RATE-OPTIMAL MIMO TRANSMISSION WITH MEAN AND COVARIANCE FEEDBACK AT LOW SNR

Ramy H. Gohary, Wessam Mesbah, and Timothy N. Davidson

Department of Electrical and Computer Engineering
McMaster University, Hamilton, Ontario, Canada

ABSTRACT

We consider a multiple-input multiple-output (MIMO) wireless communication scenario in which the channel follows a general spatially-correlated complex Gaussian distribution with non-zero mean. We derive an explicit characterization of the optimal input covariance from an ergodic rate perspective for systems that operate at low SNRs. This characterization is in terms of the eigen decomposition of a matrix that depends on the mean and the covariance of the channel, and typically results in a beamforming strategy along the principal eigenvector of that matrix. Simulation results show the potential impact of (jointly) exploiting the mean and the covariance of the channel on the ergodic achievable rate at both low and moderate-to-high SNRs.

***Index Terms*—** MIMO Communication, Partial CSI, Mean and Covariance Feedback, Ergodic Capacity, Low SNR

1. INTRODUCTION

The nature of the channel state information (CSI) that is available at the transmitter of a multiple-input multiple-output (MIMO) wireless communication system has an impact on the ergodic capacity of the system [1]. However, the significance of this impact depends on the signal-to-noise ratio (SNR) at which the system operates [2]. If the realizations of the channel are full rank with high probability, the high-SNR ergodic capacity is dominated by an SNR-dependent term that depends only on the number of transmit and receive antennas, and not on the channel state information [3, 4]. However, at low SNRs, the availability of channel information at the transmitter can have a fundamental impact [5]. Although it is desirable, from many design perspectives, for the transmitter to have full information about the actual channel realization (i.e., instantaneous CSI), in many practical communication systems this information cannot be made available to the transmitter in a timely fashion and it is more appropriate to consider systems in which the transmitter has access only to statistical information about the channel [1]. This is particularly true for systems opera ting at low-SNRs. For this reason, there have been several analyses of the impact of channel correlation on low-SNR communications (e.g., [5], [6]), but these analyses have been made under rather restricted models for the channel statistics. The goal of this paper is to provide an explicit characterization of the low-SNR rate-optimal signalling strategy for an arbitrarily correlated channel model with non-zero mean.

In some of the early work on rate-optimal signalling (in the ergodic sense) for wireless systems with multiple antennas, optimal signalling strategies for multiple-input single-output (MISO) systems were developed under certain channel models in [7] and [8] for two classes of channels; the one in which the channel mean is known at the transmitter and the channel covariance is assumed to be a (scaled) identity, and another in which the channel covariance is known at the transmitter and the mean is assumed to be zero. That work was extended to MIMO systems in [9] and [10], where rate-optimal transmission strategies were developed for the case in which the channel has zero mean, correlated rows and independent columns. For this scenario, necessary and sufficient conditions under which the covariance matrix of the optimal signalling scheme is of rank one were given in [9] and [10]. (This signalling scheme is typically referred to as beamforming [1].) Using a similar technique to the one in [9], rate-optimal signalling strategies and necessary and sufficient conditions for the optimality of beamforming for the case in which both the columns and the rows of the channel matrix are correlated were developed in [11]. For that case, a closed-form expression of the exact channel capacity was derived in [12]. In addition, sufficient (but not nece ssary) conditions under which beamforming is rate-optimal for the case in which the MIMO channel has non-zero mean and (scaled) identity covariance were derived in [13].

While MIMO channel models in which the channel is assumed to be zero mean with correlated columns and correlated rows are sufficient to characterize some practical communication scenarios, these models are unable to characterize more general correlation structures that typically occur in practice [14]. However, the extension of the above mentioned analyses to more general correlation models can be quite unwieldy. A characterization of the structure that the optimal input covariance matrix must possess for a zero-mean arbitrarily correlated channel model was provided in [14], but that work did not contain an explicit construction of this matrix.

