STATISTICAL PREFILTER DESIGN FOR MIMO ZF AND MMSE RECEIVERS BASED ON MAJORIZATION THEORY

Mario Kiessling[1,2], Joachim Speidel[1]

[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

Recently, the authors have proposed statistical prefilters for MMSE and ZF receivers that minimize SER. In this paper, we give a general derivation of their structure based on majorization theory. For both receiver types it is shown that the optimal prefilter essentially transmits on the strongest long-term eigenmodes of the channel with proper power allocation. Moreover, simple closed-form power allocation schemes are presented for Rayleigh and Ricean fading environments. Interestingly, while the statistical prefilters exhibit the same basic mathematical structure as their short-term counterparts, they require only statistical information of the correlation properties and the Ricean component of the channel. Monte-Carlo simulations show that the proposed filters can achieve a considerable performance gain. Specifically, it is demonstrated that they can completely counteract the SER degradation due to a Ricean channel component.

1. INTRODUCTION

With the availability of short-term (ST) channel state information (CSI) at the transmitter, it is well-known that the symbol error rate (SER) of wireless MIMO systems can be reduced by adequate linear prefilters. Filter designs are available for zero-forcing (ZF) [1] and minimum mean squared error (MMSE) [2][3][4] receivers. However, in many cases it is impossible to acquire ST CSI at the transmitter, e.g. with high user mobility or in frequency division duplex (FDD) systems. Therefore, in [8][9] the authors have proposed long-term (LT) statistical prefilters that are based on the correlation properties of Rayleigh fading MIMO channels only. In this paper, we extend these results by providing exact proofs of the optimal prefilter structures based on majorization theory [5]. It is shown that in both the ZF and MMSE case, the prefilter transmits independent data streams along the strongest LT eigenmodes of the channel with a suitable power allocation (PA). On the other hand, the optimum PA strategy in general is a complicated function of the prevailing fading statistics of the wireless channel and becomes hardly analytically tractable. One possible remedy to this problem is a costly numerical Monte-Carlo optimization. We avoid this by introducing low-complexity closed-form power allocation policies for Rayleigh fading environments based on bounding techniques, which are also effectively deployed after simple manipulations in the presence of Ricean fading. Interestingly, there is a duality between the prefilter structures in the ST and LT case, i.e. it turns out that the LT prefilters can be derived by simply replacing the ST eigenvectors and eigenvalues of the channel by their LT equivalents. Monte-Carlo simulations demonstrate the effectiveness of the proposed prefilter schemes and a significant improvement in the symbol error rate (SER).

2. SIGNAL AND CHANNEL MODEL

We consider a flat fading MIMO link modeled by

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

where **s** is the L×1 TX symbol vector, **F** is a M_{TX}×L linear matrix transmit prefilter, **H** is the M_{RX}×M_{TX} MIMO channel matrix with correlated fading elements, **n** is the M_{RX}×1 noise vector, and **y** is the M_{RX}×1 receive vector. 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. In the following we assume additive Gaussian noise with covariance matrix $\mathbf{R}_{nn}$.
Using a widely accepted channel model, the Rayleigh fading MIMO channel with TX correlation can be described by the matrix product

$$\boldsymbol{H}_{Ray} = \boldsymbol{H}_w \boldsymbol{B}, \tag{2}$$

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

$$\boldsymbol{BB}^H = \boldsymbol{B}^H\boldsymbol{B} = \boldsymbol{R}_{TX} = \frac{1}{M_{RX}} \cdot E_{\boldsymbol{H}_w}[\boldsymbol{H}_{Ray}^H \boldsymbol{H}_{Ray}], \tag{3}$$

whereas we assume without loss of generality that **B** is Hermitian and $\mathbf{R}_{TX}$ is the long-term stable (normalized) transmit correlation matrix, respectively. E_x denotes expectation with respect to x.
In the presence of a deterministic component, i.e. Ricean fading, the channel model in (2) can be extended to (see e.g. [6])

$$\boldsymbol{H}_{Rice} = \sqrt{\frac{1}{1+K}} \cdot \boldsymbol{H}_w \boldsymbol{B} + \sqrt{\frac{K}{1+K}} \cdot \boldsymbol{M}, \tag{4}$$

where **M** is the non-fading component, normalized such that $\mathrm{tr}(\boldsymbol{M}^H\boldsymbol{M}) = M_{RX} \cdot M_{TX}$, and K is the Ricean factor ranging from 0 (purely stochastic) to infinity (purely deterministic).

