Statistical Prefiltering for MIMO Systems with Linear Receivers in the Presence of Transmit Correlation

Mario Kiessling[1,2], Joachim Speidel[1], Ingo Viering[2], Markus Reinhardt[2]
[1]Institute of Telecommunications, University of Stuttgart, Germany
[2]Siemens AG, Information and Communication Mobile, Ulm, Germany
mario.kiessling@gmx.net, speidel@inue.uni-stuttgart.de

Abstract—**Fading correlation at the transmit antenna array of MIMO systems causes significant performance degradation, especially in case of non-adaptive transmitters and linear receivers. We propose statistical transmit matrix prefiltering schemes for zero-forcing and minimum mean squared error receivers that are controlled by knowledge of the transmit correlation matrix only, thus avoiding the need for instantaneous channel state information at the transmitter. To this end, closed-form approximations are derived for average error probability and average mean squared error, respectively, under a flat Rayleigh fading channel assumption. Based on these results, the optimum linear matrix prefilters are derived for a well-known MIMO channel model. Monte-Carlo simulations demonstrate the effectiveness of the proposed statistical prefilters.**

I. INTRODUCTION

Linear zero-forcing (ZF) and minimum mean squared error (MMSE) MIMO receivers are especially appealing due to their low complexity and are the basic building blocks of more sophisticated successive interference cancellation receivers. However, linear receivers are very sensitive to fading correlation between the antenna elements. Therefore, in typical outdoor scenarios with noticeable correlation between the transmit antennas at the base station one can expect significant performance degradation of the MIMO link.

Theoretical results on MIMO channel capacity [1][11] suggest that exploiting channel state information (CSI) at the transmitter to control an adaptive transmit matrix filter is an adequate means to counteract impairments due to fading correlation at the transmit antenna array. It has been demonstrated that even long-term adaptation of the filter via statistical knowledge of the correlation matrix is very valuable. A major advantage of this concept is that under the assumption of a quasi-stationary propagation scenario, the correlation matrix at the transmitter is long-term stable and can be accurately determined even in FDD systems (e.g. via feedback or frequency transformation between uplink and downlink). These theoretical considerations have already effectively been applied to practical MIMO systems [2][12][13].

In this paper, we outline statistical prefiltering schemes for MIMO systems with ZF and MMSE receivers that beneficially exploit the statistical CSI contained in the transmit correlation matrix in order to improve system performance, thus avoiding the need for accurate instantaneous CSI. To this end, we derive closed-form approximations for the average probability of error and the average mean squared error, respectively, as a function of the transmit correlation matrix and a linear transmit matrix filter, whereas the channel is assumed to be characterized by flat Rayleigh fading. Based on these results, we study the design and optimization of the transmit filter in a statistical sense. It should be emphasized that matrix prefiltering comprises transmit antenna selection [2] as a special case, where the degrees of freedom of the prefilter are restricted. For this particular scenario the prefilter optimization problem simplifies to a discrete problem. We demonstrate the potential of the generalized statistical prefiltering schemes via Monte-Carlo simulations. Depending on the particular propagation conditions, gains of several dB can be achieved.

II. SIGNAL AND CHANNEL MODEL

We consider a flat fading MIMO link modeled by

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

where **s** is the $L\times 1$ TX symbol vector, **F** is a $M_{TX}\times L$ linear matrix transmit prefilter, **H** is the $M_{RX}\times M_{TX}$ MIMO channel matrix with correlated Rayleigh fading elements, **n** is the $M_{RX}\times 1$ noise vector, and **y** is the $M_{RX}\times 1$ receive vector (see Fig. 1). By $M_{RX}\geq L$ we denote the number of RX antennas, M_{TX} is the number of TX antennas and L is the number of independent subchannels. Note that L can in general be smaller than the number of transmit antennas.

