Efficient Use of Fading Correlations in MIMO Systems

Michel T. Ivrlač[1], Tobias P. Kurpjuhn[1], Christopher Brunner[2], Wolfgang Utschick[1]

1. Institute for Circuit Theory and Signal Processing
Munich University of Technology
80290 Munich, Germany
phone: +49 (89) 289285 {11,09,24}
email: {ivrlac, kurpjuhn, utschick}@nws.ei.tum.de

2. Arraycomm
2480 N. First Street, Suite 200
San Jose, CA 95131, USA
phone: +1 (408) 452 1922
email: cbrunner@arraycomm.com

Abstract— **We investigate the effects of both fading correlations and transmitter channel knowledge in multiple element antenna (MEA) communication systems. While for independent and identically distributed, fades between receive and transmit antennas, pioneering work showed that a huge increase in capacity is possible for MEAs compared to a single antenna system, recent contributions warn that fading correlations destroy most of this advantage. While this is true for zero transmitter channel knowledge, we will show however, that long-term average channel state information enables the transmitter to efficiently use the fading correlations to its advantage and offers the potential to even increase capacity beyond the one possible for independent fading. A conceived transmit technique is presented that efficiently makes use of fading correlations, and also provides optimum choice of digital modulation schemes that carry the information.**

I. Introduction

Communication systems making use of multiple antennas at both sides of the link – so called multiple-input multiple-output (MIMO) antenna systems – recently have drawn considerable attention in the area of wireless communications. If the fades between pairs of transmit and receive antennas are independent and identically 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 zero transmitter channel knowledge. 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. After introducing the system model we will discuss both the impact of fading correlations and transmitter channel knowledge on capacity and propose an efficient scheme to use fading correlations in advantage. We will also consider the effect of real digital modulation schemes on system performance by cutoff rate analysis and deal with the problem of optimum choice of modulation schemes. Finally we will show how to apply fading correlation knowledge to orthogonal frequency division multiplexing (OFDM) in a frequency selective fading environment.

II. System Model

In the following we will focus on the problem of transmitting L independent data streams over a wireless channel using $N \geq L$ transmit and M receive antennas. Even though a broadband communication system in general will experience a frequency selective channel, OFDM can be used to transform this frequency selective channel into many frequency flat channels. In the following we will therefore assume a frequency flat but possibly correlated Rayleigh fading wireless channel, leading to the following 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 spatial 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 Gaussian noise $\mathbf{n} \in \mathcal{C}^M$ with covariance matrix $\mathrm{E}\left\{\mathbf{n}\mathbf{n}^\mathrm{H}\right\} = \sigma_n^2\mathbf{R}_n$, where σ_n^2 is the average noise power per receive antenna, i.e. $\sigma_n^2 = \mathrm{tr}\,\mathrm{E}\left\{\mathbf{n}\mathbf{n}^\mathrm{H}\right\}/M$. Note that $\mathrm{tr}\,\mathbf{R}_n = M$.

III. Channel Capacity

Applying an eigenvalue decomposition to

$$\begin{aligned}\mathbf{H}^\mathrm{H}\mathbf{R}_n^{-1}\mathbf{H} &= \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\mathrm{H} \\ &= [\mathbf{V}_1\ \mathbf{V}_2]\begin{bmatrix}\mathbf{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_2\end{bmatrix}[\mathbf{V}_1\ \mathbf{V}_2]^\mathrm{H},\end{aligned} \tag{2}$$

where $\mathbf{\Lambda}_1$ contains the L largest and $\mathbf{\Lambda}_2$ the $N-L$ remaining eigenvalues, while the eigenvector matrix $\mathbf{V}$ is partitioned accordingly into sub-matrices $\mathbf{V}_1$ and $\mathbf{V}_2$, respectively, the ergodic capacity of this system can be expressed as [1]

$$C = \max_{\mathbf{T},\mathbf{P},L} \mathrm{E}\left\{\log_2 \det\left(\mathbf{I} + \frac{1}{\sigma_n^2}\mathbf{T}^\mathrm{H}\mathbf{V}\mathbf{\Lambda}\mathbf{V}^\mathrm{H}\mathbf{T}\mathbf{P}\right)\right\}, \tag{3}$$

where the expectation is carried out over the different realizations of the channel matrix $\mathbf{H}$. The transmitter can maximize the average transinformation by beam-forming via $\mathbf{T}$, power control via $\mathbf{P}$, and choice of the number of data streams L. To what extend this maximization can be carried out, depends both on the statistical properties of $\mathbf{H}$ and the amount of knowledge the transmitter can acquire about them.

