Impact of correlation on linear precoding in QSTBC coded systems with linear MSE detection

Aydin Sezgin and Arogyaswami Paulraj
Stanford University
Information Systems Laboratory
350 Serra Mall
{sezgin,apaulraj}@stanford.edu

Mai Vu
Harvard University
School of Engineering and Applied Sciences
33 Oxford Street, MD342
maivu@deas.harvard.edu

***Abstract*— In this paper, we study a wireless multiple-input multiple-output system in a Rayleigh flat-fading environment with correlation among the transmit antennas. We assume that the receiver has perfect CSI and the transmitter only knows the correlation matrix. The transmitter employs a quasi-orthogonal space-time block code in combination with a linear precoder; the receiver uses a linear MMSE detector.**

We analyze the optimal transmit precoding strategy that minimizes the average sum MSE at the receiver. We show that, as expected, the optimal precoding directions coincide with the eigenvectors of the transmit correlation matrix. The optimal power allocation, however, only supports at most 2 directions at all SNRs independent of the number of transmit antennas, which correspond to the 2 largest eigenvalues of the transmit correlation matrix. We characterize this optimal power allocation by the necessary and sufficient optimality conditions. At high SNRs, the optimal allocation approaches equal power on the two supported modes. At low SNRs, the weaker mode is dropped and the precoding matrix becomes single-mode beamforming. We provide a closed-form expression characterizing this low-SNR range. Numerical simulations confirm our theoretical analysis.

I. Introduction

It has been shown that in a sufficiently rich scattering environment the capacity of multiple antenna systems grows linearly with the minimum of antennas used at the transmitter and receiver [1]. Different types of channel state information (CSI) at the transmitter were studied, e.g. no CSI, perfect CSI [1] and imperfect CSI [2] as mean-feedback [3]–[5] or covariance feedback [4], [6], [7]. For MIMO systems with partial CSI the eigenbeamforming concept was coined in [8] and the impact of imperfect CSI was analyzed recently in [9], [10]. Furthermore, transmit strategies using a linear precoder in combination with space-time codes like orthogonal space-time block codes (OSTBC) and quasi-orthogonal space-time block codes (QSTBC) have been analyzed in [5], [11] and [12], [13], respectively. A framework for the case in which mean and covariance information at the transmitter is available is treated in [14].

In the optimization of the capacity and the error rate performance it is often assumed implicitly that maximum-likelihood (ML) detection will be performed at the receiver. This task, however, requires an exhaustive search, which becomes more and more computationally infeasible especially for high number of transmit antennas and higher modulation orders. In contrast to the capacity and error rate considerations with the optimal receiver structure, in this work, we assume that the receiver applies a linear minimum mean square error (MMSE) receiver. The transmitter employs a QSTBC in combination with a linear precoder. As a result of using the MMSE, the performance metric to optimize is the average normalized MSE. Optimizing the MSE is also relevant for detectors employing a linear front end like the VBLAST algorithm or sphere detectors. Given the aforementioned scenario, we analyze the optimal transmit strategy, i.e. the optimal transmit direction, which determines the eigenvectors of the precoding matrix, and the optimal power allocation, which determines the eigenvalues of the precoding matrix. We show that it is optimal to transmit into the direction of the eigenvectors of the transmit correlation matrix. The optimal power allocation is characterized using the necessary and sufficient optimality conditions. It turns out, that for high SNR only the two largest eigenvalues of the correlation matrix are supported, which is counterintuitive. We show that for small SNR values, the optimal precoding matrix has rank one, i.e. single mode beamforming is optimal. A closed form expression characterizing the SNR range, in which only one eigenvalue is supported, is derived. Finally, we illustrate the theoretical results by numerical simulations.

II. System model, QSTBC and normalized average MSE

A. System model

Suppose that the transmitter has N antennas and the receiver has M antennas. Consider a quasi-static flat fading channel $\mathbf{H}$ with transmit antenna correlation, the channel can be modeled as $\mathbf{H} = \mathbf{W}\mathbf{R}_t^{1/2}$, where $\mathbf{W}$ has identically independently distributed (iid) complex Gaussian entries ($\mathbf{W} \sim \mathcal{CN}(0, \mathbf{I})$) and $\mathbf{R}_t$ is the transmit correlation matrix. Let the eigenvalue decomposition of $\mathbf{R}_t$ be $\mathbf{R}_t = \mathbf{U}_R\mathbf{\Lambda}_R\mathbf{U}_R^H$, with the eigenvalues $[r_1, ..., r_N] = \text{diag}\left(\mathbf{\Lambda}_R\right)$ decreasingly ordered as $r_1 \geq r_2 \geq \ldots \geq r_N \geq 0$. The power of the $\mathbf{R}_t$ is normalized such that $\text{trace}(\mathbf{R}_t) = N$.