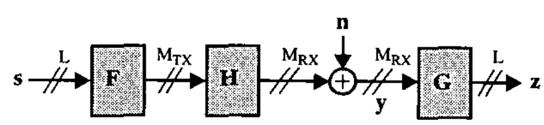

Fig. 1: System model

In the remainder of the paper, by **I** we denote an identity matrix, tr(**X**) is the trace of matrix **X**, diag(**X**) returns a diagonal matrix with the main diagonal taken from **X**, $\mathbf{X}^H$ means Hermitian (conjugate transpose), and $[\mathbf{X}]_{kk}$ is the kth diagonal

0-7803-7757-5/03/$17.00 ©2003 IEEE.

Vehicular Technology Conference, 2003. VTC 2003-Spring. The 57th IEEE Semiannual ,
Volume: 1, April 22-25, 2003, Page(s): 267 –271

element of **X**. Furthermore, $E_x[\ ]$ is the expectation with respect to x and by '≅' we denote 'has the same distribution as'.

We define the linear $L \times M_{RX}$ receive matrix **G** (Fig. 1) and the $L \times 1$ vector **z** that is used for the subsequent symbol detection. In the following we assume additive white Gaussian noise (AWGN), i.e. the noise covariance matrix is given by $\mathbf{R}_{nn}=N_0 \cdot \mathbf{I}$, where N_0 is the noise power. Furthermore, we assume that the symbol energy is normalized such that the covariance matrix of the symbol vector **s** is $\mathbf{R}_{ss}=\mathbf{I}$.

Using a widely accepted channel model, the MIMO channel with TX correlation can be described by the matrix product

$$H = H_w B, \tag{2}$$

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

$$BB^H = B^H B = R_{TX}, \tag{3}$$

whereas we assume that without loss of generality **B** is Hermitian and $\mathbf{R}_{TX}$ is the long-term stable (normalized) transmit correlation matrix, respectively.

III. Prefiltering For Zero-Forcing Receivers

The receiver zero-forcing matrix filter **G** is given by the pseudo-inverse

$$G = (HF)^{\dagger} = (F^H H^H H F)^{-1} F^H H^H, \tag{4}$$

with **GHF=I**. Under the AWGN assumption it is then straightforward to show that the signal-to-noise ratio on subchannel k can be expressed as

$$\gamma_k = \frac{1}{N_0[(HF)^{\dagger}((HF)^{\dagger})^H]_{kk}} = \frac{1}{N_0[(F^H H^H H F)^{-1}]_{kk}}. \tag{5}$$

Using results from [3] and [14] on the probability density function (PDF) of γ_k, namely the chi-squared PDF

$$p(\gamma_k) = \frac{\exp(-\gamma_k/\alpha_k)}{\alpha_k \Gamma(M_{RX}-L+1)} \left(\frac{\gamma_k}{\alpha_k}\right)^{M_{RX}-L} \tag{6}$$

with $\alpha_k = 1/(N_0[(F^H R_{TX} F)^{-1}]_{kk})$, the average (averaged over Rayleigh fading) probability of a symbol error $P_{e,k}$ on subchannel k can be bounded by

$$P_{e,k} \le \beta \cdot (N_0[(F^H R_{TX} F)^{-1}]_{kk})^{M_{RX}-L+1}, \tag{7}$$

where β is a constant depending on the modulation. This is a high SNR approximation based on an exponential bound of the complementary error function.

In order to find the optimum prefilter **F** that minimizes the cumulative error probability of all subchannels, we have to solve the equivalent constrained optimization problem ('s.t.' stands for 'subject to')

$$F_{opt} = \arg \underbrace{\min}_{F} \sum_{k=1}^{L} ([(F^H R_{TX} F)^{-1}]_{kk})^{M_{RX}-L+1} \quad \text{s. t. } \operatorname{tr}(FF^H) = \rho. \tag{8}$$

The transmit power is restricted to ρ. It is possible to solve (8) via Lagrange optimization in closed-form for the special case $L=M_{RX}$, which will be later extended to the general case $L \le M_{RX}$.

