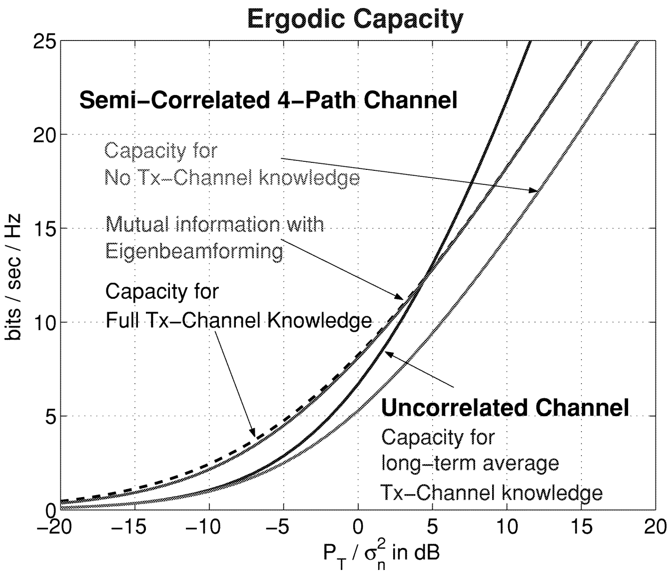


Fig. 7. Comparison of ergodic capacity and mutual information for semicorrelated and uncorrelated channels with and without average channel knowledge. Note that in the uncorrelated case, having no channel knowledge is equivalent to having average channel knowledge.

IX. Transmitter-Sided Channel Knowledge

To what extend the maximization of mutual information can be carried out, now depends on the amount of knowledge the *transmitter* has about the channel. We will distinguish three cases.

A. Complete Channel Knowledge

Assuming that the transmitter exactly knows the channel matrix $\boldsymbol{H}$ at each transmit time instant, it is well known [16] that by following the procedure:

- EVD: $\boldsymbol{H}^H\boldsymbol{H} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^H$;
- set $\boldsymbol{T} = \boldsymbol{V}$ and choose $\boldsymbol{P}$ by waterfilling based on $\boldsymbol{\Lambda}$;

the instantaneous mutual information (14) is maximized.

B. No Channel Knowledge

The other extreme is having the transmitter be completely ignorant about the channel. We follow this procedure:

- set $\boldsymbol{T} = \boldsymbol{I}_N$ and $\boldsymbol{P} = (P_T/N) \cdot \boldsymbol{I}_N$;
- hope for the best.

In this scenario, each antenna transmits an independent data stream with the power being shared equally among the streams. While for uncorrelated channels most of the capacity achievable with complete knowledge can be retained, it turns out to be disastrous in the case of semi- or fully correlated channels.

C. Average Channel Knowledge

While complete channel knowledge may be too demanding a request in practice, assuming no transmitter channel knowledge may well be over conservative. In most cases, the transmitter should be able to acquire knowledge about the channel *on average*. Assuming knowledge of the transmit correlation matrix $\boldsymbol{R}_T$ and following the procedure:

- EVD: $\mathrm{E}\left\{\boldsymbol{H}^H\boldsymbol{H}\right\} = \boldsymbol{R}_T = \boldsymbol{V}'\boldsymbol{\Lambda}'\boldsymbol{V}'^{\,H}$;
- Set $\boldsymbol{T} = \boldsymbol{V}'$ and choose $\boldsymbol{P}$ by Water-filling based on $\boldsymbol{\Lambda}'$

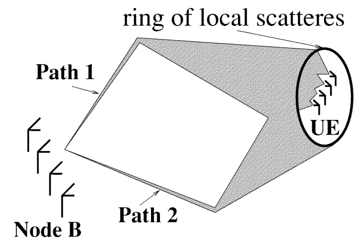


Fig. 8. Geometrical interpretation of a semicorrelated MIMO channel. The channel is spatially correlated from the transmitter's point of view, as the receiver can be reached through just two narrow spatial directions, while from the receiver's point of view the channel is uncorrelated due to its rich scattering environment.