IV. Transmitter Channel Knowledge

Let us start with the discussion of the impact of transmitter channel knowledge on transinformation. We will look at three different cases: the transmitter is allowed to know the channel instantaneously, on average only, or not at all.

A. *Instantaneous channel knowledge*

Assuming that the transmitter exactly knows the channel matrix $\mathbf{H}$ at each transmit time instant, it is well known that transinformation reaches channel capacity by setting $L = \text{rank}\,\mathbf{H}$, $\mathbf{T} = \mathbf{V}_1$ and choosing $\mathbf{P}$ by instantaneous water-filling [1], [6] based on the eigenvalues $\mathbf{\Lambda}_1$. This is, of course the best case scenario.

B. *No channel knowledge*

If there is no channel knowledge at all available to the transmitter, setting $L = N$, $\mathbf{T} = \mathbf{I}$, obviously is the only reasonable choice. Because of lack of channel knowledge, water-filling cannot be performed either and has to be replaced by an equal power distribution, i.e. $\mathbf{P} = (P_\mathrm{T}/N) \cdot \mathbf{I}$. In this scenario each antenna transmits an independent data stream with the power being shared equally.

C. *Long term average channel knowledge*

While instantaneous 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 we know $\mathrm{E}\{\mathbf{H}^\mathrm{H}\mathbf{R}_n^{-1}\mathbf{H}\}$, an eigenvalue decomposition leads to

$$\begin{aligned}\mathrm{E}\{\mathbf{H}^\mathrm{H}\mathbf{R}_n^{-1}\mathbf{H}\} &= \mathbf{V}'\mathbf{\Lambda}'\mathbf{V}'^{\,\mathrm{H}} \\ &= [\mathbf{V}_1'\ \mathbf{V}_2'] \begin{bmatrix} \mathbf{\Lambda}_1' & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_2' \end{bmatrix} [\mathbf{V}_1'\ \mathbf{V}_2']^\mathrm{H}, \end{aligned} \tag{4}$$

where $\mathbf{\Lambda}_1'$ contains the L largest and $\mathbf{\Lambda}_2'$ the $N-L$ remaining eigenvalues of $\mathrm{E}\{\mathbf{H}^\mathrm{H}\mathbf{R}_n^{-1}\mathbf{H}\}$, while the eigenvector matrix $\mathbf{V}'$ is partitioned accordingly into sub-matrices $\mathbf{V}_1'$ and $\mathbf{V}_2'$, respectively. By setting $L = \text{rank}\,\mathrm{E}\{\mathbf{H}^\mathrm{H}\mathbf{R}_n^{-1}\mathbf{H}\}$, $\mathbf{T} = \mathbf{V}_1'$ and choosing $\mathbf{P}$ by water-filling based on the average eigenvalues $\mathbf{\Lambda}_1'$, the function

$$J(\mathbf{T},\mathbf{P},L) = \log_2 \det \mathrm{E}\left\{\mathbf{I} + \frac{1}{\sigma_n^2}\mathbf{T}^\mathrm{H}\mathbf{V}\mathbf{\Lambda}\mathbf{V}^\mathrm{H}\mathbf{T}\mathbf{P}\right\}, \tag{5}$$

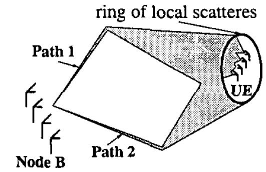

Fig. 1. A semi-correlated 2-path channel: from the transmitter's point of view the channel is spatially correlated as the receiver can be reached through just two narrow spatial directions, while from the receiver's point of view the channel has no spatial structure due to its rich scattering environment.

is maximized. This is of course *not* equivalent to transinformation, but actually an upper bound, for comparing to (3) the expectation operator has moved inside the $\log_2$ and det operators. Later we will show however (cf. Fig 2), that maximizing (5) is almost equivalent to maximizing transinformation (3). Viewing $\mathbf{T}$ as beam-forming, setting $\mathbf{T} = \mathbf{V}_1'$ will be called *eigenbeamforming*. Each data stream is said to be transmitted over an *eigenbeam*.

V. Fading Correlations

Let us now have a look at some statistical properties of the channel. In the following we will investigate two different cases, namely channels having spatial fading correlations and channels that are spatially uncorrelated.