Special Case: $L=M_{RX}$

With the Lagrange multiplier λ we get from (8) the equivalent unconstrained optimization problem for complex **F**

$$F_{opt} = \arg \underbrace{\min}_{F,\lambda} [\operatorname{tr}((F^H R_{TX} F)^{-1}) + \lambda \cdot (\operatorname{tr}(FF^H) - \rho)]. \tag{9}$$

Differentiating with respect to **F** (see appendix) and equating to 0 we find

$$F^H R_{TX} = \lambda \cdot (F^H R_{TX} F)^2 F^H. \tag{10}$$

We introduce the eigenvalue decomposition (EVD)

$$R_{TX} = V \Lambda V^H, \tag{11}$$

where the diagonal matrix Λ contains the sorted eigenvalues with decreasing size and decompose the matrix filter **F** without loss of generality as

$$F = V \Phi U^H, \tag{12}$$

where **U** is an arbitrary $L \times L$ unitary matrix and $\mathbf{\Phi}$ is a general matrix of size $M_{TX} \times L$. Note that the particular choice of the matrix **U** has no influence on the problem in (9). Inserting (11) and (12) in (10) we find

$$\Phi \Phi^H \Lambda = \lambda \cdot \Phi \Phi^H \Lambda \Phi \Phi^H \Lambda \Phi \Phi^H. \tag{13}$$

With $F = \tilde{V} \tilde{\Phi} U^H$, where the $M_{TX} \times L$ matrix $\tilde{V}$ contains the eigenvectors corresponding to the L largest eigenvalues and $\tilde{\Phi}$ is of size $L \times L$, this can be reduced to

$$I_L = \lambda \cdot \tilde{\Phi} \tilde{\Phi}^H \tilde{\Lambda} \tilde{\Phi} \tilde{\Phi}^H, \tag{14}$$

where the $L \times L$ diagonal matrix $\tilde{\Lambda}$ contains the L largest eigenvalues, resulting in the solution

$$\tilde{\Phi} \tilde{\Phi}^H = \lambda^{-1/2} \tilde{\Lambda}^{-1/2}. \tag{15}$$

The particular solution for the prefilter **F** follows intuition. Only the strongest L eigenmodes of the channel are used for data transmission, thereby maximizing received signal energy.

Note that the optimization problem in (9) depends only on $\tilde{\Phi} \tilde{\Phi}^H$, whereas the specific structure of $\tilde{\Phi}$ is not relevant. Therefore, without loss of generality we choose a diagonal $\tilde{\Phi}$ with real elements. In this case we find from (15)

$$\tilde{\Phi} = \lambda^{-1/4}\tilde{\Lambda}^{-1/4}. \quad (16)$$

The power constraint finally yields

$$F = \left(\frac{\rho}{tr\left(\tilde{\Lambda}^{-1/2}\right)}\right)^{1/2} \cdot \tilde{V}\tilde{\Lambda}^{-1/4}U^H. \quad (17)$$

General Case: $L \leq M_{RX}$

Using again the results on complex matrix differentiation in the appendix and the general decomposition for the prefilter

$$F = \tilde{V}\tilde{\Phi}, \quad (18)$$

via the Lagrange method with multiplier λ we can find with $N = M_{RX} - L + 1$ the conditional equation for $\tilde{\Phi}$

$$\frac{N}{\lambda} \cdot I = \tilde{\Phi} \cdot \mathrm{diag}((\tilde{\Phi}^H\tilde{\Lambda}\tilde{\Phi})^{-1})^{-(N-1)} \cdot \tilde{\Phi}^H\tilde{\Lambda}\tilde{\Phi}\tilde{\Phi}^H, \quad (19)$$

where we have used the chain rule of complex differentiation.