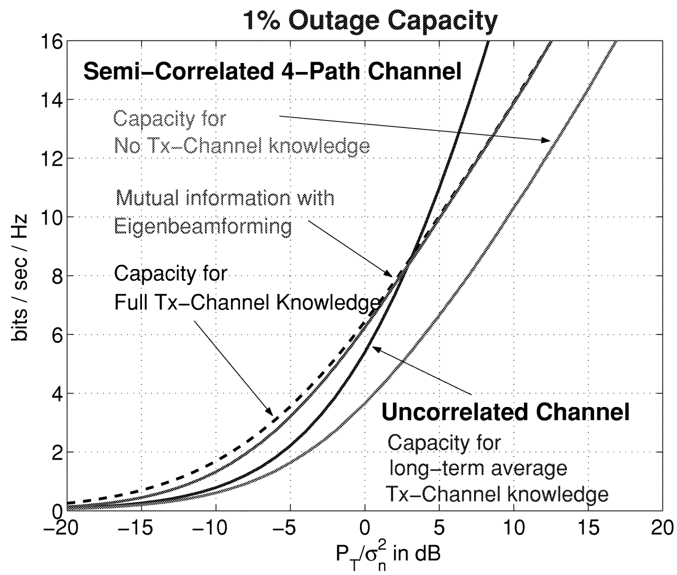


Fig. 9. Comparison of outage capacity and mutual information for semicorrelated and uncorrelated channels with and without average channel knowledge.

will maximize $\mathcal{I}_{\boldsymbol{s}\to\boldsymbol{y}}(\mathrm{E}\{\boldsymbol{H}^H\boldsymbol{H}\}, \boldsymbol{T}, \boldsymbol{P})$, i.e., the mutual information of the average channel. While complete channel knowledge allows for maximizing the average mutual information, *average* channel knowledge allows for maximizing the mutual information of the *average channel*. As we shall see in the next section, this yields close to optimum performance in MIMO systems in semicorrelated fading environments, where the performance may be even better than in the uncorrelated case. The procedure above is called *MIMO downlink eigenbeamforming* [9].

X. Capacity of Semicorrelated Rayleigh Channels

To evaluate the capacity of semicorrelated channels with and without average channel knowledge, we simulate a $M = N = 8$ antenna system, where the antennas form an omni-directional uniform linear array. We use a four-path semicorrelated channel and an uncorrelated channel for comparison. The four paths have zero angle spread and random directions of departure. Averaging over both the Rayleigh fading path coefficients and the random directions of departure is made, where the latter are assume to be uniformly and independentily distributed in the azimuthal range of $-90° \ldots 90°$. The Figs. 7 and 9 show the results. When using long-term transmit channel knowledge, the mutual information is computed by evaluating (14) with the matrices $\boldsymbol{T}$ and $\boldsymbol{P}$ chosen by the method outlined in Section IX-C. There are four major points we like to stress.

1) If there is no transmit channel knowledge, the spatial correlations reduce capacity compared with the uncorrelated case. This is true both with time diversity (ergodic capacity) and without time diversity (outage capacity) available.
2) If *average* transmit channel knowledge is used, the picture changes: for low transmit powers up to a cross over point, the semicorrelated channel indeed offers higher capacity than the uncorrelated one. This is true for both ergodic and outage capacities.
3) For the semicorrelated channel, the difference between average and complete channel knowledge is marginal and disappears for high transmit powers. As this is true even for the outage capacity, average channel knowledge is in practical terms sufficient even if no additional time or frequency diversity are available.
4) Due to full channel rank, the uncorrelated channel, used at high transmit powers, gets better and better compared with the semicorrelated case—or so it would seem. However, note, that any real communication system will have to use finite constellation-size modulation schemes, which will limit the achievable capacity. Taking realistic modulation schemes into account will again change the picture, as we shall see shortly.

XI. MIMO Antenna Gain

The capacity advantage of correlated fading MIMO channels at low transmit power can be explained by the notion of *MIMO antenna gain*. An attempt to define such an antenna gain has been done in [1]. However, that definition was made with single-stream transmission in mind and does not cover the influence of long-term average transmitter channel knowledge. In the following, we therefore propose a more general definition of MIMO antenna gain, which takes both multistream transmission and different transmitter sided channel knowledge into account.

A. Instantaneous Transmitter Channel Knowledge

Assuming there is no noise at the receiver, the *received* power P_R for a given channel $\boldsymbol{H}$ equals

$$P_R = \mathrm{E}\left\{\|\boldsymbol{y}\|_2^2 \mid \boldsymbol{H}\right\} = \mathrm{tr}\, \boldsymbol{T}^H \boldsymbol{H}^H \boldsymbol{H} \boldsymbol{T} \boldsymbol{P} = \mathrm{tr}\, \boldsymbol{\Lambda} \boldsymbol{P} \tag{16}$$

as $\boldsymbol{H}^H\boldsymbol{H} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^H$ and $\boldsymbol{T} = \boldsymbol{V}$. If instead, just a single pair of receive and transmit antennas was used to form a SISO system, connecting the ith receive with the jth transmit antenna, the receive power would read as

$$(P_R')_{i,j} = P_T \cdot |H_{i,j}|^2 . \tag{17}$$

Taking the average receive power over all possible pairs of receive and transmit antennas, we define the receive power of a reference SISO system as

$$P_R' = \frac{P_T}{NM} \cdot \sum_{i=1}^{M} \sum_{j=1}^{N} |H_{i,j}|^2 = \frac{(\mathrm{tr}\, \boldsymbol{P}) \cdot (\mathrm{tr}\, \boldsymbol{\Lambda})}{NM}. \tag{18}$$