A. *Uncorrelated Rayleigh fading*

Such a channel may arise if both transmitter and receiver live in a rich scattering environment. The result will be independent Rayleigh fading from each transmit to each receive antenna. The channel matrix can be modeled as

$$\mathbf{H} \in \mathcal{N}_\mathcal{C}^{M \times N}(0,1). \tag{6}$$

The entries are i.i.d. zero mean, unity variance complex Gaussian random variables. Note that the total power amplification of this channel is given by $\mathrm{E}\{\|\mathbf{H}\|_F^2\} = N \cdot M$.

B. *Semi-Correlated K-path channels*

Imagine a scenario where the *transmitter* is 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 be only a few dominant remote scattering or reflecting objects (see Fig. 1). This assumption is validated for urban mobile radio channels, by a recent measurement campaign taken in downtown Helsinki [7]. We will model such a scenario by

$$\mathbf{H} = \sqrt{\frac{N}{\text{tr}\,\mathbf{A}\mathbf{A}^\mathrm{H}}} \cdot \mathbf{G}\mathbf{A}^\mathrm{T}, \tag{7}$$

where $\mathrm{A} \in \mathcal{C}^{N \times K}$ is an array steering matrix containing K array response vectors of the transmitting antenna array corresponding to K directions of departure (DOD), and $\mathbf{G} \in$

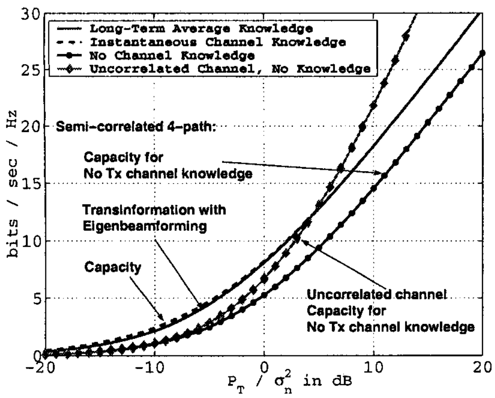

Fig. 2. Comparison of capacity and transinformation for semi-correlated and uncorrelated channels with and without long-term channel knowledge. Note that in the uncorrelated case, having no channel knowledge is equivalent to having long-term channel knowledge.

$\mathcal{N}_C^{M\times K}(0,1)$ has zero mean i.i.d. Gaussian random entries. Angle spread is easily modeled by a high enough number of discrete DODs. The total power amplification of this channel is normalized to $\mathrm{E}\left\{||\mathbf{H}||_F^2\right\} = N\cdot M$, which is the same as in the uncorrelated case. While both $\mathbf{G}$ and $\mathbf{A}$ are random variables, they vary on fairly different time scales, as $\mathbf{G}$ models fast Rayleigh fading induced by *small scale* movements of the mobile receiver, while $\mathbf{A}$ represents the geometrical structure of the propagation channel, and varies with *large scale* movements, that usually take place at much longer a time-scale than fast fading, especially for large receiver-transmitter distances. From (7) follows

$$\mathbf{R}_\mathrm{T} := \mathrm{E}_\mathbf{G}\left\{\mathbf{H}^\mathrm{H}\mathbf{R}_n^{-1}\mathbf{H}\right\} = \frac{M\cdot N}{\mathrm{tr}\,\mathbf{A}\mathbf{A}^\mathrm{H}}\cdot\mathbf{A}^*\mathbf{A}^\mathrm{T} \in \mathcal{C}^{N\times N}, \quad (8)$$

which is called *transmitter covariance matrix*, and is independent of $\mathbf{R}_n$. The operator $\mathrm{E}_\mathbf{G}\{.\}$ denotes expectation with respect to $\mathbf{G}$, i.e. averaging over fast fading. Usually $\mathbf{R}_\mathrm{T}$ will exhibit spatial correlations, possibly with numerical rank deficiency. Note that the *receiver covariance matrix*

$$\mathbf{R}_\mathrm{R} := \mathrm{E}_\mathbf{G}\left\{\mathbf{H}\mathbf{H}^\mathrm{H}\right\} = N\cdot\mathbf{I} \in \mathcal{C}^{M\times M}, \quad (9)$$

corresponds to a spatially uncorrelated scenario as requested by the model (see also Fig. 1).