An obvious solution obeying (19) is again given by a diagonal matrix $\tilde{\Phi}$ [together with (18)]

$$\tilde{\Phi} = \sqrt{\frac{\rho}{\mathrm{tr}\left(\tilde{\Lambda}^{-\frac{N}{(N+1)}}\right)}} \cdot \tilde{\Lambda}^{-\frac{N}{2(N+1)}}. \quad (20)$$

However, it turns out that diagonal $\tilde{\Phi}$ yields only a local minimum of the optimization problem in (9). The authors could not find a general closed-form solution of (19), yet a study of the structure of $\tilde{\Phi}$ and numerical results reveal that the optimum $\tilde{\Phi}$ fulfills

$$\mathrm{diag}((\tilde{\Phi}^H\tilde{\Lambda}\tilde{\Phi})^{-1}) = \kappa \cdot I, \quad (21)$$

where κ is a constant and $\tilde{\Phi}\tilde{\Phi}^H$ is diagonal. We emphasize that simulation results indicate that the closed-form local optimum corresponding to diagonal $\tilde{\Phi}$ is very close to the global optimum, thus motivating its application.

IV. Prefiltering For MMSE Receivers

A natural choice of the optimization criterion in case of the MMSE receiver is the mean squared error (MSE) averaged over Rayleigh fading. To this end, we first give analytical expressions for the average MSE that can be optimized via the linear prefilter. Equal SER is then forced on all subchannels via DFT preprocessing at the transmitter, resulting in minimum overall SER [4][13].

In case of the MMSE receiver, the receive matrix filter **G** is given by [5]

$$G = (F^H H^H H F + N_0 \cdot I)^{-1} F^H H^H, \quad (22)$$

The resulting MSE summed over all subchannels with general decomposition $F = \tilde{V}\tilde{\Phi}U^H$ and (2) reads

$$\varepsilon = N_0 \cdot \mathrm{tr}((\tilde{\Phi}^H\tilde{\Lambda}^{1/2}H_w^H H_w\tilde{\Lambda}^{1/2}\tilde{\Phi} + N_0 \cdot I)^{-1}). \quad (23)$$

We note that the MSE ε is independent of the unitary matrix **U**. As was mentioned above, we can choose **U** to be a DFT matrix to minimize the resulting overall SER (see [4]). Introducing the singular value decomposition

$$\Lambda^{1/2}\tilde{\Phi} = WDZ^H, \quad (24)$$

it is straightforward to see that we can preserve the statistical properties (pre- and post-multiplication by a unitary matrix does not change the statistics of $\mathbf{H}_w$) by writing

$$\varepsilon \cong N_0 \cdot \mathrm{tr}((D\tilde{H}_w^H\tilde{H}_w D + N_0 \cdot I)^{-1}) = \sum_{k=1}^{L} \varepsilon_k, \quad (25)$$

where $\tilde{H}_w$ is a matrix of i.i.d. complex Gaussian variables of size $M_{RX} \times L$. The MSE ε_k on subchannel k is

$$\varepsilon_k = N_0[(D\tilde{H}_w^H\tilde{H}_w D + N_0 \cdot I)^{-1}]_{kk}. \quad (26)$$

The MSE averaged over fast Rayleigh fading, namely ξ, is then given by the expectation with respect to $\tilde{H}_w$

$$\xi = \mathrm{E}_{\tilde{H}_w}[\varepsilon] = \sum_{k=1}^{L} \mathrm{E}_{\tilde{H}_w}[\varepsilon_k] = \sum_{k=1}^{L} \xi_k. \quad (27)$$

For the equivalent system described by equation (25) and (26) it can be shown that the SNR γ_k on subchannel k is given by

$$\gamma_k = \frac{1}{\varepsilon_k} - 1 = \frac{1}{N_0[(D\tilde{H}_w^H\tilde{H}_w D + N_0 \cdot I)^{-1}]_{kk}} - 1, \quad (28)$$

which can be reformulated by exploiting the matrix inversion lemma and results on partitioned inverses to yield

$$\begin{aligned} \gamma_k &= d_k^2 \cdot u^H(\hat{H}_w\tilde{D}^2\hat{H}_w^H + N_0 \cdot I)^{-1}u \\ &= u^H\left(\hat{H}_w\Gamma\hat{H}_w^H + \frac{1}{\gamma} \cdot I\right)^{-1}u \end{aligned}, \quad (29)$$