In this paper we consider the design of the rate-optimal input covariance for low-SNR signalling over an arbitrarily correlated channel with non-zero mean. For this case, we derive an explicit closed-form expression for the low-SNR ergodic capacity, and we use that expression to derive an explicit characterization of the optimal input covariance matrix. Similar to [5], we show that whenever the maximum eigenvalues of a certain matrix are distinct, beamforming remains optimal in the presence of a non-zero mean, and an arbitrary correlation. Computing the beam direction using the approach in [5] can be quite difficult for general correlation models, and in [2] this direction was only computed for a particular model known as the UIU model. The explicit characterization provided in this paper enables us to avoid these difficulties and allows us to compute the optimal beam directions for general correlation models with non-zero mean. Furthermore, we show that this computation can be significantly simplified in scenarios in which the channel covariance matrix is structured. Our general approach and its specializations enable us to elucidate the impact of channel correlation and mea n

This work is supported in part by a Premier's Research Excellence Award from the government of Ontario. The work of the third author is also supported by The Canada Research Chair program.

on the maximum achievable data rate at low SNRs.

While the focus of this paper is on achievable rate objectives, it is worth pointing out that the availability of statistical channel information at the transmitter can be used to improve other performance metrics. For example, linear precoders have been designed to minimize various probability of error criteria [15, 16].

2. ERGODIC CAPACITY: FIRST-ORDER APPROXIMATION AT LOW SNR

In this section we provide a first-order approximation of the low-SNR ergodic capacity of a MIMO system with M transmit and N receive antennas. The channel matrix, H, is assumed to be Gaussian block fading and perfectly known at the receiver, but only its mean, $\bar{H}$, and its covariance, Φ, are known at the transmitter. Under the usual additive white Gasussian noise model, the ergodic capacity of this MIMO system is given by [4]

$$C = \max_{Q\succeq 0, \operatorname{Tr}(Q)=P} \mathrm{E}_H\big\{\log\det\big(I + HQH^\dagger\big)\big\}, \tag{1}$$

where Q is the covariance matrix of the input signal, P is the total power budget, $(\cdot)^\dagger$ denotes the Hermitian transpose operation, and the expectation is taken over the ensemble of non-zero mean spatially-correlated realizations of the channel matrix, H.

We begin our analysis of (1) by stating the following identity: For a positive definite matrix $X \succ 0$,

$$\log\det(X) = \operatorname{Tr}\big(\log(X)\big), \tag{2}$$

where the log function of a positive definite matrix is defined as the inverse of the exponential function that maps the set of Hermitian matrices to the set of positive semidefinite matrices [17]. Observing that the matrix $(I + HQH^\dagger)$ is strictly positive definite, we can invoke the identity in (2) to express the capacity in (1) as

$$C = \max_{Q\succeq 0, \operatorname{Tr}(Q)=P} \mathrm{E}_H\Big\{\operatorname{Tr}\Big(\log\big(I + HQH^\dagger\big)\Big)\Big\}. \tag{3}$$

If A is a Hermitian matrix whose maximum eigenvalue is less than unity, then the Taylor expansion of $\log(I + A)$ is,

$$\log(I + A) = A - A^2/2 + A^3/3 + \dots \tag{4}$$

Now, the matrix $HQH^\dagger$ is Hermitian, and for sufficiently low input signal power, its maximum eigenvalue, $\lambda_{\max}(HQH^\dagger)$, satisfies

$$\lambda_{\max}(HQH^\dagger) \le \epsilon \ll 1 \tag{5}$$

with high probability. In this case, the expansion in (4) is dominated by the linear term and one can approximate the capacity by

$$C \approx \max_{Q\succeq 0, \operatorname{Tr}(Q)=P} \mathrm{E}_H\Big\{\operatorname{Tr}\big(HQH^\dagger\big)\Big\}. \tag{6}$$

3. EXPLICIT LOW-SNR ERGODIC CAPACITY

In order to compute the low-SNR capacity and the corresponding optimal signalling scheme for an arbitrarily correlated Gaussian channel with non-zero mean, we will first compute the expectation in (6). Under the considered model, each channel realization can be represented by

$$\operatorname{vec}(H) = \Phi^{1/2}\operatorname{vec}(H_w) + \operatorname{vec}(\bar{H}), \tag{7}$$

where $\bar{H}$ is the channel mean and H_w is a (white) matrix whose i.i.d. entries are drawn from the standard complex Gaussian distribution, Φ is an $NM \times NM$ positive semidefinite Hermitian matrix, and $\operatorname{vec}(\cdot)$ is the operator that stacks the columns of the matrix on top of each other. We will let $\operatorname{unvec}(\cdot)$ denote the inverse of the $\operatorname{vec}(\cdot)$ operator, and for notational convenience, we will define

$$\tilde{H} = \operatorname{unvec}\big(\Phi^{1/2}\operatorname{vec}(H_w)\big). \tag{8}$$

