SHORT-TERM AND LONG-TERM DIAGONALIZATION OF CORRELATED MIMO CHANNELS WITH ADAPTIVE MODULATION

Mario Kiessling[1,2], Ingo Viering[1], Markus Reinhardt[1], Joachim Speidel[2]

[1] Siemens AG, Lise-Meitner-Strasse 7, 89081 Ulm, Germany
[2] Institute of Telecommunications, University of Stuttgart

Abstract - Theoretical results on MIMO capacity maximization suggest the decoupling of the MIMO channel into independent subchannels with optimum waterfilling on these subchannels. Those results inspire the development of real systems that diagonalize the MIMO channel and adaptively control the modulation on each of the resulting subchannels. In this paper we study the design of a fully adaptive transmitter, where transmit filtering and adaptive modulation is controlled by short-term as well as long-term channel state information (channel correlation), where the focus is on channels with correlated fading at transmitter and receiver array. To this end we are motivating the use of long-term channel Eigen modes by capacity-independent considerations. Simulation results confirm the potential of fully adaptive transmit processing.

Keywords – MIMO; correlation; transmit processing; adaptive modulation

I. INTRODUCTION

Multiple-input multiple-output (MIMO) wireless systems with non-adaptive transmitter experience a significant performance degradation when the fading at transmitter and receiver array is correlated. This is especially true for systems with linear receivers (like minimum mean squared error MMSE). Fading correlation heavily influences the Eigen value (EV) spectrum of the MIMO channel matrix, i.e. with correlation there are on average only a few dominant EVs and many EVs close to zero. Transmitters devoid of any channel state information (CSI) are not capable of adapting to the different fading states of the MIMO channel and thus transmission power is wasted by pouring it into poor Eigen modes associated with the small EVs.

In low mobility TDD systems with the assumption that the transmitter can acquire full short-term (ST) CSI it is possible to considerably improve performance e.g. by antenna selection [1] or linear prefiltering [2]. However, in FDD systems the ST channel state in uplink and downlink is different due to the frequency gap. It is thus very hard if not impossible for the transmitter to get accurate ST CSI. Nevertheless, the transmitter can exploit long-term (LT) CSI (determined by the large-scale scattering scenario), which has been shown to noticeably improve performance ([13]).

Theoretical results on capacity maximization via waterfilling on the MIMO subchannels (cf. [3][4][5][12]) give the motivation to apply waterfilling-like schemes in real systems with discrete constellation sizes and finite codes (see for example [6]). The potential of the wireless MIMO channel can very efficiently be exploited by adaptive power allocation (PA) (the waterfilling analogue) and adaptive modulation (AM) on the subchannels.

In this paper we outline the application of AM to a MIMO wireless system in correlated channels. We are demonstrating the potential of ST as well as LT stochastic processing. An alternative capacity independent motivation for the exploitation of LT channel Eigen modes is given. Simulation results for minimum BER and maximum throughput, respectively, show the effectiveness of the proposed schemes. Especially the LT CSI based scheme is very appealing due to its reduced complexity and its applicability to FDD systems.

II. SIGNAL AND CHANNEL MODEL

We consider a flat fading MIMO link modeled by

$$\mathbf{y} = \mathbf{HFs} + \mathbf{n}\,, \tag{1}$$

where $\mathbf{s}$ is the L×1 TX symbol vector, $\mathbf{F}$ is a M_{TX}×L TX linear matrix filter, $\mathbf{H}$ is the M_{RX}×M_{TX} MIMO channel matrix, $\mathbf{n}$ is the M_{RX}×1 noise vector, $\mathbf{y}$ is the M_{RX}×1 receive vector (see Fig. 1). L is the number of independent subchannels, M_{RX} is the number of RX antennas and M_{TX} is the number of TX antennas. Moreover, $\mathbf{x}$ in Fig. 1 is the vector that is transmitted over the TX antenna array.

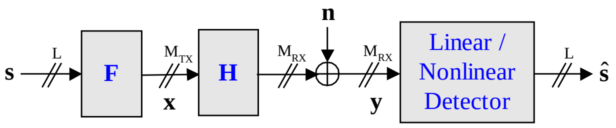

Fig. 1: System model with linear precoding

From the receive vector $\mathbf{y}$, a linear or nonlinear detection process can decouple the symbols on the subchannels and an estimated symbol vector $\hat{\mathbf{s}}$ results. In the following we denote the noise covariance matrix by $\mathbf{R}_{nn}$ and the TX signal covariance is generally given by $\mathbf{R}_{ss}$, whereas in the remainder of the paper we shall assume $\mathbf{R}_{ss}$=$\mathbf{I}_L$ ($\mathbf{I}_L$ denotes the L×L identity matrix). Furthermore, we assume that the

0-7803-7589-0/02/$17.00 ©2002 IEEE

transmit symbol vector **s** is adaptively modulated, i.e. each component of $\mathbf{s}=(s_1, \ldots, s_l, \ldots, s_L)^T$ has a different constellation size Q_l. We will assume that BPSK and square QAM constellations, i.e. $Q_l \in \{2, 4, 16, 64, \ldots\}$ are allowed for each symbol.

With the common assumptions on the MIMO propagation model given e.g. in [4][5][14], the correlated channel can mathematically be expressed as the matrix product

$$\mathbf{H} = \mathbf{A}^H \mathbf{H}_W \mathbf{B}, \tag{2}$$

where $\mathbf{H}_W$ is a $M_{RX} \times M_{TX}$ matrix of complex i. i. d. Gaussian variables of unity variance (i.e. we assume Rayleigh fading) and

$$\mathbf{A}^H \mathbf{A} = \mathbf{R}_{RX} \qquad \mathbf{B}^H \mathbf{B} = \mathbf{R}_{TX}, \tag{3}$$

where $\mathbf{R}_{RX}$ and $\mathbf{R}_{TX}$ is the LT stable receive and transmit correlation matrix, respectively.

III. SHORT-TERM PROCESSING

A. Short-Term Diagonalization of the MIMO Channel

Given full ST CSI, it is possible to perfectly diagonalize the MIMO channel matrix. A detailed view of part of the transmission chain with noise whitening filter $\mathbf{R}_{nn}^{-1/2}$ and matched filter (MF) $\mathbf{F}^H \mathbf{H}^H \mathbf{R}_{nn}^{-H/2}$ is depicted in Fig. 2. It is clear that the components of the vector **z** at the output of the matched filter of the inner system without matrix filter **F**

$$\mathbf{z} = \mathbf{H}^H \mathbf{R}_{nn}^{-1} \mathbf{H} \mathbf{x} + \mathbf{H}^H \mathbf{R}_{nn}^{-1} \mathbf{n} \tag{4}$$

in general contain a linear combination of the components of **x** plus a noise contribution.

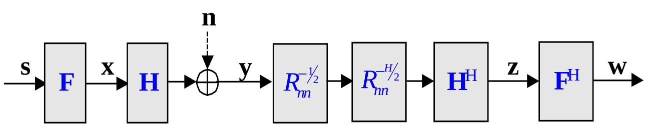

Fig. 2: Tranceiver chain with noise whitening and MF

However, with appropriate filtering at transmitter and receiver the channel matrix of the equivalent system may be diagonalized, i.e. we can achieve that each component of **w**, namely w_l, is just a scaled version of s_l plus noise. This can be shown as follows. With the Eigenvalue decomposition (EVD)

$$\mathbf{H}^H \mathbf{R}_{nn}^{-1} \mathbf{H} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^H \tag{5}$$

we can choose the matrix filter as

$$\mathbf{F} = \mathbf{V} \cdot \mathbf{\Phi}_f \tag{6}$$

with diagonal matrix $\mathbf{\Phi}_f$. We will refer to the Eigen vectors with their corresponding EVs as ST Eigen modes. Then we get

$$\begin{aligned} \mathbf{w} &= \mathbf{F}^H \mathbf{H}^H \mathbf{R}_{nn}^{-1} \mathbf{H} \mathbf{F} \mathbf{s} + \mathbf{F}^H \mathbf{H}^H \mathbf{R}_{nn}^{-1} \mathbf{n} \\ &= \mathbf{\Phi}_f \mathbf{\Lambda} \mathbf{\Phi}_f \mathbf{s} + \tilde{\mathbf{n}} \end{aligned}, \tag{7}$$