where d_k is the kth diagonal element of **D**, **u** is the kth column of $\tilde{H}_w$, $\hat{H}_w$ results from deleting the kth column from $\tilde{H}_w$, $\tilde{D}$ is the diagonal matrix resulting from deleting the kth diagonal element from **D**, finally $\Gamma = d_k^{-2}\tilde{D}^2$ and $\gamma = d_k^2/N_0$. The reliability function $R(\gamma_k)$ of the random variable γ_k was given in [6] in the context of optimum combining

$$R(\gamma_k) = e^{-\frac{\gamma_k}{\gamma}} \cdot \left(\sum_{i=1}^{M_{RX}} \bar{\beta}_i \cdot \gamma_k^{i-1}\right) \Big/ \left(\prod_{j=1}^{L-1} (1 + \Gamma_j\gamma_k)\right), \quad (30)$$

where Γ_j is the jth diagonal element of Γ and the $\bar{\beta}_i$ are the coefficients of z^i in the series expansion of

$$\exp(z/\gamma)\cdot\prod_{j=1}^{L-1}(1+\Gamma_j z). \tag{31}$$

We can directly calculate the MSE on subchannel k averaged over fast Rayleigh fading via integration by parts

$$\xi_k = 1-\int_0^{\infty}\frac{1}{(1+\gamma_k)^2}R(\gamma_k)d\gamma_k. \tag{32}$$

Due to the space limitation, we omit details of the exact calculation of the ξ_k. However, we note that the overall MSE is a (very complicated) function of the prefilter matrix, namely $\tilde{\Phi}$, that can be optimized as in case of the ZF receiver, i.e.

$$\tilde{\Phi}_{opt} = \arg\underbrace{\min}_{\tilde{\Phi}}\xi \quad \text{s. t. } \operatorname{tr}(\tilde{\Phi}\tilde{\Phi}^H)=\rho \tag{33}$$

Numerical results indicate that the prefilter can without loss of generality be assumed to be real diagonal, i.e. $D = \tilde{\Lambda}^{1/2}\tilde{\Phi}$ is diagonal. This conjecture is substantiated by the fact that for large M_{RX} the statistical matrix $H_w^H H_w$ in (23) tends to a scaled identity matrix and ε becomes deterministic. For this particular structure of ε it can shown that the optimizing $\bar{\Phi}$ is diagonal.

V. Simulation Results

In the following simulations we assume uniform linear arrays with 0.5 lambda element spacing at both transmitter and receiver, one main direction of departure (DOD) at 20 degrees with respect to the array perpendicular, and a Laplacian power distribution with 10 degrees angular spread (AS). The fading at the receiver array shall be uncorrelated.

We study an AWGN (i.e. no colored interference) scenario with the SNR given by

$$SNR = 10\cdot\log_{10}\left(\frac{M_{TX}\cdot E_b}{N_0}\right) \qquad [dB], \tag{34}$$

where E_b is the energy per information bit. Throughout our simulations we normalize the total transmitted energy to ρ=M_{TX} and assume QPSK modulation.

We have plotted the overall BER curves for a 4×4 system with L=4 independent subchannels and ZF receiver with and without prefiltering (blind transmission) in Fig. 2 for the correlated scenario. For comparison, we have also included the uncorrelated case.

One can observe a 13.5 dB loss in SNR (high SNR region) due to the transmit correlation with blind transmission. The performance degradation is severe, as all possible subchannels of the MIMO channel are utilized. In this case, the performance of the system is dominated by eigenmodes corresponding to weak eigenvalues of the correlated channel. However, note that an asymptotic 2.8 dB improvement in SNR can be achieved by the proposed statistical prefiltering design in (17).