With the $N \times T$ transmit matrix $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_\mathbf{T}]$, where $\mathbf{x}_t$ is the signal vector at time instant t, $1 \leq t \leq T$, the

The work of A.Sezgin is supported in part by the Deutsche Forschungsgemeinschaft (DFG) and by NSF Contract NSF DMS-0354674 ONR Contract ONR N00014-02-1-0088-P00006.

received signal vector $\mathbf{y}_t$ at time instant t is then given by

$$\mathbf{y}_t = \mathbf{H}\mathbf{x}_t + \mathbf{n}_t \tag{1}$$

where $\mathbf{n}_t \sim \mathcal{CN}(0, \sigma_n^2 \mathbf{I})$ is the additive white Gaussian noise (AWGN). The transmit signal has the covariance matrix $\mathbf{Q} = \mathbb{E}\left[\mathbf{x}\mathbf{x}^H\right]$ with the eigenvalue decomposition $\mathbf{Q} = \mathbf{U}_Q \mathbf{\Lambda}_Q \mathbf{U}_Q^H$. Again we order the eigenvalues $[p_1, ..., p_N] = \mathrm{diag}\left(\mathbf{\Lambda}_Q\right)$ decreasingly as $p_1 \geq p_2 \geq ... \geq p_N \geq 0$. Assuming a total transmit power P, then $\mathrm{tr}(\mathbf{Q}) \leq P$. We define the transmit SNR as $\rho = \frac{P}{\sigma_n^2}$.

B. QSTBC

We consider a system with a QSTBC $\mathbf{G}$, with $\mathbf{X} = \mathbf{Q}^{1/2}\mathbf{G}$, at the transmitter. The QSTBC $\mathbf{G}$ for N transmit antenna is constructed as follows. Starting with the well known Alamouti scheme [15] for $N = 2$, $\mathbf{G}_2(s_1, s_2) = \left[\begin{smallmatrix} s_1 & s_2 \\ s_2^* & -s_1^* \end{smallmatrix}\right] = [\mathbf{x}_1, \mathbf{x}_2]$, then for $N = 2^n$ $(N \geq 4)$ [16]

$$\mathbf{G}_N\left(\mathbf{s}\right) = \left[\begin{array}{cc} \mathbf{G}_{\frac{N}{2}}\left(s_1, \ldots, s_{\frac{N}{2}}\right) & \mathbf{G}_{\frac{N}{2}}\left(s_{\frac{N}{2}+1}, \ldots, s_N\right) \\ \mathbf{G}_{\frac{N}{2}}\left(s_1, \ldots, s_{\frac{N}{2}}\right)\mathbf{\Theta} & -\mathbf{G}_{\frac{N}{2}}\left(s_{\frac{N}{2}+1}, \ldots, s_N\right)\mathbf{\Theta} \end{array}\right],$$

where $\mathbf{\Theta} = \mathrm{diag}\left(\{(-1)^{k-1}\}_{k=1}^{N/2}\right)$ and $\mathbf{s} = [s_1, \ldots, s_j, \ldots, s_N]^T$ are the information symbols drawn from a PSK/QAM modulation alphabet. The channel is assumed to be constant for the whole code block.

1) Equivalent channel representation: Before analyzing the performance metric, which is the average normalized MSE, let us introduce the equivalent channel representation. First of all, let us split the vector $\mathbf{s}$ into two vectors, $\mathbf{s}_\mathrm{o}$ and $\mathbf{s}_\mathrm{e}$. The elements of $\mathbf{s}$ with odd index j are collected in $\mathbf{s}_\mathrm{o}$ and with even index in $\mathbf{s}_\mathrm{e}$, respectively. After some preprocessing (mainly complex conjugating some elements of the receive vector $\mathbf{y}$), channel matched filtering, and noise-whitening [16], [17], the received signal vector $\tilde{\mathbf{y}}$ in the equivalent channel model is given by

$$\tilde{\mathbf{y}} = \underbrace{\left[\begin{array}{cc} \mathbf{H}' & \mathbf{0} \\ \mathbf{0} & \mathbf{H}' \end{array}\right]}_{\tilde{\mathbf{H}}} \left[\begin{array}{c} \mathbf{s}_\mathrm{o} \\ \mathbf{s}_\mathrm{e} \end{array}\right] + \tilde{\mathbf{n}} \tag{2}$$

with additive white Gaussian noise (AWGN) vector $\tilde{\mathbf{n}} \sim \mathcal{CN}(0, \sigma_n^2 \mathbf{I})$. The properties of the equivalent $N/2 \times N/2$ channel $\mathbf{H}'$ are discussed in the following The Gram of the equivalent $N/2 \times N/2$ channel $\mathbf{H}'$ in (2) has the following eigenvalue decomposition $\mathbf{H}'^H\mathbf{H}' = \mathbf{V}_{H'}\mathbf{S}\mathbf{V}_{H'}^H$.