where we have introduced a new noise variable $\tilde{\mathbf{n}}$ with covariance matrix

$$\begin{aligned} E\left[\tilde{\mathbf{n}}\tilde{\mathbf{n}}^H\right] &= E\left[\mathbf{F}^H \mathbf{H}^H \mathbf{R}_{nn}^{-1} \mathbf{n}\mathbf{n}^H \mathbf{R}_{nn}^{-1} \mathbf{H}\mathbf{F}\right] \\ &= \mathbf{F}^H \mathbf{H}^H \mathbf{R}_{nn}^{-1} \mathbf{R}_{nn} \mathbf{R}_{nn}^{-1} \mathbf{H}\mathbf{F} \\ &= \mathbf{\Phi}_f \mathbf{\Lambda} \mathbf{\Phi}_f \end{aligned}. \tag{8}$$

Note that the noise covariance matrix is diagonal, i.e. the noise on the parallel subchannels is uncorrelated. The resulting system model is depicted in Fig. 3, where we have omitted the multiplication by $\phi_{f,l}$ on the receiver side and introduced rescaling multipliers on each subchannel $\phi_{g,l}$ to allow for a maximum likelihood detection by simple threshold detectors on the $\hat{s}_l$.

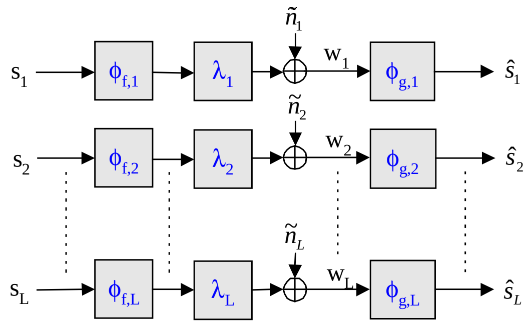

Fig. 3: Diagonalized MIMO model

It is then straightforward to see that the rescaling matrix is

$$\mathbf{\Phi}_g = \left(\mathbf{\Phi}_f \mathbf{\Lambda}\right)^{-1} \tag{9}$$

and the SNR on each subchannel is given by

$$\gamma_l = \frac{E\left[\left(\mathbf{\Phi}_f \mathbf{\Lambda} \mathbf{\Phi}_f\right)^2 \mathbf{s}\mathbf{s}^H\right]_{ll}}{E\left[\tilde{\mathbf{n}}\tilde{\mathbf{n}}^H\right]_{ll}} = \left[\mathbf{\Phi}_f^2 \mathbf{\Lambda}\right]_{ll} = \phi_{f,l}^2 \lambda_l, \tag{10}$$

where $[\mathbf{X}]_{ll}$ is the lth diagonal element of the matrix **X**.

B. Adaptive Modulation on Subchannels

With a total power constraint of ρ at the transmitter, which results in

$$tr\left(\mathbf{F}\mathbf{F}^H\right) = tr\left(\mathbf{\Phi}_f^2\right) \le \rho, \tag{11}$$

we can search for the optimum PA and modulation on each subchannel that maximizes a given criteria like BER or maximum throughput given a BER limit. The detailed structure of the transmitter incorporating AM and power adaptation is depicted in Fig. 4. After calculating **V** and **Λ**, the AM controller has to find the optimum number of subchannels L, the optimum constellation sizes on the subchannels given in the vector $\mathbf{q}=\{Q_1, \ldots, Q_L\}^T$ and the optimum PA matrix $\mathbf{\Phi}_f$. The noise covariance matrix $\mathbf{R}_{nn}$ has to be fed back by the receiver to the transmitter.

In principle, the same AM algorithms that are used for OFDM systems can be applied to the MIMO wireless system. However, the number of subchannels in the MIMO system is normally considerably lower, thus simplifying the optimization process for PA and modulation.

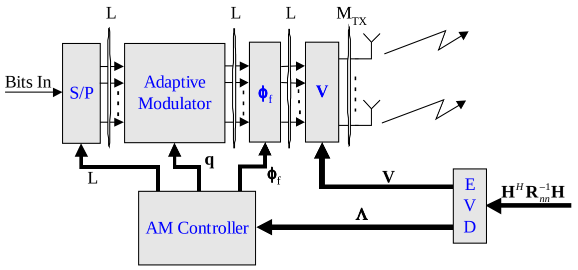

Fig. 4: System model with linear precoding

With M_{TX}=M_{RX}=4 the optimization can be carried out with brute-force. In the simulations presented in this paper, we are basically employing modified algorithms proposed by Huber and Fischer (see [10]) for the constant throughput/minimize BER case and Hughes-Hartogs for the maximum throughput/BER constraint case (see for example [11]).

C. Short-Term Simulation Results

In the simulations of this paper, a new random channel state is determined via (2) for each channel use, while the correlation matrices $\mathbf{R}_{TX}$ and $\mathbf{R}_{RX}$ are held constant for the entire simulation. Both RX and TX arrays have an antenna element spacing of 0.5 wavelengths.

The receiver is assumed to be surrounded by a huge number of scatterers, thus resulting in an angular spread (AS) of 360 degrees with equally distributed power. On the transmit side we assume a single direction of departure (DOD) of 20 degrees with respect to the array perpendicular and a root mean square AS of 10 degrees with a Laplacian power distribution, as it was proposed for the macrocellular vehicular environment in [9]. We will refer to this scenario as a semi-correlated channel (SCC).

The signal-to-noise ratio (SNR) for BER simulations is given as

$$SNR = 10\log_{10}\frac{M_{TX}E_b}{N_0} \quad [dB], \qquad (12)$$

where N_0 is the variance of the complex Gaussian noise on each RX antenna and E_b is the energy per information bit. Throughout our simulations we normalize the total transmitted energy to ρ=M_{TX}, allowing us (in combination with the noise normalization) to compare different TX processing structures. For throughput simulations we equivalently define

$$SNR = 10\log_{10}\frac{M_{TX}E_s}{N_0} = 10\log_{10}\frac{\rho}{N_0} \quad [dB], \qquad (13)$$

where E_s=1 is the symbol energy.

BER simulation results for the SCC with optimum ST PA and AM are depicted in Fig. 5, where 8 bit per channel use are transmitted. The modulation on each subchannel is restricted to maximally 64 QAM and M_{RX}=M_{TX}=4. For comparison, we have also plotted the performance of a system with constant QPSK modulation and a constant number of subchannels L=4 for an uncorrelated as well as the correlated channel with ML detection. Obviously, with the flexible scheme applying adaptive PA and AM we can achieve in a correlated channel the same performance as the non-adaptive scheme in an uncorrelated channel.

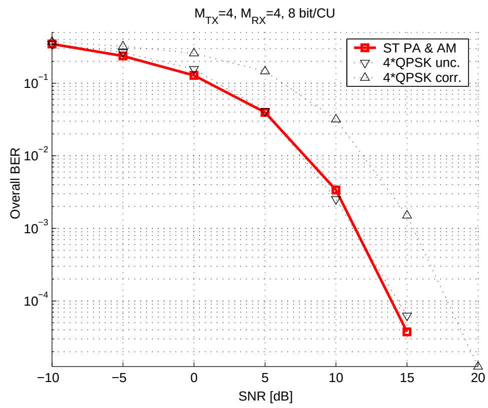

Fig. 5: BER performance of ST PA and AM

In Fig. 6 we have depicted simulation results again for the SCC for maximum throughput simulations with a BER constraint of 10^{-2}, maximum constellation size of 64 QAM and M_{RX}=M_{TX}=4.

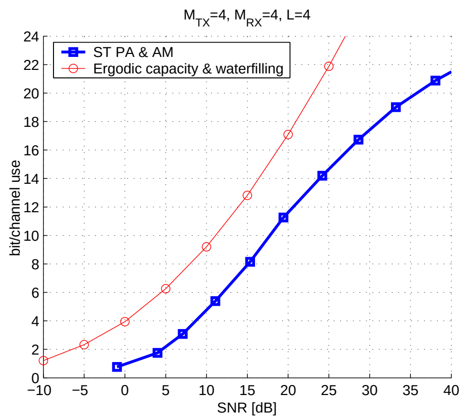

Fig. 6: Data throughput of ST PA and AM (BER 10^{-2})

We have also plotted a curve of the ergodic Shannon capacity with optimum ST waterfilling. It can be seen that there is the expected gap between theoretical capacity and uncoded QAM transmission. The throughput curve saturates at 4·log 64=24 bit per channel use due to the constellation size restriction.

IV. LONG-TERM PROCESSING

A. Long-Term Diagonalization of the MIMO Channel

A transmitter that is not aware of the ST channel state is not capable of instantaneously diagonalizing the MIMO channel via the matrix filter **F**. However, it will be shown in the following that it can beneficially exploit the LT CSI. To this end, note that from (2) and (23) we can derive

$$E\left[\mathbf{H}^H\mathbf{R}_{nn}^{-1}\mathbf{H}\right]=tr\left(\mathbf{R}_{nn}^{-1}\mathbf{R}_{RX}\right)\cdot\mathbf{R}_{TX}\,, \tag{14}$$

i.e. the LT correlation properties of the channel are mainly dominated by the transmit correlation matrix $\mathbf{R}_{TX}$. Following the proceeding in the ST case, we are introducing a LT EVD

$$E\left[\mathbf{H}^H\mathbf{R}_{nn}^{-1}\mathbf{H}\right]=\mathbf{V}_{LT}\mathbf{\Lambda}_{LT}\mathbf{V}_{LT}^H\,. \tag{15}$$

We will refer to the Eigen vectors and the corresponding Eigen values as LT Eigen modes. Clearly, while we can now achieve channel diagonalization in a LT sense by applying the LT Eigenvectors of the channel

$$E\left[\mathbf{V}_{LT}^H\mathbf{H}^H\mathbf{R}_{nn}^{-1}\mathbf{H}\mathbf{V}_{LT}\right]=\mathbf{\Lambda}_{LT} \tag{16}$$

this is in general not true for each snapshot

$$\mathbf{V}_{LT}^H\mathbf{H}^H\mathbf{R}_{nn}^{-1}\mathbf{H}\mathbf{V}_{LT}\neq diagonal\,. \tag{17}$$

However, besides interesting results on ergodic Shannon capacity (cf. [3][4][12]) that suggest transmission on the LT Eigen modes to maximize capacity, we will substantiate this claim in a different manner. To this end, define

$$\mathbf{D}=\mathbf{V}_{LT}^H\mathbf{H}^H\mathbf{R}_{nn}^{-1}\mathbf{H}\mathbf{V}_{LT}-\mathbf{\Lambda}_{LT}\,, \tag{18}$$

which is the deviation of (17) from the LT mean. We will calculate the variance of the elements of **D**, namely

$$d_{r,c}=\underbrace{\left[\mathbf{V}_{LT}^H\mathbf{H}^H\mathbf{R}_{nn}^{-1}\mathbf{H}\mathbf{V}_{LT}\right]_{r,c}}_{u_{r,c}}-\left[\mathbf{\Lambda}_{LT}\right]_{r,c}=u_{r,c}-\left[\mathbf{\Lambda}_{LT}\right]_{r,c} \tag{19}$$

and

$$E\left[d_{r,c}^2\right]=E\left[u_{r,c}^2\right]-\left[\mathbf{\Lambda}_{LT}\right]_{r,c}^2\,. \tag{20}$$

Omitting details, one can find by applying equation (24) in the appendix

$$E\left[d_{r,c}^2\right]=\frac{tr\left(\left[\mathbf{R}_{nn}^{-1}\mathbf{R}_{RX}\right]^2\right)}{tr\left(\mathbf{R}_{nn}^{-1}\mathbf{R}_{RX}\right)^2}\cdot\lambda_r\lambda_c\,, \tag{21}$$

where the λ_l are the LT Eigenvalues of the channel. Writing (21) in matrix notation we have the variance matrix

$$\mathbf{D}_{Var}=\frac{tr\left(\left[\mathbf{R}_{nn}^{-1}\mathbf{R}_{RX}\right]^2\right)}{tr\left(\mathbf{R}_{nn}^{-1}\mathbf{R}_{RX}\right)^2}\begin{pmatrix}\lambda_1\lambda_1 & \lambda_1\lambda_2 & \cdots & \lambda_1\lambda_{M_{TX}}\\ \lambda_1\lambda_2 & \lambda_2\lambda_2 & & \vdots\\ \vdots & & \ddots & \\ \lambda_1\lambda_{M_{TX}} & \cdots & & \lambda_{M_{TX}}\lambda_{M_{TX}}\end{pmatrix}. \tag{22}$$

In a channel with strong fading correlation on the TX side, there will be only a few Eigen values of noticeable size, whereas most of the Eigen values vanish. It becomes then obvious from (22) that the Eigen modes to the corresponding vanishing Eigen values convey almost no power and can thus be discarded. Furthermore, one can observe that there is a considerable variance of the strong Eigen modes. A LT PA scheme needs to take this into account.

Simple ML decisions by threshold detectors on the subchannels are not possible with LT processing at the transmitter. In order to decouple the subchannels, we need to apply standard MIMO receiver techniques like MMSE, ML, successive interference cancellation (SIC) etc. Unfortunately, this fact makes the control of PA and AM more challenging. The subchannels have now an impact on each other, i.e. increasing the power level in one subchannel increases the interference level for the other subchannels. Equivalently, the interference level increases by incrementing the number of subchannels thus leading to a coupled optimization problem. This is especially true for linear receivers like MMSE, where the degrees of freedom of the RX array are used to cancel interference while simultaneously losing diversity.

Calculation of the subchannel BERs in a LT sense in correlated channels with standard receiver architectures is a non-trivial task. The BER formulas should be given in closed form to allow for an easy control of the PA and AM in the transmitter. Random matrix and bounding techniques are currently under investigation to find a solution to this problem (similar problems have appeared in the context of optimum combining [7][8]).

B. Long-Term Simulation Results

In this paper, we simplify the LT PA and AM problem by neglecting the influence of the receiver and the assumption that the subchannel SNRs are solely determined by $\mathbf{\Lambda}_{LT}$. Furthermore, we do not take into account the power variance (fading) of the LT Eigen modes [see (22)], i.e. we are using an additive white Gaussian noise (AWGN) assumption for the calculation of the BER on the subchannels. Given these preliminaries, one can simply apply the ST PA and AM schemes for the LT case just by replacing the ST Eigen vectors and Eigen values by their corresponding LT equivalents.

Simulation results for LT power PA and AM are given in Fig. 7 for MMSE, ML and SIC receivers. For comparison we have also included a curve for optimum ST adaptation.

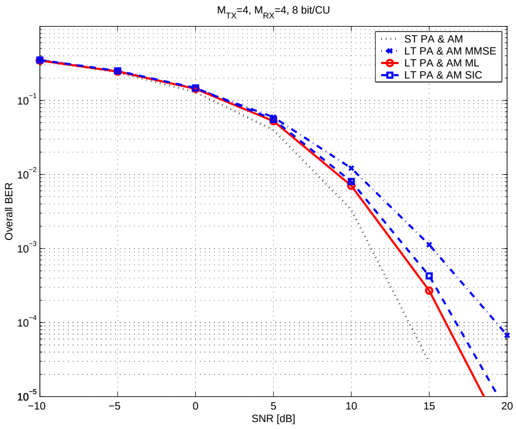

Fig. 7: BER of ST and LT based PA & AM

Despite the simplifications that were applied to find a low complexity LT PA and AM scheme, the performance degradation of the LT based scheme is quite moderate, especially in the range of interest (prior to channel coding) at a BER of 10^{-1} to 10^{-2}.

CONCLUSION

We have studied the application of adaptive modulation with power adaptation for a wireless MIMO system with correlated channels. The diagonalization of the matrix channel was outlined in a short-term and long-term sense. Simulation results have proven the potential of the proposed schemes. Especially long-term based (stochastic) transmit processing seems to be an interesting approach to improve MIMO performance in correlated channels due to its reduced complexity and its applicability in FDD systems. A simpfied long-term power allocation and adaptive modulation scheme was shown to achieve acceptable performance in the BER range of interest.

APPENDIX

For an arbitrary m×m matrix **A** and an m×n matrix **X** with i. i. d. unity variance complex Gaussian distributed entries one can derive by straightforward algebraic evaluation

$$E\left[\mathbf{X}^H\mathbf{A}\mathbf{X}\right]=tr\left(\mathbf{A}\right)\cdot\mathbf{I}_{n\times n}\,. \qquad (23)$$

Equivalently, one can derive the following matrix expectation by explicitly carrying out the matrix multiplication and taking the expectation

$$E\left[\mathbf{G}^H\mathbf{A}\mathbf{G}\mathbf{B}\mathbf{G}^H\mathbf{C}\mathbf{G}\right]=tr\left(\mathbf{A}\right)\cdot tr\left(\mathbf{C}\right)\cdot\mathbf{B}+tr\left(\mathbf{B}\right)\cdot tr\left(\mathbf{AC}\right)\cdot\mathbf{I}\,, \qquad (24)$$

where **G** is a matrix of i. i. d. unitry variance complex Gaussian entries and **A**, **B**, **C** are deterministic matrices.

REFERENCES

[1] Heath R. W., Sandhu S., Paulraj A., “Antenna selection for spatial multiplexing systems with linear receivers”, *IEEE Communications Letters*, vol. 5, no. 4, April 2001

[2] Scaglione A., Giannakis G. B., Barbarossa S., “Redundant filterbank precoders and equalizers part 1: unification and optimal designs”, *IEEE Transactions on Signal Processing*, vol. 47, no. 7, pp. 1988-2006, July 1999

[3] D. Shiu, J. M. Kahn, “Power allocation strategies for wireless systems with multiple transmit antennas”, *http://www.eecs.berkeley.edu/~jmk/res.areas/mawc.html*, July 1998

[4] M. T. Ivrlac, T. P. Kurpjuhn, C. Brunner, W. Utschick, “Efficient use of fading correlations in MIMO systems”, Proc. VTC 2001, Oct. 2001

[5] D. Shiu, G. J. Foschini, M. J. Gans, J. M. Kahn, “Fading correlation and its effect on the capacity of multielement antenna systems”, *IEEE Transactions on Communications*, vol 48, no. 3, pp. 502-513, March 2000

[6] Raleigh G. G., Cioffi J. M., “Spatio-temporal coding for wireless communications”, *IEEE Transactions on Communications*, vol. 46, no. 3, March 1998

[7] H. Gao, P. J. Smith, “Exact SINR calculations for optimum linear combining in wireless systems,” *Probability in the Engineering and Informational Sciences*, vol. 12, pp. 261-281, 1998

[8] Winters J. H., Salz J., Gitlin R. D., ”The impact of antenna diversity on the capacity of wireless communication systems”, *Transactions on Communications*, vol. 42, no. 2/3/4, Feb./March/April 1994

[9] Lucent, Nokia, Siemens, Ericsson, “A standardized set of MIMO radio propagation channels”, TSG-RAN WG 1 #23 R1-01-1179

[10] Fischer R. F. H., Huber J. B., “A new loading algorithm for discrete multitone transmission”, Proc. GLOBECOM 1996, vol. 1, pp. 724-728, May 1996

[11] Bingham J. A. C., “Multicarrier modulation for data transmission: an idea whose time has come”, *IEEE Communications Magazine*, vol. 28, no. 5, pp. 5-14, May 1990

[12] Kiessling M., Viering I., Reinhardt M., Speidel J., “A closed form bound on correlated MIMO channel capacity”, VTC Fall 2002, accepted for publication

[13] Kiessling M., Viering I., Reinhardt M., Speidel J., “Long-term linear prefiltering for MMSE and ML receivers in correlated MIMO channels”, GLOBECOM 2002, submitted

[14] Viering I. Reinhardt M., Frey T., “Statistical modelling of spatially correlated MIMO channels”, IEEE Int. Symposium on Intelligent Signal Processing, 2001