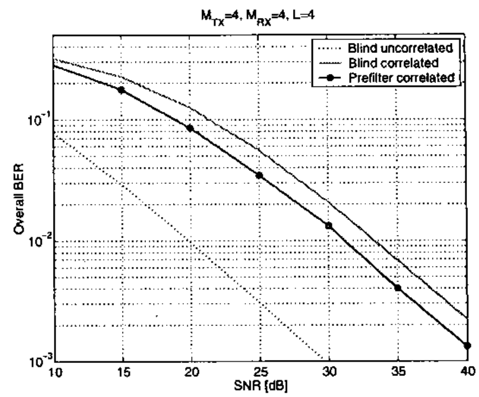

Fig. 2: ZF prefiltering (L=M_{RX}=M_{TX}=4, 10° AS)

An example of a 4×4 system, where only a subset L=2 of the available subchannels is used, is given in Fig. 2.

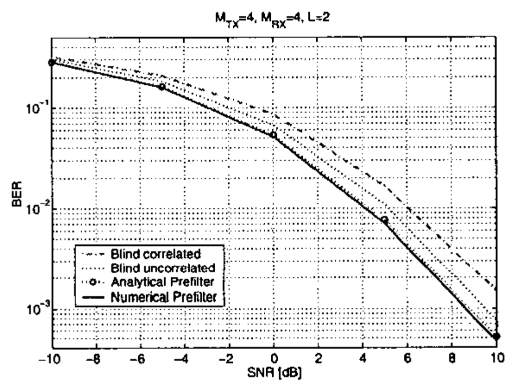

Fig. 3: ZF prefiltering (L=2, M_{RX}=M_{TX}=4, 10° AS)

Note that in the blind transmission case we are using only the outer two transmit antennas. The performance degradation due to transmit correlation in the blind case is less pronounced, as the influence of the weak channel eigenmodes is now less severe. Interestingly, with matrix prefiltering and transmit correlation, the performance of the uncorrelated system can be outperformed. The matrix prefilter effectively uses the beamforming-like capabilities of the transmit antenna array. We have included results for the analytical prefilter according to (20) and the numerically designed prefilter that globally optimizes (8). Obviously, only a minor difference in performance between the two designs can be observed.

Simulation results for a 4×4 system with MMSE receiver with prefiltering including DFT preprocessing and on the other hand without prefiltering are depicted in Fig. 4, again for an AS of 10 degrees at the transmitter and full usage of all eigenmodes L=4.

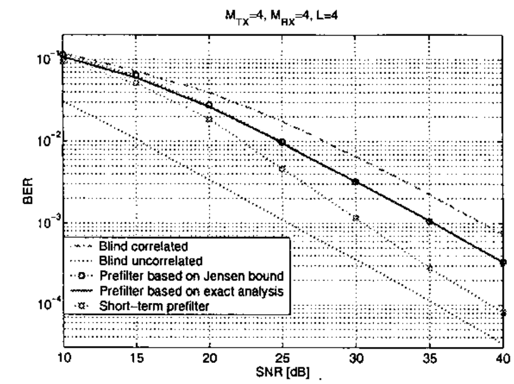

Fig. 4: MMSE prefiltering ($L=M_{RX}=M_{TX}=4$, 10° AS)

Similar to the ZF case, transmit correlation leads to a loss in SNR of more than 13 dB, while the prefiltering gain achieved by the transmit matrix filter is more than 3 dB for higher SNR. We have also included a curve for a system with optimum short-term MMSE prefilter (see e.g. [8]) requiring instantaneous CSI at the transmitter. Furthermore, we note that the low-complexity prefiltering scheme based on Jensen's bound of the MSE, which was proposed in [13], achieves a very similar performance to the prefilter based on exact analysis. The curves of the two systems are almost indistinguishable.

Similar results can be obtained for a system with only L=2 subchannels (Fig. 5). The general behavior resembles the ZF case in Fig. 3. We emphasize that the curve of the low complexity Jensen's bound based prefiltering scheme coincides with the graph based on exact analysis.