In the following, we present the eigenvalues of $\mathbf{H}'^H\mathbf{H}'$ for the case of $N = 4$ in order to gain more insight into the structure of the eigenvalues. Afterward, the structure of the eigenvalues for the general case is given.

Example II.1: *In the case of $N = 4$ transmit antennas $\mathbf{H}'$ is given by $\mathbf{H}' = \left[\begin{smallmatrix} \mu_1 & i\mu_1 \\ \mu_2 & -i\mu_2 \end{smallmatrix}\right]$, with $\mu_1 = \sqrt{\frac{\lambda+\beta}{2}}$, $\mu_2 = \sqrt{\frac{\lambda-\beta}{2}}$, $\lambda = \sum_{i=1}^M \sum_{j=1}^N |h_{ji}|^2$ and $\beta = \sum_{i=1}^M 2\mathrm{Im}(h_{1i}^* h_{3i} + h_{4i}^* h_{2i})$, where h_{ji} is the channel gain from the j-th transmitter $(1 \leq j \leq N)$ to the i-th receiver $(1 \leq i \leq M)$ of the actual quasi-static block flat fading MIMO channel [17]. Furthermore, $\mathbf{V}_{H'}$ and $\mathbf{S}^{\frac{1}{2}}$ are given as*

$$\mathbf{V}_{H'} = \frac{1}{\sqrt{2}}\left[\begin{array}{cc} 1 & 1 \\ -i & i \end{array}\right], \mathbf{S}^{\frac{1}{2}} = \sqrt{2}\left[\begin{array}{cc} \mu_1 & 0 \\ 0 & \mu_2 \end{array}\right].$$

Lemma 1 *For any N and M, the eigenvalues μ_k^2 of $\frac{2}{N}\mathbf{S}$ depend on the actual channel $\mathbf{H}$ and on the transmit strategy $\mathbf{Q}$ by*

$$\mu_k^2 = \mathrm{tr}\left(\mathbf{H}\mathbf{U}_Q {\mathbf{\Lambda}_\mathbf{Q}}^{1/2} \mathbf{A}_N^k {\mathbf{\Lambda}_\mathbf{Q}}^{1/2} \mathbf{U}_Q^H \mathbf{H}^H\right). \tag{3}$$

where the matrices $\mathbf{A}_N^k$ are given in the Appendix.

Proof: We omit the proof because it follows the same line of arguments as the proof given in [16, Lemma 4.2] and replacing $\mathbf{H}$ by $\mathbf{H} = \mathbf{H}\mathbf{U}_Q{\mathbf{\Lambda}_\mathbf{Q}}^{\frac{1}{2}}$. ■

C. Normalized average MSE

Since a MMSE receiver is applied, the performance metric introduced is the average normalized MMSE. We follow the definition and derivation of the normalized MSE in [18].The linear MMSE receiver computes the data estimate from the received signal vector in the equivalent channel model (1)

$$\hat{\mathbf{x}} = \rho\tilde{\mathbf{H}}^H\left(\mathbf{I} + \rho\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H\right)^{-1}\tilde{\mathbf{y}}.$$

The covariance matrix of the estimation error ϵ is given as

$$\mathbf{K}_\epsilon = \mathbf{I} - \rho\tilde{\mathbf{H}}^H\left(\rho\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H + \mathbf{I}\right)^{-1}\tilde{\mathbf{H}}. \tag{4}$$

From (4), it follows the average normalized MSE as

$$\mathrm{MSE} = \mathbb{E}\mathrm{tr}(\mathbf{K}_\epsilon) = N - \mathbb{E}\mathrm{tr}\left(\rho\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H\left[\mathbf{I} + \rho\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H\right]^{-1}\right). \tag{5}$$

The last term in (5) can be rewritten and we obtain for the average MSE[1] as a function of $\mathbf{R}_t$, and the transmit strategy $\mathbf{Q}$

$$\mathrm{MSE}(\mathbf{R}_t, \mathbf{Q}) = N - M + \mathbb{E}\mathrm{tr}\left(\left[\mathbf{I} + \rho\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H\right]^{-1}\right) \tag{6}$$

$$= \sum_{k=1}^N \mathbb{E}\left[\frac{1}{1 + \rho\mu_k^2(\mathbf{R}_t, \mathbf{Q})}\right]. \tag{7}$$

The eigenvalues $\mu_k^2(\cdot)$ are defined and explained in [16].

III. Optimal transmit strategy

Having analyzed the general properties of the MSE in the last section, in the following, we will solve the optimization problem

$$\min_{\mathbf{Q}\succeq 0, \mathrm{tr}\mathbf{Q} \leq P} \mathrm{MSE}(\mathbf{R}_t, \mathbf{Q}). \tag{8}$$

[1]The MSE depends on the channel matrix $\mathbf{H}$. The average is computed with respect to the random part $\mathbf{W}$ of the real channel realization.