VI. Capacity of semi-correlated K-path channels

To evaluate the capacity of semi-correlated channels with and without long-term average channel knowledge, we simulated a $M = N = 8$ antenna system, where the antennas formed a omni-directional uniform linear array. We used a 4-path semi-correlated channel and an uncorrelated channel for comparison. The four paths had zero angle spread and random directions of departure. Fig. 2 shows the results. There are four major points to stress here. First, if there is no transmit channel knowledge the spatial correlations reduce capacity compared to the uncorrelated case. Second, if long-term average transmit channel knowledge is used, the picture changes: for low transmit powers up to a cross over point, the semi-correlated channel indeed offers higher capacity than the uncorrelated one , which is due to antenna gain that can be exploited by knowing the long-term average channel structure. Third, for the semi-correlated channel the difference between long-term average and instantaneous channel knowledge is marginal and disappears for high transmit powers. Fourth, at high transmit powers the uncorrelated channel gets better and better compared to the semi-correlated 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 in the next sections.

VII. 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 [8], 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

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

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

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

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

$$\mathcal{M} = \mathcal{M}_1\times\mathcal{M}_2\times\cdots\times\mathcal{M}_L, \quad (12)$$

is the Cartesian product of the individual alphabet sets, with $|\mathcal{M}| = q = q_1\cdot q_2\cdot\cdots\cdot q_L$. By labeling the elements of $\mathcal{M} = \{\mathbf{s}_1,\mathbf{s}_2,\cdots,\mathbf{s}_q\}$ the cutoff rate can be written as

$$R_0 = \log_2(q) - \log_2\left(1+\frac{2}{q}\sum_{p=1}^{q-1}\sum_{t=p+1}^{q}\exp\left(-\frac{1}{4}||\mathbf{b}_p-\mathbf{b}_t||_2^2\right)\right), \quad (13)$$

with $\mathbf{b}_p = \frac{1}{\sigma_n}\mathbf{R}_n^{-\frac{1}{2}}\mathbf{H}\mathbf{T}\mathbf{P}^{\frac{1}{2}}\mathbf{s}_p$. The ergodic cutoff rate is the expectation of (13) taken with respect to H.

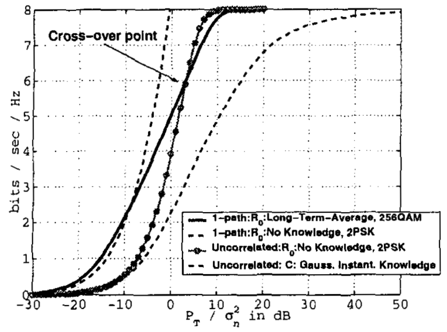

Fig. 3. Ergodic cutoff rates for semi-correlated 1-path and uncorrelated channels with and without long-term average knowledge.

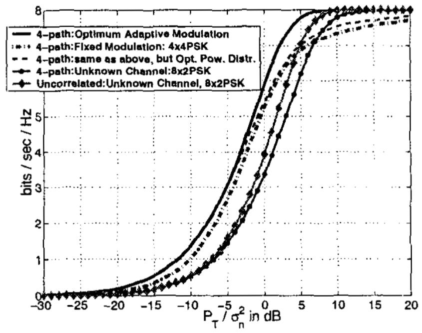

Fig. 4. Ergodic cutoff rates for semi-correlated 4-path and uncorrelated channels using fixed and adaptive modulation.

VIII. Cutoff Rate Comparison

We assume a $M = N = 8$ antenna MIMO system, and compute the cutoff rates for a 1-path semi-correlated and for an uncorrelated channel. Note that the semi-correlated channel has unity rank. We used quadrature amplitude modulation (QAM) and fixed the raw data rate to 8 bits per channel use. For the uncorrelated channel each of the 8 antennas therefore transmits a data stream with 1 bit per channel use (binary phase shift keying, 2PSK). The same holds for the semi-correlated channel with no channel knowledge. In the case of available long-term average channel knowledge, the transmitter is aware of the rank deficiency and therefore transmits a single data stream over the strongest eigenbeam only. To achieve a raw data rate of 8 bits per channel use, the modulation scheme is changed to 256QAM. Fig. 3 shows the results. Let us stress the major points: First, again we see a crossover point between semi-correlated channels using eigenbeamforming and uncorrelated channels, but since the cutoff rates are bounded, we can judge its position better than in Fig. 2: for code rates less than 3/4, the semi-correlated channel using eigenbeamforming outperforms the uncorrelated channel up to the antenna gain of 9dB, while for higher code rates the loss is limited to 4.3dB[1] , instead of growing unbound as in Fig.2. Second, having spatial fading correlations without the transmitter knowing about them is even more disastrous than suggested by the capacity analysis in Fig. 2. Not only is there a loss due to no exploitable antenna gain, for high code rates there is additional loss, which turns out to be due to distortion of the received signal constellation [9]. Third, knowing about fading correlations can actually lead to higher capacity then is possible for uncorrelated channels even in the best case of having instantaneous transmit channel knowledge and Gaussian signal distribution (see dotted capacity line in Fig. 3).