The instantaneous antenna gain is then simply the ratio P_R/P_R', i.e.,

$$A_{\mathrm{inst}} = \frac{P_R}{P_R'} = NM \cdot \frac{\mathrm{tr}\, (\boldsymbol{\Lambda}\boldsymbol{P})}{(\mathrm{tr}\, \boldsymbol{P}) \cdot (\mathrm{tr}\, \boldsymbol{\Lambda})} . \tag{19}$$

It clearly depends both on the eigenvalue profile of the instantaneous channel Gramian $\boldsymbol{H}^H\boldsymbol{H}$ and the distribution of transmit power over its eigenmodes. As the eigenvalue profile depends both on receive and transmit properties of the MIMO channel, there is no decoupling of antenna gain in a receive and a transmit part. The maximum antenna gain

$$\widehat{A}_{\mathrm{inst}} = \max_{\boldsymbol{P}}\ A_{\mathrm{inst}} = NM \cdot \frac{\lambda_{\max}}{\mathrm{tr}\, \boldsymbol{\Lambda}} \tag{20}$$

is achieved, if the strongest eigenmode only is powered up, i.e., $P_1 = P_T = \mathrm{tr}\ \boldsymbol{P}$, which leads to single-stream transmission. Note, that waterfilling power distribution has this very effect at low transmit power. To make this more clear, assume that

$$\frac{P_T}{\sigma_n^2} \leq \frac{1}{\lambda_2} - \frac{1}{\lambda_1}$$

where $\lambda_1 \geq \lambda_2 > 0$ are the two largest eigenvalues of the channel Gramian $\boldsymbol{H}^H\boldsymbol{H}$. In this case, the capacity achieving waterfilling power distribution will just power up the strongest eigenmode and the instantanteous channel capacity from (15) becomes $C = \log(1 + (P_T/\sigma_n^2)\lambda_1)$. Because of the normalization $\mathrm{tr}\ \boldsymbol{\Lambda} = M \cdot N$, we can write

$$C = \log\left(1 + \frac{P_T}{\sigma_n^2} \cdot \frac{M \cdot N}{\mathrm{tr}\, \boldsymbol{\Lambda}} \cdot \lambda_1\right) = \log\left(1 + \widehat{A}_{\mathrm{inst}} \cdot \frac{P_T}{\sigma_n^2}\right)$$

which shows a simple relationship of maximum instanteous antenna gain $\widehat{A}_{\mathrm{inst}}$ and channel capacity for low transmit powers.

B. Long-Term Average Transmitter Channel Knowledge

In the case, where only long-term average channel information is available to the transmitter, we define the MIMO antenna gain, based on *average* receive power

$$\begin{aligned}\overline{P_R} =& \mathrm{E}\left\{\mathrm{E}\left\{\|\boldsymbol{y}\|_2^2 \mid \boldsymbol{H}\right\}\right\} = \mathrm{E}\left\{\|\boldsymbol{y}\|_2^2\right\} = \mathrm{tr}\, \boldsymbol{T}^H \boldsymbol{R}_T \boldsymbol{T} \boldsymbol{P} \\ =& \mathrm{tr}\, \boldsymbol{\Lambda}' \boldsymbol{P}\end{aligned} \tag{21}$$

as $\boldsymbol{R}_T = \boldsymbol{V}'\boldsymbol{\Lambda}'\boldsymbol{V}'^H$ and $\boldsymbol{T} = \boldsymbol{V}'$. A SISO system operated over two antennas of the same MIMO channel would produce on the average over different antenna pairs the average receive power

$$\overline{P_R'} = \frac{P_T}{NM} \cdot \sum_{i=1}^{M} \sum_{j=1}^{N} \mathrm{E}\left\{|H_{i,j}|^2\right\} = \frac{(\mathrm{tr}\, \boldsymbol{P}) \cdot (\mathrm{tr}\, \boldsymbol{\Lambda}')}{NM}. \tag{22}$$