We first rewrite (7) as

$$\begin{aligned}\mathrm{MSE}(\mathbf{R}_t, \mathbf{Q}) &= \sum_{k=1}^{N} \mathbb{E}\left[\int_0^\infty e^{-t\left(1+\rho\mu_k^2(\mathbf{R}_t,\mathbf{Q})\right)}\, dt\right] \\ &= \sum_{k=1}^{N} \int_0^\infty e^{-t}\mathbb{E}\left[e^{-t\rho\mu_k^2(\mathbf{R}_t,\mathbf{Q})}\right]\, dt\end{aligned}$$

by using the identity $(1+x)^{-1} = \int_0^\infty e^{-t(1+x)}\, dt$. Using the probability density function of the channel

$$p(\mathbf{H}) = \frac{1}{\pi^{NM}\det(\mathbf{R_t})^M}\exp\left(-\mathrm{tr}\left(\mathbf{H}\mathbf{R}_t^{-1}\mathbf{H}^H\right)\right)$$

and define $\mathbf{F} = \mathbf{Q}^{1/2} = \mathbf{U}_Q{\mathbf{\Lambda_Q}}^{1/2}$ we have

$$\mathbb{E}\left[e^{-t\rho\mu_k^2(\mathbf{R}_t,\mathbf{F})}\right] = \frac{1}{\det\left(\mathbf{R}_t\right)^M\det\left(t\rho\mathbf{F}\mathbf{A}_N^k\mathbf{F}^H + \mathbf{R}_t^{-1}\right)^M}$$

It follows that

$$\mathrm{MSE}(\mathbf{R}_t, \mathbf{F}) = \sum_{k=1}^{N}\int_0^\infty \frac{\det\left(\mathbf{R}_t\right)^{-M}\exp(-t)\, dt}{\det\left(t\rho\mathbf{F}\mathbf{A}_N^k\mathbf{F}^H + \mathbf{R}_t^{-1}\right)^M}.$$

Using the eigenvalue decomposition of $\mathbf{R}_t$ results in

$$\mathrm{MSE}(\mathbf{R}_t, \mathbf{F}) = \sum_{k=1}^{N}\int_0^\infty \frac{\det\left(\mathbf{\Lambda}_R\right)^{-M}\exp(-t)\, dt}{\det\left(t\rho\mathbf{U}_R^H\mathbf{F}\mathbf{A}_N^k\mathbf{F}^H\mathbf{U}_R + \mathbf{\Lambda}_R^{-1}\right)^M}.$$

Defining $\tilde{\mathbf{Q}}^{1/2} = \mathbf{U}_R^H\mathbf{F} = \mathbf{U}_R^H\mathbf{Q}^{1/2}$ and using it in $\mathrm{tr}(\mathbf{Q}) \le P$ from (8) results in $\mathrm{tr}(\tilde{\mathbf{Q}}) \le P$ which is minimized by choosing $\mathbf{U}_Q = \mathbf{U}_R$. Thus, we have

$$\mathrm{MSE}(\mathbf{R}_t, \mathbf{F}) = \sum_{k=1}^{N}\int_0^\infty \frac{\exp(-t)}{\det\left(t\rho\mathbf{\Lambda}_R\mathbf{\Lambda_Q}\mathbf{A}_N^k + \mathbf{I}\right)^M}\, dt. \tag{9}$$

Lemma 2 *Let $\mathbf{D}$ be a diagonal matrix with arbitrary entries. Then the following statement holds*

$$\det\left(t\rho\mathbf{D}\mathbf{A}_N^k + \frac{1}{a}\mathbf{I}\right) = \det\left(t\rho\mathbf{D}\mathbf{A}_N^{k'} + \frac{1}{a}\mathbf{I}\right),$$

for $1 \le k, k' \le N$. Further, since the matrices $\mathbf{A}_N^k$ have rank two the only eigenvalues of $\mathbf{D}\mathbf{A}_N^k$, $\forall k, 1 \le k \le \frac{N}{2}$, unequal to zero are given as

$$\frac{2}{N}\sum_{k=1}^{\frac{N}{2}} d_{2k-1}, \quad \textit{and}\ \frac{2}{N}\sum_{k=1}^{\frac{N}{2}} d_{2k}.$$

Proof: The proof is given in [19]. ∎

Thus, the eigenvalues of $\mathbf{\Lambda}_R\mathbf{\Lambda_Q}\mathbf{A}_N^k$, $\forall k, 1 \le k \le \frac{N}{2}$, unequal to zero are given as

$$\nu_1 = \frac{2}{N}\sum_{k=1}^{\frac{N}{2}} p_{2k-1}r_{2k-1}, \quad \text{and}\ \nu_2 = \frac{2}{N}\sum_{k=1}^{\frac{N}{2}} p_{2k}r_{2k}.$$