Using the model in (7), we have that

$$\begin{aligned} &\mathrm{E}_H\Big\{\operatorname{Tr}\big(HQH^\dagger\big)\Big\} \\ &\quad= \mathrm{E}_{H_w}\Big\{\operatorname{Tr}\Big(Q^{1/2}\big(\tilde{H}^\dagger + \bar{H}^\dagger\big)\big(\tilde{H} + \bar{H}\big)Q^{1/2}\Big)\Big\} \\ &\quad= \mathrm{E}_{H_w}\Big\{\operatorname{Tr}\Big(Q^{1/2}\tilde{H}^\dagger\tilde{H}Q^{1/2}\Big)\Big\} + \operatorname{Tr}(\bar{H}^\dagger\bar{H}Q). \end{aligned} \tag{9}$$

We can write the first term on the right hand side of (9) as

$$\begin{aligned} \mathrm{E}_{\tilde{H}}\Big\{\operatorname{Tr}\big(Q^{1/2}\tilde{H}^\dagger\tilde{H}Q^{1/2}\big)\Big\} &= \mathrm{E}_{\tilde{H}}\Big\{\big(\operatorname{vec}(\tilde{H})\big)^\dagger(Q^T\otimes I)\operatorname{vec}(\tilde{H})\Big\} \\ &= \operatorname{Tr}\Big(\Phi(Q^T\otimes I)\Big), \end{aligned} \tag{10}$$

where A^T and A^* denote the transpose and conjugate of the matrix A, respectively, and in (10) we have used the identity $\operatorname{Tr}(A^\dagger B) = (\operatorname{vec}(A))^\dagger\operatorname{vec}(B)$. Using (10), we have that for the general model in (7), the low-SNR rate-optimal covariance matrix maximizes

$$\operatorname{Tr}\Big(\Phi(Q^T\otimes I_N)\Big) + \operatorname{Tr}(\bar{H}^\dagger\bar{H}Q) = \operatorname{Tr}\Big(X(Q\otimes I_N)\Big), \tag{11}$$

where, $X = \big(\Phi^T + \frac{1}{N}(\bar{H}^\dagger\bar{H}\otimes I_N)\big)$. Now, the ijth $N\times N$ block of the $NM\times NM$ matrix $(Q\otimes I_N)$ is in the form of a scaled identity. That is,

$$[Q\otimes I_N]_{(i-1)N+r,(j-1)N+s} = q_{ij}\delta_{rs}, \tag{12}$$

where q_{ij} denotes the ijth entry of Q. Using (12) in (11)

$$\begin{aligned} \operatorname{Tr}\Big(X(Q^T\otimes I_N)\Big) &= \sum_{i,j=1}^{M}\sum_{s=1}^{N} q_{ij}X_{(i-1)N+s,(j-1)N+s} \\ &= \sum_{i,j=1}^{M} q_{ij}\operatorname{Tr}\big(X_{[i,j]}\big), \end{aligned} \tag{13}$$

where we have used $X_{[i,j]}$ to denote the ijth $N\times N$ block of the matrix X. Now, define the $M\times M$ Hermitian matrix Z via:

$$[Z]_{ij} = \operatorname{Tr}\big(X_{[i,j]}\big). \tag{14}$$

Using this notation, (13) can be expressed as

$$\operatorname{Tr}\big(ZQ\big). \tag{15}$$

It may be useful at this point to particularize the results obtained so far to the separable case in which the covariance matrix can be decomposed into a finite sum of the Kronecker product of positive semidefinite matrices [14]; that is, to the case in which

$$\Phi = \sum_{s=1}^{S} T_s^T\otimes R_s, \tag{16}$$

where $\{R_s\}_{s=1}^S$ and $\{T_s\}_{s=1}^S$ are sets of $N\times N$ and $M\times M$ Hermitian positive semidefinite matrices that characterize the correlation between the receiver elements and the correlation between the transmitter elements, respectively. In that case, using steps similar to the ones used to derive (15), it can be shown that the left hand side of (9) can be expressed as $\text{Tr}(YQ)$, where

$$Y = \sum_{r,s=1}^{S} \text{Tr}(R_r^{1/2} R_s^{1/2}) T_r^{1/2} T_s^{1/2} + \bar{H}^\dagger \bar{H}. \qquad (17)$$

An interesting observation from (17) is that for the separable model in (16) the matrix Y is positive semidefinite. Actually, a consequence of the following lemma, which may be of independent interest, is the fact that the corresponding matrix Z in (14) for the case of the general channel correlation model in (7) is also positive semidefinite. We will use that fact in Section 4 to find the low-SNR rate-optimal input covariance matrix.

Lemma 1 *Let A be an $MN \times MN$ positive semidefinite matrix, and let A be partitioned into $M \times M$ blocks, each of size $N \times N$. Then the matrix constructed in such a way that each block is replaced by its main diagonal is also positive semidefinite. Moreover, the $M \times M$ matrix constructed by replacing each block by its trace is also positive semidefinite.*

See the Appendix for a proof.

4. THE OPTIMAL INPUT COVARIANCE MATRIX

We now proceed to find the covariance matrix of the input signals that maximize the low-SNR ergodic achievable rate. That is, we solve the following optimization problem:

$$\max_{Q \succeq 0,\ \text{Tr}(Q) \leq P} \quad \text{Tr}(WQ) \qquad (18)$$

where the matrix W is fixed. If the general channel correlation model in (7) is used, then $W = Z$, and if the separable model in (16) is used, then we set $W = Y$. The derivation of our solution to (18) begins with the following lemma from [18].

Lemma 2 *[18, Example 7.4.13] Let $U_A \Sigma_A V_A^\dagger$ and $U_B \Sigma_B V_B^\dagger$ denote the singular value decompositions of the $M \times M$ matrices A and B, respectively then, with the matrix A fixed*

$$\max_{U_B, V_B \text{ are unitary}} \left\{ \Re\left\{ \text{Tr}(AB^\dagger) \right\} \right\} = \text{Tr}(\Sigma_A \Sigma_B), \qquad (19)$$

where $\Re\{\cdot\}$ denotes the real part of the argument and the singular values are assumed to be arranged in the same (decreasing) order. The matrices U_B and V_B that achieve the maximum value are given by $U_B = U_A$ and $V_B = V_A$.

Since W and Q are positive semidefinite, their singular values coincide with their eigenvalues, and hence if we denote the eigen decomposition of W by $U_W \Lambda_W U_W^\dagger$ and that of Q by $U_Q \Lambda_Q U_Q^\dagger$, Lemma 2 implies that (at low SNRs) it is sufficient to optimize over the covariance matrices that satisfy

$$U_Q = U_W. \qquad (20)$$

In that case, the optimization problem in (18) can be cast as

$$\max_{\lambda_{Q_i} \geq 0,\ \sum_{i=1}^{M} \lambda_{Q_i} \leq P} \quad \sum_{i=1}^{M} \lambda_{W_i} \lambda_{Q_i} \qquad (21)$$

where λ_{W_i} and λ_{Q_i} denote the ith eigenvalue of W and Q, respectively. The optimization problem in (21) is a convex linear program that can be easily solved. If the maximum eigenvalue of W is distinct, the optimal eigenvalues of Q are given by $\lambda_{Q_1} = P$, and $\lambda_{Q_2} = \ldots = \lambda_{Q_M} = 0$. That is, beamforming along the eigenvector that corresponds to the maximum eigenvalue of W is sufficient for rate-optimal communication at low SNR. Observe that for the case in which the largest eigenvalue of W has multiplicities, any partitioning of power in the direction of the eigenvectors corresponding to these eigenvalues is optimal up to first-order approximation. However, for up to second-order optimality, power may have to be carefully distributed across these directions [6].

Remark 1 Observe that for the case in which the separable model corresponds to the so-called Kronecker model (i.e., the case of $S = 1$ in (16)), the optimum beam direction is along the eigenvector that corresponds to the maximum eigenvalue of $(\text{Tr}(R)T + \bar{H}^\dagger \bar{H})$. If the channel is zero mean, the matrix R does not affect the optimal beam direction, and, in agreement with the result in [11], this direction is along the principal eigenvector of T. In contrast, if the channel is non-zero mean, our result suggests that $\text{Tr}(R)$ acts as a weight that controls the relative impact of the eigenvectors of T and that of the eigenvectors of $\bar{H}^\dagger \bar{H}$ on the optimal beam direction. □

5. A NUMERICAL EXAMPLE

In this section we provide a numerical example that illustrates the utility of the explicit characterization of the optimal low-SNR signalling strategy derived in Section 4. We consider a MIMO system with five transmit and five receive antennas; i.e., $M = N = 5$. With the noise power normalized to unity, Figure 1 shows a plot of the achievable rates under different signalling strategies against the transmitted signal power . The channel mean $\bar{H}$ was randomly chosen and its Frobenius norm was set to be equal to 0.866. The matrix Φ in the general correlation model in (7) was also randomly chosen and its trace was set to 2.5. In order to provide a benchmark for these results, in Figure 1 we also provide an upper bound that corresponds to the maximum rate that would be achievable had the channel realizations been perfectly known to the transmitter (i.e., instantaneous CSI). The signalling strategies considered i n Figure 1 are as follows:

1. The transmitter ignores all channel information and transmits isotropically; i.e., $Q = (P/M)I$.
2. The transmitter ignores the channel covariance information and treats the mean as if it were the actual channel. Hence, transmission takes place along the eigenvectors of $\bar{H}^\dagger \bar{H}$ with the eigenvalues of Q chosen so as to "water-fill" over those of $\bar{H}^\dagger \bar{H}$. (At low SNRs this will result in beamforming along the principal eigenvector of $\bar{H}^\dagger \bar{H}$.)
3. The transmitter employs the low-SNR rate-optimal strategy derived herein. That is, beamforming along the principal eigenvector of W; cf., (18).
4. The transmitter signals along the eigenvectors of W (cf., (20)) without restricting the input covariance matrix to be rank one. For each input power constraint, the eigenvalues of Q are found by solving the problem in (1) (with $U_Q = U_W$) using gradient-based stochastic optimization techniques. (At sufficiently low input powers, this strategy reduces to Strategy 3.)

Figure 1 illustrates the optimality of beamforming along the principal eigenvectors of W (Strategy 3) at low SNRs. As expected, it also illustrates the price paid at higher SNRs for the restriction of the signal covariance to be rank one. However, by selecting the eigenvectors of the input covariance matrix as in (20) and optimizing over

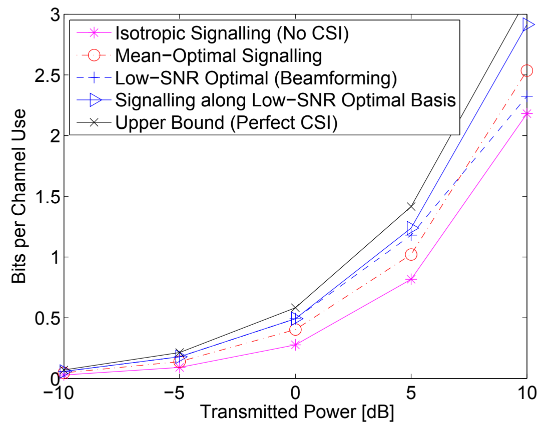

Fig. 1. Achievable rates versus transmitted signal power for different transmission schemes over a correlated channel model with non-zero mean. The signalling schemes are isotropic signalling (Strategy 1), mean-optimal signalling (Strategy 2), signalling along low-SNR optimal basis (Strategy 4) and beamforming (Strategy 3).

the set of feasible eigenvalues (Strategy 4), we obtain consistently higher achievable rates than any of the other considered strategies. While we cannot claim optimality of Strategy 4 at this point, we have performed numerous experiments, all of which suggest that this Strategy 4 is capable of achieving rates that are higher than those that can be achieved under known signalling strategies.

6. CONCLUSION

In this paper we have provided explicit characterizations of the low-SNR rate-optimal input covariance matrix when both the channel mean and covariance are known at the transmitter. We have considered both separable and general channel correlation models and have demonstrated numerically the optimality of the computed beam directions at low SNR.

7. APPENDIX: PROOF OF LEMMA 1

First, observe that Lemma 1 holds trivially for $M = 1$ and $N = 1$. Now, consider general values for M and N. For any $i \in \{1, \ldots, MN\}$, define the matrix E_i to be the $MN \times MN$ all zero matrix with the ith entry on the main diagonal replaced by one. Let the matrix constructed by replacing each block by its main diagonal be denoted by G. Then one can verify that

$$G = \sum_{i=1}^{N} \Big(\sum_{j=1}^{M} E_{(j-1)N+i} \Big) A \Big(\sum_{j=1}^{M} E^T_{(j-1)N+i} \Big). \qquad (22)$$

Since A is positive semidefinite, then so is $\Big(\sum_{j=1}^{M} E_{(j-1)N+i} \Big) A \Big(\sum_{j=1}^{M} E^T_{(j-1)N+i} \Big)$. Invoking the fact that the sum of positive semidefinite matrices is also positive semidefinite completes the proof of the first part of the lemma.