Analogously to the instantaneous case, we define the long-term average antenna gain

$$A_{\mathrm{LT}} = \frac{\overline{P_R}}{\overline{P_R'}} = NM \cdot \frac{\mathrm{tr}\ (\boldsymbol{\Lambda}'\boldsymbol{P})}{(\mathrm{tr}\, \boldsymbol{P}) \cdot (\mathrm{tr}\, \boldsymbol{\Lambda}')} = A_{\mathrm{LT}}^{\mathrm{Rx}} \cdot A_{\mathrm{LT}}^{\mathrm{Tx}} \tag{23}$$

which in contrast to the instantaneous case can be decoupled easily into a receive and a transmit part

$$A_{\mathrm{LT}}^{\mathrm{Rx}} = M, \quad \text{and} \quad A_{\mathrm{LT}}^{\mathrm{Tx}} = N \cdot \frac{\mathrm{tr}\ (\boldsymbol{\Lambda}'\boldsymbol{P})}{(\mathrm{tr}\, \boldsymbol{P}) \cdot (\mathrm{tr}\ \ \boldsymbol{\Lambda}')} \tag{24}$$

as the eigenvalue profile of the transmit correlation matrix $\boldsymbol{R}_T$ is solely a function of transmitter-sided stochastic properties of the MIMO channel. The largest antenna gain is again achieved for single-stream transmission

$$\widehat{A}_{\mathrm{LT}} = NM \cdot \frac{\lambda'_{\max}}{\mathrm{tr}\,\mathbf{\Lambda}'}. \tag{25}$$

Note that $M \leq A_{\mathrm{LT}} \leq NM$, where the lower bound is taken on for the uncorrelated case, where $\boldsymbol{R}_T$ is a scaled unity matrix. The maximum antenna gain of $A_{\mathrm{LT}} = NM$ is obtained if $\boldsymbol{R}_T$ has a rank of one, which may happen in practice, if the mobile station can be reached through just one single-dominant scattering object, with small enough angle spread. We will look at such a scenario in more detail later. For low transmit powers the maximum long-term antenna gain $\widehat{A}_{\mathrm{LT}}$ has a simple relationship with the capacity of the average channel, as defined in Section IX-C. For

$$\frac{P_T}{\sigma_n^2} \leq \frac{1}{\lambda'_2} - \frac{1}{\lambda'_1}$$

with $\lambda'_1 \geq \lambda'_2 > 0$ as the two largest eigenvalues of the transmit correlation matrix, the capacity of the average channel becomes $C' = \log(1 + (P_T/\sigma_n^2)\lambda'_1)$. Because of the normalization tr $\mathbf{\Lambda}' = M \cdot N$, we can write

$$C' = \log\left(1 + \frac{P_T}{\sigma_n^2} \cdot \frac{M \cdot N}{\mathrm{tr}\,\mathbf{\Lambda}'} \cdot \lambda'_1\right) = \log\left(1 + \widehat{A}_{\mathrm{LT}} \cdot \frac{P_T}{\sigma_n^2}\right).$$

C. No Transmitter Channel Knowledge

If the transmitter has no channel state information (CSI) available, the uniform power distribution $\boldsymbol{P} = (P_T/N)\boldsymbol{I}_N$ leads to an antenna gain of

$$A_{\mathrm{NoCSI}} = M \tag{26}$$

which is solely due to the receive antenna gain. Activation of transmitter sided antenna gain needs at least long-term information about the channel.

D. Average Instantaneous Versus Long-Term Average Antenna Gain

As the channel $\boldsymbol{H}$ is a random variable, this is also true for the instantaneous antenna gain. It is interesting to compare its expected value $\mathrm{E}\{\widehat{A}_{\mathrm{inst}}\}$ to the long-term average antenna gain $\widehat{A}_{\mathrm{LT}}$. From a result of [8] the average maximum eigenvalue of a matrix $\boldsymbol{H}^H\boldsymbol{H}$ can be upper bounded, if $\boldsymbol{H} \in \mathcal{C}^{M\times N}$ is a complex Gaussian matrix with zero-mean i.i.d. components

$$\mathrm{E}\left\{\frac{\lambda_{\max}}{\mathrm{tr}\,\mathbf{\Lambda}}\right\} < \left(\frac{1}{\sqrt{M}} + \frac{1}{\sqrt{N}}\right)^2. \tag{27}$$

From this follows:

$$\mathrm{E}\left\{\widehat{A}_{\mathrm{inst}}\right\} < \left(\sqrt{M} + \sqrt{N}\right)^2 \tag{28}$$

that the average maximum instantaneous antenna gain of an uncorrelated Gaussian channel grows much slower than the product of the number of receive and transmit antennas. On the other hand, the long-term average antenna gain of a correlated fading MIMO channel takes on the value $\widehat{A}_{\mathrm{LT}} = M \cdot N$ for rank $\boldsymbol{R}_T = 1$. This explains the possibility of having performance gain in the low transmit power region, when dealing with correlated fading MIMO channels. This effect gets stronger with larger antenna number, as is illustrated in Fig. 10. Please note, that for low transmit power, single-stream transmission is capacity achieving, which makes antenna gain an important issue in the low power region. On the other hand, in the high transmit power domain, antenna gain is no figure of merrit anymore, as true multistream transmission is needed to achieve capacity.