Therefore, we can rewrite (9) as

$$\mathrm{MSE} = N\int_0^\infty \frac{\exp(-t)}{\left(t\rho\nu_1 + 1\right)^M\left(t\rho\nu_2 + 1\right)^M}\, dt. \tag{10}$$

Theorem 1: The optimal transmit precoding activates at most two eigen-modes at all SNRs. Specifically, let the eigenvalues of the transmit correlation matrix be decreasingly ordered as $r_1 \ge r_2 \ge \cdots \ge r_N \ge 0$. Then it is optimal to transmit into at most two eigen-directions which correspond to r_1 and r_2.

Proof: With $\sum_{k=1}^{N/2} p_{2k-1} = P_o$, $\sum_{k=1}^{N/2} p_{2k} = P_e = P - P_o$ and $\sum_{k=1}^{N/2}\alpha_k = 1$, $\sum_{k=1}^{N/2}\beta_k = 1$, we can rewrite (10) as

$$\begin{aligned}\mathrm{MSE} = \int_0^\infty &\frac{N\exp(-t)}{\left(t\rho P_o\frac{2}{N}\sum_{k=1}^{N/2}\alpha_k r_{2k-1} + 1\right)^M} \\ &\times \frac{1}{\left(t\rho(P-P_o)\frac{2}{N}\sum_{k=1}^{N/2}\beta_k r_{2k} + 1\right)^M}\, dt\end{aligned} \tag{11}$$

Eq. (11) can be interpreted as follows. Given the total power constraint of P we allocate a power of P_o and $P_e = P - P_o$ to the odd and even eigenvalues of $\mathbf{R}_t$, respectively. The weights α_k and β_k thereby determine, how much power is allocated to each eigenvalue. The optimal weights α_k and β_k are obtained by using the following lemma

Lemma 3 *Let μ_i $(i = 1\ldots n)$ be nonnegative, decreasingly ordered numbers ($\mu_i > \mu_j$ for $i < j$). Then the weighted sum $s = \sum_{i=1}^n a_i\mu_i$, where $a_i \ge 0$ and $\sum_i a_i = 1$, is maximized when $a_1 = 1$.*

Proof: Given any set of weights that contains more than 1 non-zero value, say a_k and a_j, then the sum $s = a_k\mu_k + a_j\mu_j$ can always be increased by creating new weights as $\tilde{a}_1 = a_k + a_j$ and other $\tilde{a}_i$ equal to zero, which leads to the new sum $\tilde{s} = (a_k + a_j)\mu_1 > a_k\mu_k + a_j\mu_j = s$.

If there are multiple maximum values of μ_i, the optimal weights can be distributed equally among these values. ∎

Using Lemma 3, the optimal weights for (11) are $\alpha_1 = \beta_1 = 1$ and $\alpha_i = \beta_i = 0$, $2 \le i \le N/2$, i.e. only the two largest eigenvalues are supported, independent of the SNR. This concludes the proof. ∎

Using this in (11) gives

$$\begin{aligned}\mathrm{MSE} &= N\int_0^\infty \frac{\exp(-t)}{\left(t\rho P_o\frac{2}{N}r_1 + 1\right)^M\left(t\rho(P-P_o)\frac{2}{N}r_2 + 1\right)^M}\, dt \\ &= N\mathbb{E}_t\left[\frac{1}{\left(t\rho P_o\frac{2}{N}r_1 + 1\right)^M\left(t\rho P_e\frac{2}{N}r_2 + 1\right)^M}\right],\end{aligned} \tag{12}$$

with the pdf of the random variable t given as $p_t(t) = \exp(-t)$. In the next step, we characterize the optimal power allocation in terms of necessary and sufficient optimality conditions. The convex optimization problem is then

$$\min_{\mathbf{\Lambda_Q}}\ N\mathbb{E}_t\left[\prod_{k=1}^{2}\frac{1}{\left(t\rho p_k\frac{2}{N}r_k + 1\right)^M}\right]$$

$$\text{subject to}\quad \sum_{k=1}^{2} p_k \le P \text{ and } p_k \ge 0, k = 1, 2.$$

The Lagrangian function for the above problem is given by

$$\mathcal{L}(\mathbf{\Lambda_Q}, \mu, \mathbf{v}) = N\mathbb{E}_t \left[\prod_{k=1}^{2} \frac{1}{\left(t\rho p_k \frac{2}{N} r_k + 1\right)^M} \right] + \mu \left(\sum_{k=1}^{2} p_k - P \right) - \sum_{k=1}^{2} v_k p_k.$$