[1] Using 256QAM asymptotically needs 22.3dB more power than 2PSK, but as the power is concentrated onto one stream instead of being shared on 8 streams, there is a gain of 9dB and because of an additional antenna gain of 9dB, the asymptotic loss turns out to be 22.3-9-9 = 4.3dB. If the number of antennas is reduced below 6, the loss turns into gain, e.g. 2dB for N=M=4.

IX. Adaptive Modulation and Rank Search

We now want to address the problem of finding the number L of transmitted data streams that is optimum in a given situation. For Gaussian distributed signals the answer is simple: set $L = N$ and use the water-filling policy to optimally share the transmit power. For modulated signals that is no longer applicable, as each data stream has a finite raw data rate. It makes more sense to ask: "How many bits should be transmitted over each data stream?". The answer to this one is *adaptive modulation*. The idea is to transmit more bits over a stream where the associated eigenbeam has a high eigenvalue, and transmit less bits over other streams. To illustrate, we use a $M = N = 3$ system, where the transmit antennas form a ULA with $\lambda/2$ antenna spacing and look into a semi-correlated channel that supports two DODs with different angle spread and attenuation, as depicted in Fig. 5. The eigenvalues of $\mathrm{E}\{\mathbf{H}^{\mathrm{H}}\mathbf{H}\}$ compute to 4.75, 3.23, and 1.02, respectively. We fix the raw data rate to 6 bits per channel use, and compute the average cutoff rate for different distributions of bits per data stream. The averaging is done by computing realizations of the channel matrix according to

$$\frac{1}{\sqrt{N}} \cdot \mathbf{G} \cdot \left(\mathrm{E}\left\{\mathbf{H}^{\mathrm{H}}\mathbf{H}\right\}\right)^{\frac{1}{2}}, \quad \text{where } \mathbf{G} \in \mathcal{N}_{\mathcal{C}}^{N\times N}(0,1).$$

Note, that using a non Gaussian random matrix $\mathbf{G}$ above, would lead to other fading statistics than Rayleigh. The transmit power is shared equally between data streams. Using QAM the results are given in Table I. For low transmit power it is best to focus on the strongest eigenbeam only and use

TABLE I

Ergodic cutoff rates for the scenario from Fig. 5

$P_{\mathrm{T}} / \sigma_n^2 \longrightarrow$			-10dB	4dB	12dB	17dB
Bits per stream ↓			ergodic cutoff rate			
6	0	0	**0.160**	1.95	4.13	5.41
5	1	0	0.138	1.96	4.22	5.45
4	2	0	0.138	**2.17**	4.94	5.83
3	3	0	0.137	2.15	**5.07**	**5.89**
4	1	1	0.103	1.74	4.37	5.62
3	2	1	0.104	1.84	4.88	**5.85**
2	2	2	0.104	1.87	4.71	5.75

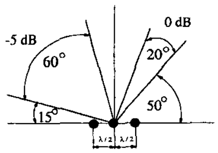

Fig. 5. Example of transmitter side angle spread

64QAM. For medium transmit powers it pays off to open up and share the power with a second data stream and switch to 16QAM/4QAM or at a little higher power to 2×8QAM. Only at very high transmit powers a full rank transmission is reasonable. The optimum number of data streams therefore depends on the long-term average channel situation, the used transmit power, the modulation schemes, and also on the raw data rate that has to be kept up. To show the effects of an optimum adaptive modulation, Fig. 4 shows the average cutoff rates for a $M = N = 8$ antenna system, transmitting at a raw data rate of 8 bits per channel use, over a 4-path semi-correlated channel (supporting independent transmission of up to four data streams). The 4 paths have zero angle spread and random DODs. The averaging is done both over short-term (fading) and long-term (DOD) properties of the channel. Let us state the major points. First, use of the fixed 4x4PSK modulation is inferior as the transmitter cannot react to changing average eigenvalue profile. It gets disastrous at higher code rates, were it gets even outperformed by dropping eigenbeamforming altogether. Second, applying additional optimum power distribution among the eigenbeams improves the performance at lower code rates, but also suffers at higher code rates. Third, optimum adaptive modulation saves the day, as it constantly improves performance at all code rates, especially at higher ones, yielding always the best performance. Note, that there is no cross-over point with the uncorrelated case any more.