3. ZF STATISTICAL PREFILTER DESIGN

Using majorization theory, we calculate a general solution of the prefilter optimization problem for Rayleigh fading channels given in [9]. Via a simple moment fit the result is then extended to Ricean fading.

3.1. Rayleigh fading

By considering the Chernoff bound (valid for the high SNR region) on the symbol error rate (SER), in [9] the authors have derived that the SER optimizing linear prefilter matrix **F** for a ZF receiver with transmit correlated Rayleigh fading channel and additive white Gaussian noise (AWGN) can be obtained by solving the following constrained optimization problem

$$\boldsymbol{F}_{opt} = \arg\min_{\boldsymbol{F}} \sum_{k=1}^{L} ([(\boldsymbol{F}^H\boldsymbol{R}_{TX}\boldsymbol{F})^{-1}]_{kk})^N \text{ s.t. } \mathrm{tr}(\boldsymbol{F}\boldsymbol{F}^H) = \rho \quad (5)$$

with the diversity parameter $N = M_{RX} - L + 1$ and transmit power restriction ρ. Using a direct differentiation approach, in [9] a general solution for arbitrary system parameters could not be found. This gap is now closed by

Theorem 1. The optimum SER minimizing prefilter with ZF receiver in the sense of (5) is given by

$$\boldsymbol{F}_{opt} = \tilde{\boldsymbol{V}}\Phi\boldsymbol{D}_L, \quad (6)$$

where $\mathbf{D}_L$ is a L×L discrete Fourier transform (DFT) matrix, $\tilde{\boldsymbol{V}}$ is a matrix of the eigenvectors corresponding to the L strongest eigenvalues of the eigenvalue decomposition (EVD) in

$$\boldsymbol{R}_{TX} = \begin{bmatrix}\boldsymbol{V} & \tilde{\boldsymbol{V}}\end{bmatrix}\begin{bmatrix}\Lambda & \\ & \tilde{\Lambda}\end{bmatrix}\begin{bmatrix}\boldsymbol{V} & \tilde{\boldsymbol{V}}\end{bmatrix}^H, \quad (7)$$

where the matrix $\tilde{\Lambda}$ contains the L largest eigenvalues in increasing order. The diagonal power allocation matrix Φ reads

$$\boldsymbol{\Phi} = \left[\rho/\mathrm{tr}\left(\tilde{\Lambda}^{-1/2}\right)\right]^{1/2} \cdot \tilde{\Lambda}^{-1/4}. \quad (8)$$

Proof: In this paper, we use majorization theory for solving problem (5). To this end, we first introduce the auxiliary L×L matrix

$$\boldsymbol{X} = (\boldsymbol{F}^H\boldsymbol{R}_{TX}\boldsymbol{F})^{-1} \quad (9)$$

with the vector of real diagonal elements

$$\boldsymbol{x} = \mathrm{diag}(\boldsymbol{X}) = \begin{bmatrix}x_1(\boldsymbol{F}) & x_2(\boldsymbol{F}) & \dots & x_L(\boldsymbol{F})\end{bmatrix}. \quad (10)$$

Without loss of generality, assume that the elements of **x** are arranged in decreasing order. Now we can reformulate the problem in (5) as

$$\boldsymbol{F}_{opt} = \arg\min_{\boldsymbol{F}} \sum_{k=1}^{L} [x_k(\boldsymbol{F})]^N \text{ s.t. } \mathrm{tr}(\boldsymbol{F}\boldsymbol{F}^H) = \rho. \quad (11)$$