The necessary and sufficient optimality conditions (or Karush-Kuhn-Tucker (KKT) conditions [20]) are obtained by setting the derivative of $\mathcal{L}$ wrt p_k to zero. For $k = 1$ and $k = 2$, these conditions are given by

$$\mathbb{E}_t \left[\frac{-NMt\rho \frac{2}{N} r_1}{\left(t\rho P_o \frac{2}{N} r_1 + 1\right)^{M+1} \left(t\rho(P - P_o)\frac{2}{N} r_2 + 1\right)^M} \right] = v_1 - \mu \tag{13}$$

and

$$\mathbb{E}_t \left[\frac{-NMt\rho \frac{2}{N} r_2}{\left(t\rho P_o \frac{2}{N} r_1 + 1\right)^{M} \left(t\rho(P - P_o)\frac{2}{N} r_2 + 1\right)^{M+1}} \right] = v_2 - \mu. \tag{14}$$

From these conditions we observe the following behavior of the optimal power allocation for asymptotically high SNR values and for small SNR values.

Corollary 1: *Let the eigenvalues of the transmit correlation matrix be ordered in descending order $r_1 \geq r_2 \geq \cdots \geq r_N \geq 0$. For asymptotically high SNR values, the optimal power allocation is to allocate equal power to the two largest eigenvalues r_1 and r_2.*

Proof: Suppose that $P_o^\star$ $(0 < P_o^\star < P)$ is the optimal power as $\rho \to \infty$. Break the integral in (12) into two integrals as follows

$$\int_0^\infty [...]dt = \int_0^{K/\rho} [...]dt + \int_{K/\rho}^\infty [...]dt$$

where K is a constant chosen such that

$$KP_o^\star \frac{2}{N} r_1 \gg 1 \text{ and } K(P - P_o^\star)\frac{2}{N} r_2 \gg 1.$$

Then, as $\rho \to \infty$, the expression in the denominator of the second integral can be approximated as

$$\left(t^2 K P_o^\star \frac{2}{N} r_1 K (P - P_o^\star) \frac{2}{N} r_2 \right)^M$$

while the first integral approaches 0. Thus minimizing the MSE is equivalent to maximizing the product $P_o^\star(P - P_o^\star)$, which leads to $P_o^\star = P/2$. ∎

Theorem 2: Let the eigenvalues of the transmit correlation matrix be ordered in descending order $r_1 \geq r_2 \geq \ldots, \geq r_N \geq 0$. The necessary and sufficient condition for the optimality of single-mode beamforming, i.e. $p_1 = P_o = P$, $p_2 = p_3 = \cdots = p_N = 0$ is given by

$$\mathbb{E}_t \left[\frac{t\rho^* \frac{2}{N} r_1}{\left(t\rho^* P \frac{2}{N} r_1 + 1\right)^{M+1}} \right] \geq \mathbb{E}_t \left[\frac{t\rho^* \frac{2}{N} r_2}{\left(t\rho^* P \frac{2}{N} r_1 + 1\right)^{M}} \right] \tag{15}$$

Proof: The optimal transmit strategy if only one direction is supported, i.e. the SNR $\rho \leq \rho^*$, is to allocate the complete power in direction of the largest eigenvalue r_1, i.e. $p_1 = P_0 = P$. The KKT conditions require that $v_k p_k = 0$ at the optimal values, thus $v_1 = 0$ and $v_2 \geq 0$. Coupled with (13) and (14), we obtain eq. (15). ∎

Now define the following function $\gamma(\rho)$ given as

$$\gamma(\rho) = \mathbb{E}_t \left[\frac{t\rho^* \frac{2}{N} r_2}{\left(t\rho^* P \frac{2}{N} r_1 + 1\right)^{M}} \right] - \mathbb{E}_t \left[\frac{t\rho^* \frac{2}{N} r_1}{\left(t\rho^* P \frac{2}{N} r_1 + 1\right)^{M+1}} \right].$$

As long as $\gamma(\rho) \leq 0$, from Theorem 2 it follows that single-mode beamforming is optimal. Interestingly, the above expression can be evaluated in closed form resulting in

$$\gamma(\rho) = e^{\frac{1}{\frac{2}{N}\rho r_1 P}} \left(\frac{2}{N} \rho r_1 P \right)^{-M} \left[\frac{1}{\frac{2}{N}\rho r_1 P} \Gamma\left(-M, \frac{1}{\frac{2}{N}\rho r_1 P}\right) + \left(\frac{2}{N}\rho r_2 P \right) \Gamma\left(2 - M, \frac{1}{\frac{2}{N}\rho r_1 P}\right) - \Gamma\left(1 - M, \frac{1}{\frac{2}{N}\rho r_1 P}\right)\left(\frac{r_2}{r_1} - 1\right) \right], \tag{16}$$

where $\Gamma(a, x)$ is the incomplete Gamma function [21]. In the following section, the theoretical results obtained are illustrated by numerical simulations.