X. Application to OFDM

A broadband communication system usually will experience a frequency selective channel. Assuming a multi-path MIMO channel with path delay times τ_k:

$$\mathbf{H}(t,\tau) = \sum_{k=1}^{d} \mathbf{H}_k\, \delta(t-\tau_k), \tag{14}$$

and cyclic prefixed OFDM with N_c sub-carriers with baseband frequencies $f_n = \frac{1}{T} \cdot \frac{n}{N_c}$, where T is the time for a channel use, and $0 \le n < N_c$, the frequency selective channel (14) evolves into N_c frequency flat MIMO channels described by N_c channel matrices

$$\tilde{\mathbf{H}}_n = \sum_{k=1}^{d} \mathbf{H}_k \cdot \exp\left(-\mathrm{j}2\pi \frac{\tau_k}{T} \frac{n}{N_c}\right). \tag{15}$$

The ergodic capacity therefore reads as

$$C = \max_{\mathbf{T},\mathbf{P},L} \mathrm{E}\left\{ \frac{1}{N_c} \sum_{n=0}^{N_c-1} \log_2 \det\left(\mathbf{I} + \frac{1}{\sigma_n^2} \mathbf{T}^{\mathrm{H}} \tilde{\mathbf{H}}_n^{\mathrm{H}} \tilde{\mathbf{H}}_n \mathbf{T}\mathbf{P}\right) \right\}. \tag{16}$$

By moving all averaging operations inside the $\log_2$ and det operators, we define a cost function

$$J(\mathbf{T},\mathbf{P},L) = \log_2 \det\left(\mathbf{I} + \frac{1}{\sigma_n^2} \mathbf{T}^{\mathrm{H}} \mathbf{R} \mathbf{T} \mathbf{P}\right), \tag{17}$$

where

$$\mathbf{R} = \frac{1}{N_c} \sum_{n=0}^{N_c-1} \mathrm{E}\left\{\tilde{\mathbf{H}}_n^{\mathrm{H}} \tilde{\mathbf{H}}_n\right\}. \tag{18}$$

For temporally uncorrelated channel taps, $\mathrm{E}\left\{\mathbf{H}_k^{\mathrm{H}} \mathbf{H}_{k'}\right\} = \mathrm{E}\left\{\mathbf{H}_k^{\mathrm{H}} \mathbf{H}_k\right\} \cdot \delta_{k,k'}$, (18) simplifies to

$$\mathbf{R} = \sum_{k=1}^{d} \mathrm{E}\left\{\mathbf{H}_k^{\mathrm{H}} \mathbf{H}_k\right\}, \tag{19}$$

and eigenbeamforming can be applied by eigenanalysis of $\mathbf{R}$.

XI. Conclusion

The capacity of MIMO systems depends both on the statistical properties of the channel and on the knowledge about those properties. While for no transmitter channel knowledge correlated fading is disastrous for capacity, having the transmitter acquire the channel properties on average can actually lead to capacity improvement over uncorrelated fading channels. A transmit scheme was presented that efficiently exploits fading correlations while depending solely on average channel properties. Cutoff rate analysis showed that for real digital modulation schemes, correlated fading channels in practice offer superior performance in the whole transmit power range. A key to this performance gain turns out to be adaptive modulation. A method for achieving optimum adaptive modulation was presented that is based on the channel's average cutoff rate. Finally, we showed how to make efficient use of fading correlations in OFDM based broadband communication systems.

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 Personla Communications*, vol. 6-3, 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. 43-4, pp. 1049–1057, Nov. 1994.

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

[5] D-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading Correlation, and its Effect on the Capacity of Multielement Antenna Systems," *IEEE Trans. Communications*, vol. 48-3, pp. 502–513, 2000.

[6] R. G. Gallager, *Information Theory and Reliable Communication*, John Wiley & Sons, 1968.

[7] 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 and Propagation (in press)*, 2001.

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

[9] M. T. Ivrlac, "On capacity of correlated mimo channels," in *internal technical memo, Munich University of Technology, unpublished*, 2001.