Note that $\boldsymbol{F}^H\boldsymbol{R}_{TX}\boldsymbol{F}$ is a positive definite matrix, i.e. all minors are positive definite and thus the diagonal of the inverse is positive. Now consider the objective function in (11). Obviously, it is symmetric in its arguments. Furthermore, it is convex in each of its arguments. It follows from Theorem 4 in the appendix that it is a Schur-convex function. Using Theorem 3, the objective function is minimized for equal elements in **x**. From Theorem 6, we can find a real symmetric matrix **Q** such that $\mathbf{Q}^H\mathbf{X}\mathbf{Q}$ has identical diagonal elements, i.e. we use prefilter matrix

$$\boldsymbol{F} = \tilde{\boldsymbol{F}}\boldsymbol{Q}. \quad (12)$$

Note that this does not change the transmit power. Another straightforward choice for the matrix **Q** is a DFT matrix of size L×L. Using the transmit filter structure in (12), we find

$$[(\boldsymbol{F}^H\boldsymbol{R}_{TX}\boldsymbol{F})^{-1}]_{kk} = \frac{1}{L}\mathrm{tr}((\boldsymbol{F}^H\boldsymbol{R}_{TX}\boldsymbol{F})^{-1}). \quad (13)$$

Obviously, with (13) the optimization problem in (5) can be reduced to (minimum trace minimizes each diagonal element)

$$\tilde{\boldsymbol{F}}_{opt} = \arg\min_{\tilde{\boldsymbol{F}}} \mathrm{tr}((\tilde{\boldsymbol{F}}^H\boldsymbol{R}_{TX}\tilde{\boldsymbol{F}})^{-1}) \text{ s.t. } \mathrm{tr}(\tilde{\boldsymbol{F}}\tilde{\boldsymbol{F}}^H) = \rho. \quad (14)$$

Now (14) is a Schur-concave function of the diagonal elements of $(\tilde{\boldsymbol{F}}^H\boldsymbol{R}_{TX}\tilde{\boldsymbol{F}})^{-1}$. Applying Theorem 5 we can choose **F** by proper application of a rotation matrix such that $(\tilde{\boldsymbol{F}}^H\boldsymbol{R}_{TX}\tilde{\boldsymbol{F}})^{-1}$ is diagonal with elements in decreasing order or equivalently such that the diagonal elements of $\tilde{\boldsymbol{F}}^H\boldsymbol{R}_{TX}\tilde{\boldsymbol{F}}$ are in increasing order. The diagonalizing rotation matrix consists of L eigenvectors of $\mathbf{R}_{TX}$.
It can then be shown by Lemma 2 that we have to chose the eigenvectors corresponding to the largest L eigenvalues of $\mathbf{R}_{TX}$, such that the optimum **F** has the structure in (6). Constrained Lagrange optimization of the problem in (14) then leads to (see also [9]) the power allocation matrix in (8). *QED.*

3.2. Ricean fading

The SER performance analysis that led to the optimum prefilter design problem in (5) becomes extremely complicated in the case of Ricean fading and to the author's best knowledge there are no solutions available in literature. Therefore, we propose an approximation of the Ricean fading statistics (essentially, we approximate a non-central Wishart distribution via a central Wishart distribution [7]) by a moment fit

$$E_{H_w}[\tilde{\boldsymbol{H}}_{Ray}^H\tilde{\boldsymbol{H}}_{Ray}] \overset{!}{=} E_{H_w}[\boldsymbol{H}_{Rice}^H\boldsymbol{H}_{Rice}], \quad (15)$$

where $\tilde{\boldsymbol{H}}_{Ray}$ is the Rayleigh fading channel matrix approximating the Ricean statistics. Calculating the expected values in (15) and equating results in an approximating transmit covariance of $\tilde{\boldsymbol{H}}_{Ray}$

$$\tilde{\boldsymbol{R}}_{TX} = \frac{1}{1+K}\cdot\boldsymbol{R}_{TX} + \frac{1}{M_{RX}}\cdot\frac{K}{1+K}\cdot\boldsymbol{M}^H\boldsymbol{M}. \quad (16)$$

Using the approximation in (16), for the prefilter design in Ricean fading we can just use equations (6)(7)(8), with $\mathbf{R}_{TX}$ replaced by $\tilde{\boldsymbol{R}}_{TX}$. We note that the proposed scheme is applicable to arbitrary MIMO channel statistics.

4. MMSE STATISTICAL PREFILTER DESIGN

4.1. Rayleigh fading

The resulting average mean squared error (MSE) summed over all subchannels with the signal and channel models in (1) and (2) and MMSE receiver reads [8]

$$\varepsilon(\boldsymbol{F}) = E_{\boldsymbol{H}_w}[\mathrm{tr}((\boldsymbol{F}^H\boldsymbol{B}^H\boldsymbol{H}_w^H\boldsymbol{R}_{nn}^{-1}\boldsymbol{H}_w\boldsymbol{B}\boldsymbol{F} + \boldsymbol{I})^{-1})]. \quad (17)$$

Now we have the optimization problem for the prefilter (such that it minimizes the average MSE)

$$\boldsymbol{F}_{opt} = \arg\min_{\boldsymbol{F}} \varepsilon(\boldsymbol{F}) \text{ s.t. } \mathrm{tr}(\boldsymbol{F}\boldsymbol{F}^H) = \rho. \quad (18)$$

The solution is given in

Theorem 2. The optimum prefilter with MMSE receiver in the sense that it minimizes the average MSE and forces equal SER on each subchannel in a Rayleigh fading environment is given by

$$\boldsymbol{F}_{opt} = \tilde{\boldsymbol{V}}\boldsymbol{\Phi}\boldsymbol{D}_L, \quad (19)$$

with diagonal PA matrix Φ, L×L DFT matrix $\mathbf{D}_L$ and the EVD in (7).

Proof: Introducing the singular value decomposition (SVD)

$$\boldsymbol{BF} = \boldsymbol{YDZ}^H, \tag{20}$$

where $\mathbf{Y}$ contains the left singular vectors, $\mathbf{Z}$ the right singular vectors and $\mathbf{D}$ the singular values we get

$$\varepsilon(\boldsymbol{D}) = E_{\boldsymbol{H}_w}[\mathrm{tr}((\boldsymbol{D}\tilde{\boldsymbol{H}}_w^H\boldsymbol{R}_{nn}^{-1}\tilde{\boldsymbol{H}}_w\boldsymbol{D}+\boldsymbol{I})^{-1})] \tag{21}$$

with a M_{RX}×L matrix of complex Gaussian i.i.d. entries $\tilde{\boldsymbol{H}}_w$. Obviously, the average MSE is a function of the singular values in $\mathbf{D}$ only. Now noting that left multiplication of $\mathbf{F}$ with unitary $\mathbf{Z}$ does not change the objective function and the constraint in (18), we can introduce

$$\tilde{\boldsymbol{F}} = \boldsymbol{FZ} \Leftrightarrow \tilde{\boldsymbol{F}}\boldsymbol{Z}^H = \boldsymbol{F} \tag{22}$$

and solve problem (18) with $\mathbf{F}$ being replaced by $\tilde{\boldsymbol{F}}$. We can arbitrarily chose the unitary matrix $\mathbf{Z}$. Here, we let $\mathbf{Z}^H$=$\mathbf{D}_L$ be a DFT matrix for minimizing the overall SER [3]. We find from (20) and (22)

$$\tilde{\boldsymbol{F}}^H\boldsymbol{B}^H\boldsymbol{B}\tilde{\boldsymbol{F}} = \tilde{\boldsymbol{F}}^H\boldsymbol{R}_{TX}\tilde{\boldsymbol{F}} = \boldsymbol{D}^2. \tag{23}$$