IV. Simulations

In Fig. 1, the MSE for two different correlation scenarios and power allocation strategies for a 4×1 system is depicted. From the figure, we observe that correlation has a negative impact on the performance in case equal power allocation is applied. The performance is improved for low to average SNRs, if single-mode beamforming, i.e. allocation the full power to the largest eigenvalue, is applied. Single-mode beamforming is optimal up to -1 dB in the given scenario. For higher SNRs, the optimal power allocation strategy is to support two eigenvalues.

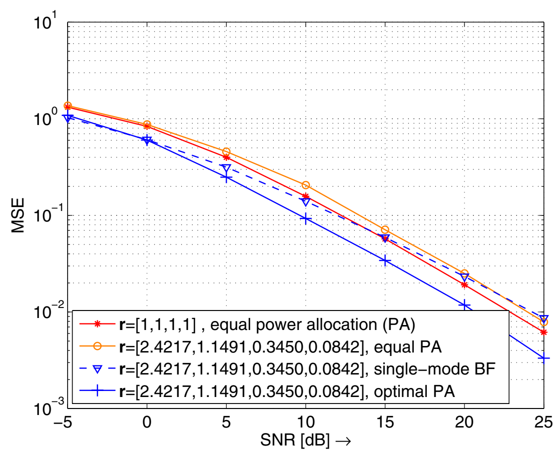

Fig. 1. MSE as a function of the SNR: Uncorrelated case with equal power allocation (PA), and the correlated case with equal PA, single-mode beamforming (BF) and optimal PA.

In Fig. 2, the optimal power allocation for a 4×1 system is illustrated. We observe that at low SNRs up to ρ^*, the complete power is allocated to the largest eigenvalue, i.e. single-mode beamforming is optimal. In addition to the power allocation, γ given in (16) is depicted. In Fig. 2, it can be observed that the null of the function γ corresponds with the occurrence of the power in the direction of the second largest eigenvalue. Furthermore, in the SNR range considered here, the optimal power allocation is still different from equal power allocation along the eigenvalues. Nevertheless, the optimal power allocation converges slowly towards equal power allocation.

V. Conclusion

In conclusion, we studied the optimal linear precoder for a wireless MIMO system with transmitter correlation applying a QSTBC at the transmitter and a linear MMSE detector at the receiver. We showed that it is optimal to transmit into the direction of the eigenvectors of the transmit correlation matrix. The optimal power allocation is characterized using the necessary and sufficient optimality conditions. It turns out that for high SNRs, only the two largest eigenvalues of the correlation matrix are supported, which is a surprising result. We showed that for small SNR values, the optimal precoding matrix has rank one, i.e. single mode beamforming is optimal. A closed form expression characterizing the SNR range for optimal single-mode beamforming is derived.

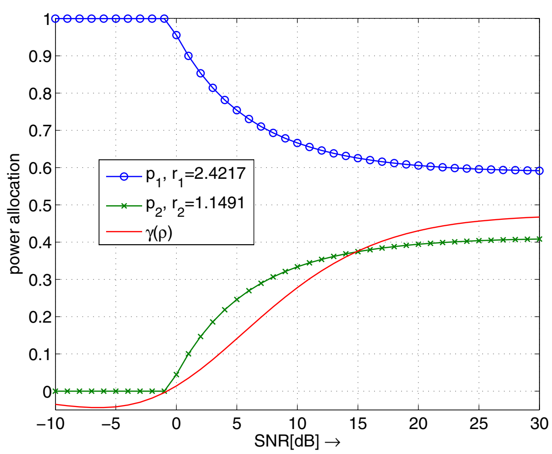

Fig. 2. Optimal power allocation and single-mode beamforming region with $\mathbf{R}_t$ with eigenvalues $r_1 = 2.4217$ and $r_2 = 1.1491$.

Appendix

The matrices $\mathbf{A}_N^k$, $1 \le k \le N$ are idempotent (the eigenvalues consist of ones and zeros) with rank two and $\sum_{k=1}^{N} \mathbf{A}_N^k = \mathbf{I}_N$. They are obtained recursively as follows

$$\mathbf{A}_N^j = \frac{1}{2}\begin{bmatrix} \mathbf{A}^{j'}_{\frac{N}{2}} & -\mathbf{B}^{j'}_{\frac{N}{2}} \\ \mathbf{B}^{j'}_{\frac{N}{2}} & \mathbf{A}^{j'}_{\frac{N}{2}} \end{bmatrix}, \mathbf{A}_N^{j+1} = \frac{1}{2}\begin{bmatrix} \mathbf{A}^{j'}_{\frac{N}{2}} & \mathbf{B}^{j'}_{\frac{N}{2}} \\ -\mathbf{B}^{j'}_{\frac{N}{2}} & \mathbf{A}^{j'}_{\frac{N}{2}} \end{bmatrix}$$