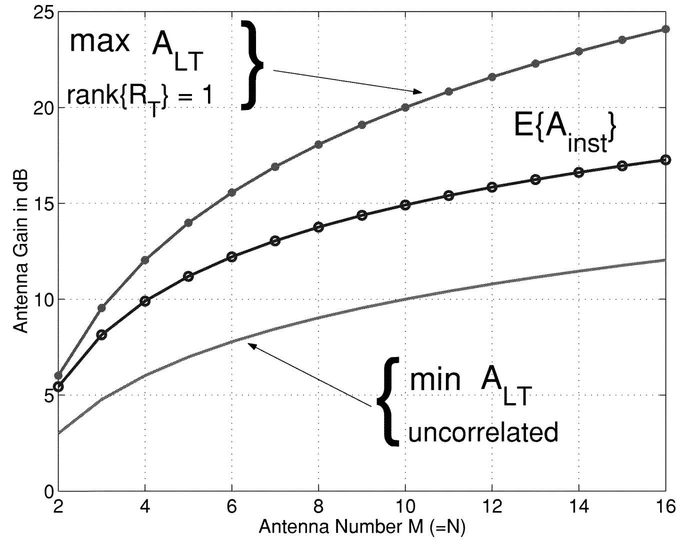


Fig. 10. Comparison of average instantaneous MIMO antenna gain of an uncorrelated channel ($\mathrm{E}\{A_{\mathrm{inst}}\}$) and long-term average MIMO antenna gain (A_{LT}) for different number of transmit and receive antennas (here equal number of receive and transmit antennas was assumed). Clearly, long-term average antenna gain in correlated fading can be way larger than the average instantaneous antenna gain in the uncorrelated case. This happens for highly correlated fading, where the transmit correlation matrix has significant (numerical) rank deficiency.

XII. Cutoff Rate

While capacity is a theoretical limit for infinite block length codes and zero-error probability, the cutoff rate gives a bound for finite block length and error probability. Furthermore, it is computationally less demanding to compute cutoff rates than capacities for linear digital modulation schemes in MIMO systems. The cutoff rate is useful because of the cutoff-rate theorem [13], which states that there exist $(n,k)_q$ block codes, with code word error probability P_w after maximum- likelihood decoding being upper bounded by $P_w < 2^{-n\cdot(R_0 - R_b)}$, provided the binary code rate $R_b := k/n \cdot \log_2 q$ is less than the cutoff rate

$$R_0 = -\log_2 \int_{\mathcal{C}^M} \left(\sum_{\boldsymbol{s}\in\mathcal{M}} \frac{1}{q}\sqrt{p(\boldsymbol{y}|\boldsymbol{s})}\right)^2 d\boldsymbol{y} \tag{29}$$

where $\mathcal{M}$, with $|\mathcal{M}| = q$ is the set of code symbols (input alphabet) and $p(\boldsymbol{y}|\boldsymbol{s})$ is the probability density function of the received signal $\boldsymbol{y}$ given the transmitted code symbol $\boldsymbol{s}$. To apply this to our MIMO system, we consider the data vector $\boldsymbol{s}$ as a q-ary code symbol, where each component s_k, with $1 \leq k \leq L$ can take on q_k values from a discrete modulation alphabet $\mathcal{M}_k$, with $|\mathcal{M}_k| = q_k$. The input alphabet $\mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2 \times \cdots \times \mathcal{M}_L$, is the Cartesian product of the individual alphabet sets, with $|\mathcal{M}| = q = q_1 \cdot q_2 \cdots q_L$. By labeling the elements of $\mathcal{M} = \{\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_q\}$ the instantaneous cutoff rate can be written as [11]

$$R_0 = \log_2(q) - \log_2 \cdot \left(1 + \frac{2}{q}\sum_{p=1}^{q-1}\sum_{t=p+1}^{q} \exp\left(-\frac{1}{4}\left\|\boldsymbol{b}_p - \boldsymbol{b}_t\right\|_2^2\right)\right) \quad (30)$$

with $\boldsymbol{b}_p = (1/\sigma_n)\boldsymbol{HTP}^{1/2}\boldsymbol{s}_p$. Ergodic, outage, and sample-mean outage cutoff rates can be computed accordingly to the discussion in Section VII.

XIII. Cutoff-Rate Performance

Some insight in MIMO system performance in correlated fading can be gained by evaluating the cutoff rate for realistic modulation schemes and antenna numbers. We simulate a system consisting of $N = 2$ transmit antennas, separated half a wavelength apart and $M = 2$ receive antennas, that is operated either over an uncorrelated channel, or over a one-path semicorrelated channel. The latter could result in practice from a scenario, where the receiver can be reached by remote scattering from just *one* single tall object—like a church tower or a tall lamp mast—located in adequate distance from the transmitter. Note, that the transmit covariance matrix will have numerical rank deficiency, if the angle spread is small compared with the standard beamwidth of the transmit array—in this case 60° in the bore-side direction. In the semicorrelated case, we will distinguish between *no* and *average* transmitter channel knowledge. The uncoded (raw) bit rate shall be fixed to $r_{\mathrm{raw}} = 4$ bits per use in all cases, as to implement a given service. For the uncorrelated case, we use 4-quadrature amplitude modulation (QAM) and transmit two independent data streams—one over each antenna. The same holds for the semicorrelated case with *no* channel knowledge, as the transmitter is not aware of the channel conditions. If, on the other hand, average channel information is available to the transmitter, the rank deficiency can be turned into antenna gain. Hence, just one single data stream will be transmitted over the dominant eigenbeam. To achieve the same raw bit rate, the modulation scheme is changed to 16QAM. Fig. 13 summarizes the test bed, while Figs. 11, 12, and 14 show the following results.

1) The uncorrelated channel performs best, when no time diversity is available, as can be seen from the outage cutoff rate—at least in the interesting range of code rates ($R < 0.9$).
2) The more time diversity is available—i.e., the more independent channel realizations are available during airtime—the more attractive the semicorrelated channel becomes.
3) If there is no transmit channel knowledge, however, the semicorrelated channel performs worst, no matter how much time diversity is available.
4) Arming the transmitter with average channel knowledge, the semicorrelated channel turns out to yield the best performance—in fact, beating the performance of the uncorrelated channel—provided there is enough time diversity available.

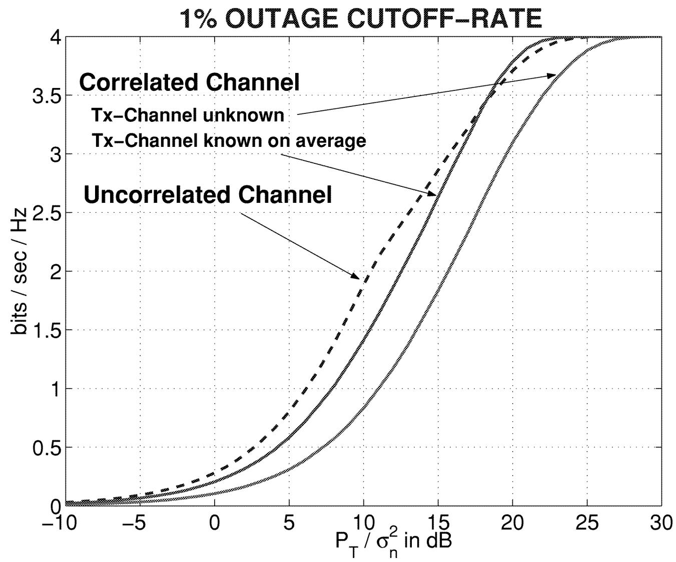


Fig. 11. Outage cutoff rates for the systems in Fig. 13. As outage cutoff rate describes a situation with no time or frequency selectivity, the only source of diversity is space. Hence, the uncorrelated channel behaves superior to the semicorrelated case. Note however, that in this particular scenario, the correlated channel gets the lead over the uncorrelated one in the *high* rather than low transmit power region, though not of much practical concern as the code rates must be larger than about 0.9 ($R_0 > 3.5$ bits per use), which rarely is the case for realistic channel coding in wireless communication systems.

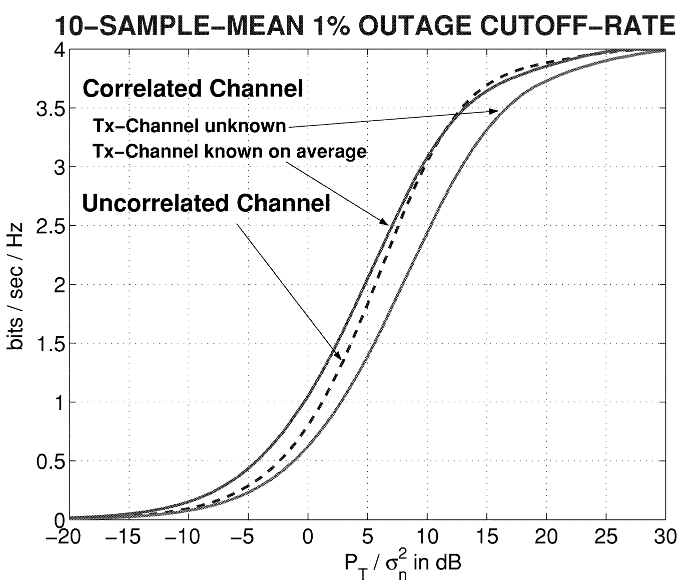


Fig. 12. Sample mean outage cutoff rates for the systems in Fig. 13, with a sample size of ten. The sample mean outage cutoff rate describes a situation with a certain amount of time or frequency selectivity available—in this case interleaving over ten coherence times of the channel is assumed. The situation now gets more in favor of the semicorrelated channel with long-term average information at the transmitter. Here, clearly, it is the low transmit power region where the fading correlations get beneficial.

5) Looking at the ten sample-mean outage cutoff rate, we see, that the amount of ten independent channel realizations during airtime, suffices for the semicorrelated channel to realize all code rates less than about 0.8 with lower transmit power than required for the uncorrelated channel.
6) If the number of independent channel realizations during airtime is further extended, the ergodic cutoff rate shows,

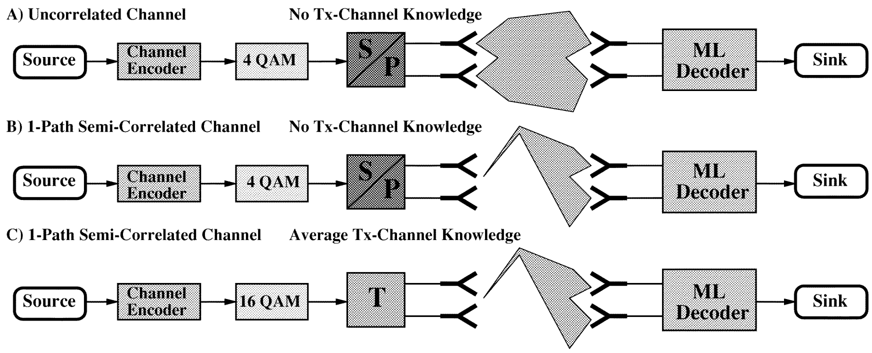


Fig. 13. Setup for comparing cutoff rates for different fading correlations and transmitter channel knowledge. The raw data rate is fixed to 4 bits per use.

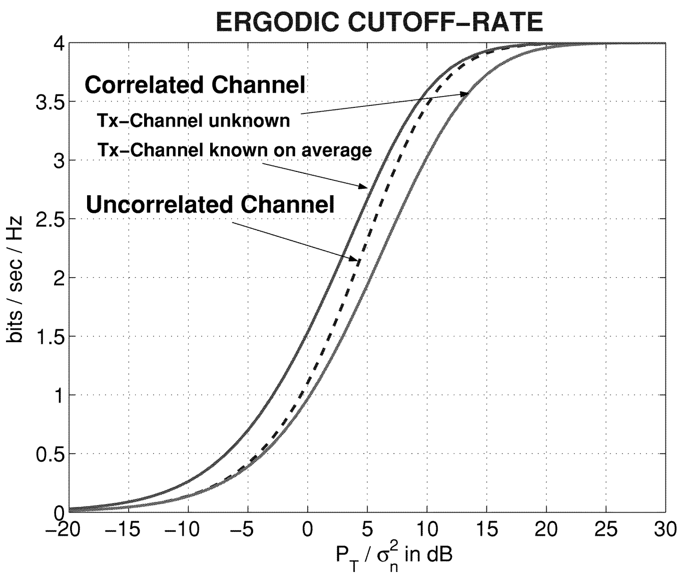


Fig. 14. Ergodic cutoff rates for the systems in Fig. 13. This is equivalent to sample mean outage cutoff rate with an infinite sample size, hence, unlimited amounts of time or frequency selectivity, e.g., provided by perfect interleaving are available. Here, the semicorrelated channel with long-term average channel information at the transmitter gives the favorable performance for virtually *all* transmit powers and code rates.

that the advantage of the semicorrelated channel is still improving and extended for virtually all code rates.

XIV. Conclusion

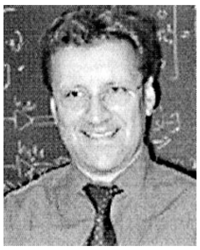
The capacity of MIMO systems depends on the statistical properties of the channel and the amount of knowledge about those properties. While for no transmitter channel knowledge correlated fading is a curse—especially if no other form of diversity, like frequency or time, is available—having the transmitter acquire the channel properties on average, can actually lead to capacity improvement over uncorrelated fading channels. This effect can be understood by the notion of antenna gain, which can be far larger for correlated fading than in the uncorrelated case. This leads to performance increase of correlated MIMO channels for low transmit powers. In the domain of high transmit powers, fading correlations decrease performance, since capacity gains due to multistream transmission and higher space diversity get the lead over antenna gain. It depends on the operating point a MIMO system is set up to work in, if fading correlations are helpful or not. The concept of sample-mean outage was defined, which allows to compute the influence of time selectivity on information theoretic measures, like capacity or cutoff rate. Cutoff-rate analysis showed, that, for linear modulation schemes, semicorrelated fading channels have potential to offer superior performance in an interesting transmit power range, provided a modest amount of time or frequency diversity is available in addition to pure space diversity.

References

[1] J. B. Andersen, "Antenna arrays in mobile communications: Gain, diversity and channel capacity," *IEEE Antennas Propagat. Mag.*, vol. 42, pp. 12–16, Apr. 2000.
[2] H. Bolcskei, D. Gesbert, and A. Paulraj, "On the capacity of OFDM-based multi-antenna systems," in *Proc. IEEE ICASSP*, June 2000, pp. 2569–2572.
[3] C. Chuah, J. M. Kahn, and D. Tse, "Capacity of multi-antenna array systems in indoor wireless environment," in *Proc. GLOBECOM*, Nov. 1998, pp. 1894–1899.
[4] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment using multiple antennas," *Wireless Pers. Commun.*, vol. 6, pp. 311–335, 1998.
[5] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.
[6] D. Gesbert, H. Bolcskei, D. Gore, and A. Paulraj, "MIMO wireless channels: Capacity and performance prediction," in *Proc. IEEE GLOBECOM Conf.*, Nov. 2000, pp. 1083–1088.
[7] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1986–1992, Nov. 1997.
[8] J. Haagerup and S. Thorbjornsson, "Random matrices with complex gaussian entries," Centre for Mathematical Physics and Stochastics, Aarhus Univ., Denmark, Res. Rep. 14, 1998.
[9] M. T. Ivrlač, T. P. Kurpjuhn, C. Brunner, and W. Utschick, "Efficient use of fading correlations in MIMO systems," in *Proc. 54th IEEE Vehicular Technology Conf. (VTC '01)*, Oct. 2001, pp. 2763–2767.
[10] S. A. Jafar, S. Vishwanath, and A. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback," in *Proc. IEEE Int. Conf. Communications, ICC 2001*, June 2001, pp. 2266–2270.
[11] T. Kurpjuhn and M. Ivrlac, "Information Theoretic Measures for MIMO-Channels," Lehrstuhl für Netzwertheorie und Signalverarbeitung, Munich, Germany, Technische Univ. München, TUM-LNS-TR-01-01, 2001.
[12] J. Laurila, K. Kalliola, M. Toeltsch, K. Hugel, P. Vainikainen, and E. Bonek, "Wideband 3-D characterization of mobile radio channels in urban environment," *IEEE Trans. Antennas Propagat.*, vol. 50, pp. 233–243, Feb. 2002.
[13] J. L. Massey, "Coding and modulation in digital communications," presented at the Int. Zürich Seminar, Sindelfingen, Germany, Mar. 1974.
[14] J. Salz and J. Winters, "Effect of fading correlation on adaptive arrays in digital mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 1049–1057, Nov. 1994.
[15] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multi-element antenna systems," *IEEE Trans. Commun.*, vol. 48, pp. 502–513, Mar. 2000.
[16] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecomm.*, vol. 10, no. 6, pp. 585–596, Nov. 1999.

Michel T. Ivrlač received the first diploma degree in electrical engineering from the Munich University of Applied Sciences (FH München), Munich, Germany, in 1994. He earned a second diploma degree in electrical engineering from the Munich University of Technology (TU München), Munich, Germany, in 1998. He is currently with the Institute for Circuit Theory and Signal Processing, Munich University of Technology, working as a Research and Teaching Assistant and toward the Ph.D. degree.

His research interests include signal processing and coding for wireless communications, quality of service control, and cross OSI-layer optimization strategies for heterogenious communication networks.



Wolfgang Utschick completed several industrial education programs before receiving the diploma degree in electrical enginerring and the Dr.-Ing. degree in 1993 and 1998, respectively, both from the Munich University of Technology (TU München), Munich, Germany.

During the Summer 2000, he was a Visting Researcher with the Eidgenössische Technische Hochschule, Zürich, Switzerland. He is currently a Full Professor for signal processing in wireless communications with the Institute for Circuit Theory and Signal Processing, TU München. His main research interests are signal processing in wireless communications, estimation theory, and nonlinear optimization.

Josef A. Nossek (F'93) received the Dipl.-Ing. and Dr. Techn. degrees in electrical engineering from the University of Technology, Vienna, Austria, in 1974 and 1980, respectively.

In 1974, he joined Siemens AG, Munich, Germany, working on the design of discrete-time filters for communications. In 1980, he became Head of the department responsible for digital microwave radio design activities. Since 1989, he has been a Full Professor for network theory and signal processing at Munich University of Technology (TU München), Munich, Germany. In 1984, he was a Guest Professor at the University of Cape Town, South Africa. He was a Guest Professor at the University of California, Berkeley, in 1992 and 1998, and a Guest Professor at the University of Technology, Vienna, in 1995.