In order to prove the second part, let F denote the $M \times M$ matrix constructed by replacing each block by its trace, and let $\mathbf{1}_N$ be the $N \times 1$ vector in which all entries are equal to unity. Then

$$F = (I_M \otimes \mathbf{1}_N^T) G (I_M \otimes \mathbf{1}_N). \qquad (23)$$

From the first statement of the lemma, we know that G is positive semidefinite. Therefore, F is also positive semidefinite, which completes the proof of the second statement of the lemma.

8. REFERENCES

[1] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, “Capacity limits of MIMO channels,” *IEEE J. Select. Areas Commun.*, vol. 21, pp. 684–701, June 2003.

[2] A. M. Tulino, A. Lozano, and S. Verdú, “Impact of antenna correlation on the capacity multiantenna channels,” *IEEE Trans. Inform. Theory*, vol. 51, pp. 2491–2509, July 2005.

[3] G. J. Foschini and M. J. Gans, “On limits of wireless communication in a fading environment when using multiple antennas,” *Wireless Personal Commun.*, vol. 6, pp. 311–335, Mar. 1998.

[4] I. E. Telatar, “Capacity of multiantenna Gaussian channels,” *Eur. Trans. Telecom.*, vol. 10, pp. 585–595, Nov. 1999.

[5] S. Verdú, “Spectral efficiency in the wideband regime,” *IEEE Trans. Inform. Theory*, vol. 48, pp. 1319–1343, June 2002.

[6] A. Lozano, A. M. Tulino, and S. Verdú, “Multiple-antenna capacity in the low-power regime,” *IEEE Trans. Inform. Theory*, vol. 49, pp. 2527–2544, Oct. 2003.

[7] E. Visotsky and U. Madhow, “Space-time transmit precoding with imperfect feedback,” *IEEE Trans. Inform. Theory*, vol. 47, pp. 2632–2639, Sept. 2001.

[8] A. L. Moustakas and S. H. Simon, “Optimizing multiple-input single-output MISO communication systems with general Gaussian channels: Nontrivial covariance and nonzero mean,” *IEEE Trans. Inform. Theory*, vol. 49, pp. 2770–2780, Oct. 2003.

[9] S. A. Jafar, S. Vishwanath, and A. Goldsmith, “Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback,” in *Proc. IEEE Int. Conf. Commun.*, (Helsinki), pp. 2266–2270, June 2001.

[10] S. H. Simon and A. L. Moustakas, “Optimizing MIMO antenna systems with channel covariance feedback,” *IEEE J. Select. Areas Commun.*, vol. 21, pp. 406–417, Apr. 2003.

[11] E. A. Jorswieck and H. Boche, “Channel capacity and capacity-range of beamforming in MIMO wireless systems under correlated fading with covariance feedback,” *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1543–1553, Sept. 2004.

[12] M. Kießling, “Unifying analysis of ergodic MIMO capacity in correlated Rayleigh fading environment,” *Europ. Trans. Telecommun.*, vol. 16, pp. 17–35, 2005.

[13] S. H. Simon and A. L. Moustakas, “Optimality of beamforming in multiple transmitter multiple receiver communication systems with partial channel knowledge,” in *Proc. DIMACS Wkshp Signal Processing Wireless Commun.*, Series Discrete Math. and Theoretical Comp. Sci., (Rutgers Univ.), Oct. 2002.

[14] I. Bjelaković and H. Boche, “Structure of optimal input covariance matrices for MIMO systems with covariance feedback under general correlated fading,” in *Proc. IEEE Int. Symp. Inform. Theory*, (Seattle), pp. 1041–1045, July 2006.

[15] M. Vu and A. Paulraj, “Optimal linear precoders for MIMO wireless correlated channels with nonzero mean in space-time coded systems,” *IEEE Trans. Signal Processing*, vol. 54, pp. 2318–2332, June 2006.

[16] A. Hjørungnes and D. Gesbert, “Precoding of orthogonal space-time block codes in arbitrarily correlated MIMO channels: Iterative and closed-form solutions,” *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1072–1082, Mar. 2007.

[17] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge: Cambridge University Press, 1994.

[18] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge: Cambridge University Press, 1999.