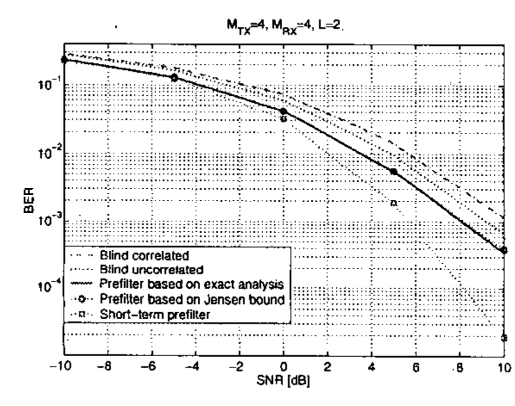

Fig. 5: MMSE prefiltering ($M_{RX}=M_{TX}=4$, L=2, 10° AS)

VI. Appendix

We are using the formal definition of the derivative with respect to the complex variable z=x+jy (see e.g. [4][9][10])

$$\frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - j\frac{\partial}{\partial y}\right) \quad (35)$$

with the formal results

$$\frac{\partial}{\partial z}(z) = 1 \qquad \frac{\partial}{\partial z}(z^*) = 0. \quad (36)$$

One can then derive for a matrix **Y** (see e.g. [7] for the real case)

$$\frac{\partial}{\partial z}(Y^{-1}) = -Y^{-1} \cdot \frac{\partial}{\partial z}(Y) \cdot Y^{-1}. \quad (37)$$

Using equation (37) and the linearity of the trace operator, we can derive

$$\frac{\partial}{\partial X}\mathrm{tr}((X^H R X)^{-1}) = -((X^H R X)^{-1}(X^H R X)^{-1} \cdot X^H R)^T. \quad (38)$$

Equivalently, it is possible to derive for complex **X**

$$\frac{\partial}{\partial X}\mathrm{tr}(X^H X) = X^*. \quad (39)$$

References

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

[2] D. Gore, R. W. Heath Jr., A. Paulraj, "Statistical antenna selection for spatial multiplexing systems", ICC, May 2002

[3] D. Gore, R. W. Heath Jr., A. Paulraj, "On performance of the zero forcing receiver in presence of transmit correlation", ISIT, June 2002

[4] Thomas T. A., Wok F. W., "MIMO strategies for equal-rate data streams", VTC, pp. 548-552, 2001

[5] Kay S. M., *Fundamentals of statistical signal processing volume 1 (estimation theory)*, Prentice Hall, 1993

[6] H. Gao, P. J. Smith, M.V. Clark, "Theoretical reliability of MMSE linear diversity combining in Rayleigh-fading additive interference channels", *IEEE Transactions on Communications*, vol. 46, no. 5, pp. 666-672, May 1998

[7] Magnus J. R., Neudecker H., *Matrix differential calculus with applications in statistics and econometrics*, John Wiley & Sons, 1998

[8] Sampath H., "Linear precoding and decoding for multiple input multiple output (MIMO) wireless channels", PhD thesis, August 2001

[9] Haykin S., *Adaptive filter theory*, Prentice Hall, 1996

[10] Brandwood D. H., "A complex gradient operator and its application in adaptive array theory", *IEE Proceedings Parts F and H*, vol. 130, no. 1, February 1983

[11] Kiessling M., Viering I., Reinhardt M., Speidel J., "A closed form bound on correlated MIMO channel capacity", VTC, Sept. 2002

[12] Kiessling M., Viering I., Reinhardt M., Speidel J., "Short-term and long-term diagonalization of correlated MIMO channels with adaptive modulation", PIMRC, Sept. 2002

[13] Kiessling M., Viering I., Reinhardt M., Speidel J., "Long-term linear prefiltering for MMSE and ML receivers in correlated MIMO channels", WCNC, March 2003

[14] Kiessling M., Speidel J., "Analytical performance of MIMO zero-forcing receivers in correlated Rayleigh fading environments", SPAWC, June 2003, submitted