Finally, we can apply Lemma 2 and arrive at (19). *QED.*

Theorem 2 describes only the general structure of the prefilter and we still have to define the PA matrix Φ. To this end, let

$$E[\boldsymbol{H}_{Ray}^H\boldsymbol{R}_{nn}^{-1}\boldsymbol{H}_{Ray}] = \mathrm{tr}(\boldsymbol{R}_{nn}^{-1})\cdot\boldsymbol{R}_{TX} = \begin{bmatrix}\boldsymbol{V}_h & \tilde{\boldsymbol{V}}_h\end{bmatrix}\begin{bmatrix}\boldsymbol{\Lambda}_h & \\ & \tilde{\boldsymbol{\Lambda}}_h\end{bmatrix}\begin{bmatrix}\boldsymbol{V}_h & \tilde{\boldsymbol{V}}_h\end{bmatrix}^H. \tag{24}$$

We note that in (24) obviously $\begin{bmatrix}\boldsymbol{V}_h & \tilde{\boldsymbol{V}}_h\end{bmatrix} = \begin{bmatrix}\boldsymbol{V} & \tilde{\boldsymbol{V}}\end{bmatrix}$ defined in (7). Via a bound on the average MSE and Lagrange optimization, in [8] a low complexity long-term power allocation (PA) policy was proposed with

$$\boldsymbol{\Phi} = (\mu^{-1/2}\tilde{\boldsymbol{\Lambda}}_h^{-1/2} - \tilde{\boldsymbol{\Lambda}}_h^{-1})_+^{1/2}, \tag{25}$$

where the constant μ is chosen according to the power constraint, resulting in

$$\mu^{1/2} = \frac{\mathrm{tr}(\tilde{\boldsymbol{\Lambda}}_h^{-1/2})}{\mathrm{tr}(\tilde{\boldsymbol{\Lambda}}_h^{-1})+\rho}. \tag{26}$$

We have to assure ϕ_l>0 for all l, which is indicated by the plus sign in (25). This means that in certain situations the weakest LT eigenmodes are not used for transmission by setting the corresponding PA coefficient to 0.

4.2. Ricean fading

Similar to (24), we now calculate for the Ricean channel

$$E[\boldsymbol{H}_{Rice}^H\boldsymbol{R}_{nn}^{-1}\boldsymbol{H}_{Rice}] = \frac{\mathrm{tr}(\boldsymbol{R}_{nn}^{-1})}{1+K}\cdot\boldsymbol{R}_{TX} + \frac{K}{(1+K)}\cdot\boldsymbol{M}^H\boldsymbol{R}_{nn}^{-1}\boldsymbol{M}. \tag{27}$$

With the corresponding EVD of (27), we can use the prefilter structure in (19) together with the power allocation (25) also for a Ricean fading environment. We note that the resulting prefilter is optimum (i.e. it is no approximation) in the sense that it minimizes the bound on the average MSE given in [8] based on Jensen's inequality.

5. SIMULATION RESULTS

In the following simulations, we assume a uniform linear array with 0.5 lambda element spacing at the transmitter and receiver. In the presence of transmit fading correlation, for the Rayleigh fading scattering component of the channel we consider the case of 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 transmit correlation matrix is determined according to these assumptions via Monte-Carlo simulation, while uncorrelated fading is assumed at the receiver.

On the other hand, we consider the case of a rank-1 deterministic channel component $\mathbf{M}$ corresponding to a single line-of-sight (LOS) component, whereas we assume a DOD at the transmitter and a DOA at the receiver of 45 degrees.

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

$$SNR = 10\cdot\log_{10}\left(\frac{M_{TX}\cdot E_b}{N_0}\right) \qquad [dB], \tag{28}$$

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

In Fig. 1 we show BER results of a ZF receiver in a uncorrelated Ricean fading channel with K=1 (equal split between deterministic and scattering component) for a 4×4 system with L=2 independent data streams. For reference we have also depicted results for an i.i.d. Rayleigh fading channel (**M=0**) with blind transmission and results for a prefilter algorithm based on ST CSI [1].

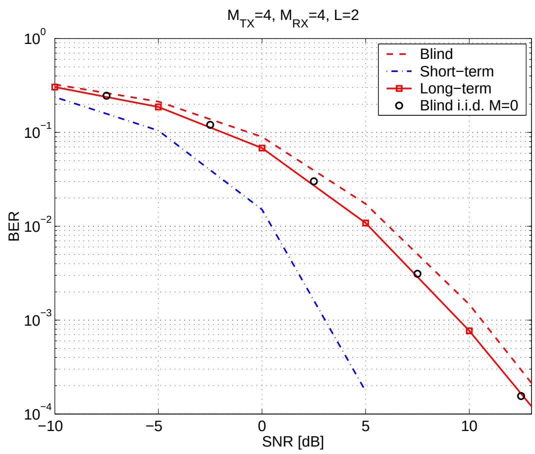

Fig. 1: ZF prefiltering (L=2, M_{RX}=M_{TX}=4, uncorrelated, K=1)

A degradation of the BER performance with non-adaptive transmitter owing to the Ricean component can be observed. Due to the rank-1 nature of the deterministic component, on average essentially an increased noise amplification takes place in the ZF receiver in the channel matrix inversion process. However, application of the novel LT CSI transmit prefilter can obviously overcome these impairments and interestingly the curves for blind transmission in the Rayleigh case and with LT prefilter in the Ricean case agree.

Similar results hold for MMSE receiver simulations for a 6×6 system with L=4 independent subchannels given in Fig. 2. We again consider a Ricean channel with K=1, however, now the scattering component exhibits transmit correlation according to an AS of 10 degrees. This results in a stronger degradation of the curve with blind transmission in a Ricean correlated channel compared to blind transmission in an i.i.d. Rayleigh fading channel with **M**=**0** (note that the 4 independent data streams are transmitted on the outer 4 transmit antennas). Obviously, via the LT CSI based prefilter we can successfully exploit the knowledge of the channel statistics and one can see a considerable performance improvement.

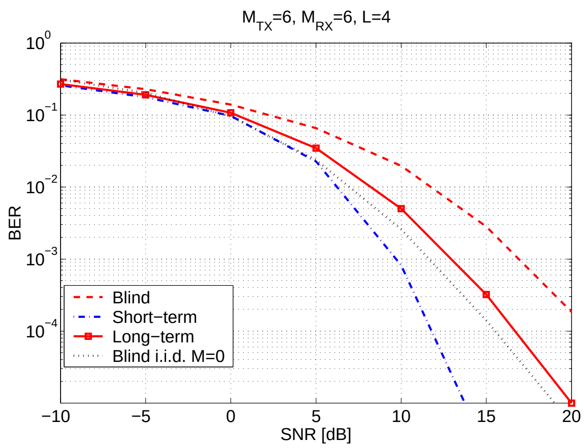

Fig. 2: MMSE prefiltering (L=4, M_{RX}=M_{TX}=6, 10° AS, K=1)

An important observation is that in the presence of strong channel correlation, it can be seen from Fig. 2 that statistical CSI becomes more valuable and the difference to the ST CSI based scheme becomes less significant.

6. APPENDIX

Definition 1. For any $x \in R^n$, let

$$x_{[1]} \geq \ldots \geq x_{[n]} \tag{29}$$

denote the elements of **x** in decreasing order (also termed order statistics).

Definition 2. [5, definition 1.A.1] Let $x, y \in R^n$. Vector **x** is majorized by vector **y** (**y** majorizes **x**) if

$$\left(\sum_{i=1}^{k} x_{[i]} \leq \sum_{i=1}^{k} y_{[i]} \quad \forall 1 \leq k \leq n-1 \right) \wedge \left(\sum_{i=1}^{n} x_{[i]} = \sum_{i=1}^{n} y_{[i]} \right) \tag{30}$$

and is represented by $\mathbf{x} \prec \mathbf{y}$.

Definition 3. [5, definition 3.A.1] A real-valued function ϕ defined on a set $A \subseteq R^n$ is said to be Schur-convex

$$\mathbf{x} \prec \mathbf{y} \text{ on } A \Rightarrow \phi(\mathbf{x}) \leq \phi(\mathbf{y}) \tag{31}$$

Similarly, ϕ is said to be Schur-concave on A if

$$\mathbf{x} \prec \mathbf{y} \text{ on } A \Rightarrow \phi(\mathbf{x}) \geq \phi(\mathbf{y}) \tag{32}$$

Theorem 3. [5, page 7] Let $x \in R^n$ and let $\mathbf{1}_n$ be a n×1 vector of all ones. Then with

$$\kappa = \left(\sum_{j=1}^{n} x_j \right) / n \tag{33}$$

it follows $\kappa \cdot \mathbf{1}_n \prec \mathbf{x}$.

Theorem 4. [5, proposition 3.C.2] If ϕ is symmetric (i.e. symmetric in its arguments) and convex, then ϕ is Schur-convex.

Theorem 5. [5, theorem 9.B.1] Let **R** be an n×n Hermitian matrix with diagonal elements denoted by the vector **d** and eigenvalues denoted by the vector λ, then

$$\mathbf{d} \prec \lambda. \tag{34}$$

Theorem 6. [5, theorem 9.B.2] If $\mathbf{h} \prec \lambda$ on R^n, then there exists a real symmetric matrix **H** with diagonal elements **h** and characteristic roots λ.

Lemma 1. [5, lemma 9.H.1.h] If **U** and **V** are n×n positive semidefinite Hermitian matrices, then

$$\mathrm{tr}(\boldsymbol{UV}) \geq \sum_{i=1}^{n} \lambda_i(\boldsymbol{U}) \cdot \lambda_{n-i+1}(\boldsymbol{V}), \tag{35}$$

where the eigenvalues are sorted in decreasing order.

Lemma 2. [4, lemma 12 in appendix] Let **R** be a full rank positive semidefinite Hermitian matrix with eigenvalue decomposition

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{U} & \tilde{\boldsymbol{U}} \end{bmatrix} \begin{bmatrix} \boldsymbol{L} & \\ & \tilde{\boldsymbol{L}} \end{bmatrix} \begin{bmatrix} \boldsymbol{U} & \tilde{\boldsymbol{U}} \end{bmatrix}^H, \tag{36}$$

where the matrix $\tilde{\boldsymbol{L}}$ contains the L largest eigenvalues in increasing order and $\tilde{\boldsymbol{U}}$ the corresponding eigenvectors. Given a M_{TX}×L (M_{TX}>L) matrix **B** such that $\mathbf{B}^H\mathbf{R}\mathbf{B}$ is diagonal with diagonal elements in increasing order, it is always possible to find a matrix $\tilde{\boldsymbol{B}}$ of the form $\tilde{\boldsymbol{B}} = \tilde{\boldsymbol{U}}\boldsymbol{\Sigma}$ with diagonal matrix Σ, such that it satisfies $\tilde{\boldsymbol{B}}\boldsymbol{R}\tilde{\boldsymbol{B}}^H = \boldsymbol{B}\boldsymbol{R}\boldsymbol{B}^H$ with $\mathrm{tr}(\tilde{\boldsymbol{B}}\tilde{\boldsymbol{B}}^H) \leq \mathrm{tr}(\boldsymbol{B}\boldsymbol{B}^H)$.

7. REFERENCES

[1] Ding Y., Davidson T. N., Zhang J.- K., Luo Z.- Q., Wong K. M., "Minimum BER block precoders for zero-forcing equalization," IEEE ICASSP, pp. 2261-2264, May 2002

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

[4] Palomar D. P., Cioffi J. M., Lagunas M. A., "Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization," *IEEE Transactions on Signal Processing*, vol. 51, no. 9, Sept. 2003

[5] Marshall A. W., Olkin I., *Inequalities: theory of majorization and its applications*, Academic Press, 1979

[6] Paulraj A., Nabar R., Gore D., *Introduction to space-time wireless communications*, Cambridge University Press, 2003

[7] Steyn H. S., Roux J. J. J., "Approximations for the non-central Wishart distribution," *South African Statistical Journal*, no. 6, pp. 165-173, 1972

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

[9] Kiessling M., Speidel J., Viering I., Reinhardt M., “Statistical prefiltering for MIMO systems with linear receivers in the presence of transmit correlation”, IEEE VTC, April 2003