with $j = 1, 3 \ldots, {}^{N}\!/2 - 1$ and $N = 2^n, n \ge 2$, $\mathbf{B}_N^{j'} = i\mathbf{\Theta}_N \mathbf{A}_N^{j'}$ and $j' = \frac{j+1}{2}$. Furthermore $\mathbf{A}_N^k = \mathbf{A}_N^{N/2+k}$, $1 \le k \le {}^{N}\!/2$ and $\mathbf{A}_2^1 = \mathbf{I}_2$. In addition to that, it holds that $\mathbf{\Theta}_N \mathbf{A}_N^j \mathbf{\Theta}_N = \mathbf{A}_N^j$, i.e. $\mathbf{\Theta}_N$ and $\mathbf{A}_N^j$ commutate and $\mathbf{A}_N^j \mathbf{A}_N^k = \mathbf{0}$, $j \ne k$ i.e. are mutually orthogonal.

References

[1] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. on Telecomm. ETT*, vol. 10, no. 6, pp. 585–596, November 1999.

[2] E. Visotsky and U. Madhow, "Space-time precoding with imperfect feedback," *IEEE Trans. on Info. Theory*, vol. 47, no. 6, pp. 2632 – 2639, Sept. 2001.

[3] A. Narula, M.J. Lopez, M.D. Trott, and G.W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE Journal on Selec. Areas in Comm.*, vol. 16, no. 8, pp. 1423 – 1436, October 1998.

[4] S. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems with imperfect feedback," *IEEE Transactions on Wireless Communications*, vol. 3, no. 4, pp. 1165 – 1175, July 2004.

[5] S. Zhou and G.B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback," *IEEE Trans. on Signal Processing*, vol. 50, no. 10, pp. 2599–2613, October 2002.

[6] S. Simon and A. Moustakas, "Optimality of beamforming in multiple transmitter multiple receiver communication systems with partial channel knowledge," *Proc. of DIMACS Conference on Signal Processing for Wireless Transmission*, 2002.

[7] 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. on Wireless Communication*, vol. 3, no. 5, pp. 1543–1553, Sept. 2004.

[8] C. Brunner, W. Utschick, and J. A. Nossek, "Exploiting the short-term and long-term channel properties in space and time: eigenbeamforming concepts for the bs in wcdma," *European Trans. on Telecommunications*, vol. 12, no. 5, pp. 365–378, 2001.

[9] M. Kiessling, "Unifying analysis of ergodic MIMO capacity in correlated rayleigh fading environments," *Europ. Trans. on Telecommunications*, vol. 16, no. 1, pp. 17–35, Jan. 2005.

[10] T. Weber, A. Sklavos, and M. Meurer, "Imperfect channel-state information in MIMO transmission," *IEEE Trans. on Communications*, vol. 54, no. 3, pp. 543–552, March 2006.

[11] G. Jöngren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. on Info. Theory*, vol. 48, no. 3, pp. 611 – 627, March 2002.

[12] L. Liu and H. Jafarkhani, "Application of quasi-orthogonal space-time block codes in beamforming," *IEEE Trans. on Signal Proc.*, vol. 53, no. 1, pp. 54–63, January 2005.

[13] M. Vu and A. Paulraj, "Linear precoding for MIMO wireles correlated channels with non-zero means: K factor analysis, extension to non-orthogonal STBC," *to appear in Proc. of IEEE ICASSP, Philadelphia, PA, USA*, March 2005.

[14] X. Zhang, D. P. Palomar, and B. Ottersten, "Robust design of linear MIMO transceivers under channel uncertainty," *Proc. of IEEE Intern. Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2006.

[15] S.M. Alamouti, "A simple transmitter diversity scheme for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. SAC-16, pp. 1451–1458, October 1998.

[16] A. Sezgin and T.J. Oechtering, "Complete characterization of the equivalent MIMO Channel for quasi-orthogonal space-time codes," *submitted to IEEE Trans. on Info. Theory. Also available at http://www.stanford.edu/∼sezgin*, 2004.

[17] A. Sezgin and T.J. Oechtering, "On the outage probability of quasi-orthogonal space-time codes," *Proc. of IEEE Info. Theory Workshop 2004, San Antonio, TX, USA*, October 2004.

[18] P. Viswanath, V. Anantharam, and D. N. C. Tse, "Optimal sequences, power control and user capacity of synchronous CDMA systems with linear MMSE multiuser receivers," *IEEE Transactions on Information Theory*, vol. 45, no. 6, pp. 1968–1983, October 1999.

[19] A. Sezgin, E.A. Jorswieck, and E. Costa, "LDC in MIMO Rician channels: Optimal transmit strategy with MMSE detection," *IEEE Trans. on Signal Processing, accepted*, 2006.

[20] S. Boyd and L. