ON THE IMPACT OF CORRELATED FADING FOR MIMO-SYSTEMS

Michel T. Ivrlač and Josef A. Nossek

Institute for Circuit Theory and Signal Processing
Munich University of Technology
Arcisstr. 16, D-80290 Munich, Germany
{ivrlac, nossek}@ei.tum.de

ABSTRACT

We investigate the effects of fading correlations on wireless communication systems employing multiple antennas at both the receiving and the transmitting side of the link, so called multiple-input multiple-output (MIMO) systems. It turns out that these effects depend on factors such as the amount of transmitter sided channel knowledge and the availability of other sources of diversity than space. For the information theoretic analysis of the latter, the concept of *sample-mean outage* is introduced, that allows to apply information theoretic measures, like capacity or cutoff-rate to time selective channels. It will be shown, that in contrast to common belief, correlated fading may offer better performance than uncorrelated fading permits.

1. INTRODUCTION

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 are independent Rayleigh distributed, it is well known [1, 2, 3], that for high enough transmit power 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 [4, 5], 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 show 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. 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 real digital modulation schemes on system performance by cutoff-rate analysis.

2. SYSTEM MODEL

In the following 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

$$\mathbf{y} = \mathbf{H}\mathbf{T}\mathbf{P}^{\frac{1}{2}}\mathbf{s} + \mathbf{n}, \tag{1}$$

where $\mathbf{s} \in \mathcal{C}^L$ is the L-dimensional data vector with zero mean and unity covariance matrix, while $\mathbf{P} \in \mathcal{R}_+^{L\times L}$ is a positive definite diagonal matrix used to set the transmit power for each data stream with total transmit power given by $P_\mathrm{T} = \mathrm{tr}\,\mathbf{P}$, and finally the matrix $\mathbf{T} \in \mathcal{C}^{N\times L}$ 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 beam-forming. The channel is modeled by the matrix $\mathbf{H} \in \mathcal{C}^{M\times N}$ with possibly correlated complex zero mean Gaussian entries. The receive signal vector $\mathbf{y} \in \mathcal{C}^M$ is corrupted by additive zero mean white Gaussian noise $\mathbf{n} \in \mathcal{C}^M$ with power σ_n^2 per receive antenna.

3. FADING CORRELATIONS

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

$$\mathbf{H} = \frac{1}{\sqrt{\mathrm{tr}\,\mathbf{R_T}}} \cdot \mathbf{R}_\mathrm{R}^{\frac{1}{2}}\,\mathbf{G}\,\mathbf{R}_\mathrm{T}^{\frac{1}{2}}\,, \tag{2}$$

where $\mathbf{R}_\mathrm{R} = \mathrm{E}\left\{\mathbf{H}\mathbf{H}^\mathrm{H}\right\}$ is the $M \times M$ receive correlation matrix, and $\mathbf{R_T} = \mathrm{E}\left\{\mathbf{H}^\mathrm{H}\mathbf{H}\right\}$ is the $N \times N$ transmit correlation matrix, while $\mathbf{G} \in \mathcal{C}^{M\times N}$ is a random matrix with independent, *zero mean, unity variance* complex entries. In the following we will assume that $\mathbf{G}$ is drawn from a complex Gaussian distribution leading to correlated Rayleigh fading. Note, that $\mathrm{tr}\,\mathbf{R_T} = \mathrm{tr}\,\mathbf{R}_\mathrm{R} = \mathrm{E}\left\{\|\mathbf{H}\|_\mathrm{F}^2\right\} = \sum_{m=1}^{M}\sum_{n=1}^{N}\mathrm{E}\left\{|\mathrm{H}_{m,n}|^2\right\}$, which is the channel's total power amplification. One can distinguish four fundamental cases

case	Rx-side	Tx-side	type of fading
1	uncorr.	uncorr.	uncorrelated
2	uncorr.	corr.	semi-correlated
3	corr.	uncorr.	semi-correlated type 2
4	corr.	corr.	fully-correlated

For brevity we will restrict the discussion to the first two cases, as furthermore the semi-correlated model is valid for urban mobile radio channels, as has been shown by a recent measurement campaign taken in downtown Helsinki [6].

0-7803-7448-7/02/$17.00 ©2002 IEEE

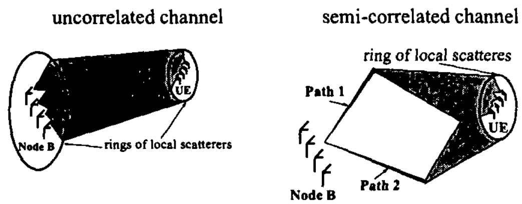

Fig. 1. Geometrical interpretation of uncorrelated and semi-correlated fading channels.

3.1. Uncorrelated Rayleigh channel

Such a channel may arise if both the transmitter and the receiver live in a rich scattering environment. The result will be independent Rayleigh fading from each transmit to each receive antenna. We have $\mathbf{R_T} = M \cdot \mathbf{I}_N$, $\mathbf{R_R} = N \cdot \mathbf{I}_M$, producing $\mathbf{H} = \mathbf{G}$ and $\text{tr}\,\mathbf{R_R} = \text{tr}\,\mathbf{R_T} = M \cdot N$. Note, that $\mathbf{I}_n$ is the $n \times n$ identity matrix.

3.2. Semi-correlated 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. Hence, we can write

$$\mathbf{H} = \sqrt{\frac{N}{\text{tr}\,\mathbf{AA}^{\text{H}}}} \cdot \mathbf{ZA}^{\text{T}},$$

where $\text{A} \in \mathcal{C}^{N \times K}$ is an array steering matrix containing K array response vectors of the transmitting antenna array corresponding to K directions of departure (DOD), and $\mathbf{Z} \in \mathcal{N}_{\mathcal{C}}^{M \times K}(0,1)$ has zero mean 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 (2) we have

$$\mathbf{R_T} = \frac{M \cdot N}{\text{tr}\,\mathbf{AA}^{\text{H}}} \cdot \mathbf{A}^{*}\mathbf{A}^{\text{T}}, \quad \mathbf{R_R} = N \cdot \mathbf{I}_M. \tag{3}$$

Note, that the total power amplification of this channel is normalized to $\text{tr}\,\mathbf{R_R} = \text{tr}\,\mathbf{R_T} = M \cdot N$, which is the same as in the uncorrelated case.

4. THE 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. Therefore their instantaneous capacity $C(t)$ is a random

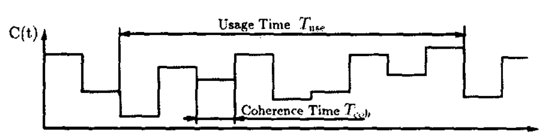

Fig. 2. Example of temporal evolution of a block fading channel

process, too. Figure 2 shows an example for $C(t)$ of a channel which properties remain constant during the *coherence time* T_{coh} and afterwards change to a new, independent realization (block fading). The capacity of such a channel depends on the ratio between usage time T_{use} and coherence time T_{coh}.

- For $T_{\text{use}} \to \infty$, 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*
$$C_{\text{erg}} = \text{E}\{C\}\ . \tag{4}$$
- 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.
$$C_{\text{out}}(p) \quad : \quad \text{Prob}\,\{C < C_{\text{out}}(p)\} = p \tag{5}$$
- 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
$$C' = \frac{1}{m} \sum_{k=1}^{m} C_k\ ,$$
where C_k are the instantaneous capacities corresponding to independent channel realizations, the appropriate characterization of capacity is the proposed *sample-mean outage capacity* $C_{soc}(p, m)$
$$C_{\text{soc}}(p,m) \quad : \quad \text{Prob}\,\{C' < C_{\text{soc}}(p,m)\} = p\ . \tag{6}$$

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

$$C_{\text{soc}}(p,1) = C_{\text{out}}, \quad \text{and} \quad C_{\text{soc}}(p,\infty) = C_{\text{erg}}\ . \tag{7}$$

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 get its hands on the instantaneous capacity of the channel, which may involve establishment of a feedback link from the receiver.
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, but comes at the price of additional delay time.

5. INSTANTANEOUS CHANNEL CAPACITY

Assuming the system model defined in section 2 and complete *receiver* side channel knowledge, the instantaneous transinformation $\mathcal{I}_{\mathbf{s} \to \mathbf{y}}$ from transmitted Gaussian signal $\mathbf{s}$ to received signal $\mathbf{y}$ is given as

$$\mathcal{I}_{\mathbf{s} \to \mathbf{y}}\left(\mathbf{H}^{\text{H}}\mathbf{H}, \mathbf{T}, \mathbf{P}\right) = \log_2 \det\left(\mathbf{I}_L + \frac{1}{\sigma_n^2}\mathbf{T}^{\text{H}}\mathbf{H}^{\text{H}}\mathbf{HTP}\right)\ . \tag{8}$$

The instantaneous channel capacity is the maximum transinformation

$$C\left(\mathbf{H}^{\text{H}}\mathbf{H}\right) = \max_{\mathbf{T},\mathbf{P}} \mathcal{I}_{\mathbf{s} \to \mathbf{y}}\left(\mathbf{H}^{\text{H}}\mathbf{H}, \mathbf{T}, \mathbf{P}\right),\ \text{s.t. tr}\,\mathbf{P} = P_{\text{T}}, \tag{9}$$

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

6. TRANSMITTER SIDED CHANNEL KNOWLEDGE

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

6.1. Complete channel knowledge

Assuming that the transmitter exactly knows the channel matrix $\mathbf{H}$ at each transmit time instant, it is well known [1] that by following the procedure:

- EVD: $\mathbf{H}^{\mathrm{H}}\mathbf{H} = \mathbf{V}\Lambda\mathbf{V}^{\mathrm{H}}$
- Set $\mathbf{T} = \mathbf{V}$ and choose $\mathbf{P}$ by Waterfilling based on Λ

the instantaneous transinformation is maximized. Hence, the average transinformation is maximized, too.

6.2. No channel knowledge

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

- Set $\mathbf{T} = \mathbf{I}_N$ and $\mathbf{P} = \frac{P_T}{N} \cdot \mathbf{I}_N$
- 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.

6.3. 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 $\mathbf{R_T}$ and following the procedure

- EVD: $\mathrm{E}\left\{\mathbf{H}^{\mathrm{H}}\mathbf{H}\right\} = \mathbf{R_T} = \mathbf{V}'\Lambda'\mathbf{V}'^{\,\mathrm{H}}$
- Set $\mathbf{T} = \mathbf{V}'$ and choose $\mathbf{P}$ by Waterfilling based on Λ'

will maximize $\mathcal{I}_{\mathbf{s}\to\mathbf{y}}\left(\mathrm{E}\left\{\mathbf{H}^{\mathrm{H}}\mathbf{H}\right\}, \mathbf{T}, \mathbf{P}\right)$, i.e. the transinformation of the average channel. While complete channel knowledge allows for maximizing the average transinformation, *average* channel knowledge allows for maximizing the transinformation of the *average channel*. As was presented in [9], this yields close to optimum performance in MIMO-systems in semi-correlated fading environments, where the performance may be even better than in the uncorrelated case. The procedure above is called *MIMO Downlink Eigenbeamforming* [8].

7. 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 feasible to compute cutoff-rates for real modulation schemes in MIMO systems. The cutoff rate is useful because of the cutoff-rate theorem [7], 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

$$P_w < 2^{-n\cdot(R_0 - R_b)}, \tag{10}$$

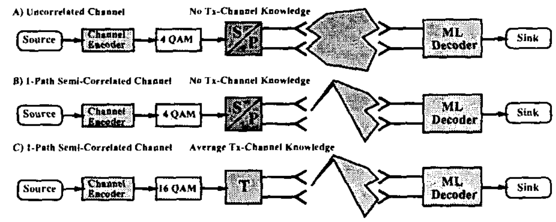

Fig. 3. System setup for comparing cutoff-rates for different fading correlations and transmitter channel knowledge. The raw data rate is fixed to 4 bits/sec/Hz.

provided the binary code rate $R_b := \frac{k}{n} \cdot \log_2 q$ is less than the cutoff-rate

$$R_0 = -\log_2 \int_{\mathbb{C}^M} \left(\sum_{\mathbf{s}\,\in\,\mathcal{M}} \frac{1}{q} \sqrt{p(\mathbf{y}|\mathbf{s})} \right)^2 d\mathbf{y}, \tag{11}$$

where $\mathcal{M}$, with $|\mathcal{M}| = q$ is the set of code symbols (input alphabet) and $p(\mathbf{y}|\mathbf{s})$ is the probability density function of the received signal $\mathbf{y}$ given the transmitted code symbol s. To apply this to our MIMO system, we look at the data vector $\mathbf{s}$ as a q-ary code symbol, where each component s_k, with $1 \le k \le L$ 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 $|\mathcal{M}| = q = q_1 \cdot q_2 \cdots \cdot q_L$. By labeling the elements of $\mathcal{M} = \{\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_q\}$ the instantaneous cutoff-rate can be written as

$$R_0 = \log_2(q) - \log_2\left(1 + \frac{2}{q}\sum_{p=1}^{q-1}\sum_{t=p+1}^{q} \exp\left(-\frac{1}{4}\|\mathbf{b}_p - \mathbf{b}_t\|_2^2\right)\right), \tag{12}$$

with $\mathbf{b}_p = \frac{1}{\sigma_n}\mathbf{HTP}^{\frac{1}{2}}\mathbf{s}_p$. Ergodic, outage and sample-mean outage cutoff-rates can be computed accordingly to the discussion in section 4.

8. 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 simulated 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 1-path semi-correlated 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 to the standard beam-width of the antenna array – in this case 60° in bore-side direction. In the semi-correlated case we will distinguish between *no*, and *average* transmitter channel knowledge. The uncoded (raw) bit rate shall be fixed to $r_{raw} = 4$ bits/sec/Hz in all cases, as to implement a given service. For the uncorrelated case we use 4-QAM modulation and transmit two independent data streams - one over each antenna. The same holds for the semi-correlated case with *no* channel knowledge, as the transmitter is not aware of the channel

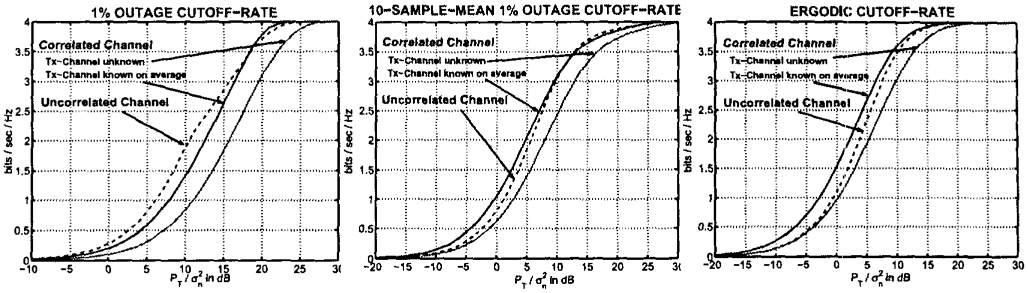

Fig. 4. Results: ergodic, outage, and sample-mean outage cutoff-rates

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 catch up with the required raw bit rate, the modulation scheme is changed to 16QAM. Figure 3 summarizes the test bed, while Figure 4 shows the 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 air-time – the more attractive the semi-correlated channel becomes.
3. If there is no transmit channel knowledge however, the semi-correlated channel performs worst, no matter how much time diversity is available.
4. Arming the transmitter with average channel knowledge, the semi-correlated channel turns out to yield the best performance – in fact beating the performance of the uncorrelated channel – provided there is enough time diversity available.
5. Looking at the 10-sample-mean outage cutoff-rate, we see, that the amount of 10 independent channel realizations during air-time (or per code-word, if interleaving is used), suffices for the semi-correlated channel to qualify for the winner for all code-rates less than about 0.8.
6. If the number of independent channel realizations during air-time (or per code-word, if interleaving is used) is further extended, the ergodic cutoff-rate shows, that the advantage of the semi-correlated channel is still improving and extended for virtually all code-rates.

9. 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. To this end a transmit scheme was presented that efficiently exploits fading correlations while depending solely on average channel properties, and making correlated fading a blessing. 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 real digital modulation schemes, semi-correlated fading channels in practice offer superior performance in the interesting transmit power range, provided a modest amount of time or frequency diversity is available.

10. REFERENCES

[1] E. Telatar, "Capacity of Multi-Antenna Gaussian Channels", *AT&T-Bell Technical Memorandum*, 1995.

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

[3] J. Salz and J. Winters, "Effect of Fading Correlation on Adaptive Arrays in Digital Mobile Radio", *IEEE Trans. Vehicular Technology*, vol. 43, no. 4, pp. 1049-1057, November 1994.

[4] C. Chuah, J. M. Kahn and D. Tse, "Capacity of multi-antenna array systems in indoor wireless environment", *Globecom*, 1998.

[5] D-S. Shiu and G. J. Foschini and M. J. Gans and J. M. Kahn, "Fading Correlation, and its Effect on the Capacity of Multielement Antenna Systems", *IEEE Trans. Communications*, vol. 48, no. 3, pp. 502-513, 2000.

[6] J. Laurila and K. Kalliola and M. Toeltsch and K. Hugel and P. Vainikainen and E. Bonek, "Wideband 3-D characterization of mobile radio channels in urban environment", *IEEE Trans. Antennas and Propagation (in press)*, 2001.

[7] J. L. Massey, "Coding and modulation in digital communications", *Int. Zürich Seminar*, Sindelfingen, Germany, March 1974.

[8] M. T. Ivrlač and T. P. Kurpjuhn and C. Brunner and W. Utschick, "Efficient use of fading correlations in MIMO systems", *Proc. 54th IEEE Vehicular Technology Conf. (VTC '01)*, pages 2763-2767, Atlantic City, USA, October 2001.

[9] M. T. Ivrlač and Josef A. Nossek, "Correlated Fading in MIMO Systems - Blessing or Curse ?", *39th Annual Allerton Conference on Communication, Controll and Computing*, (in press), Monticello, IL, USA, October 2001.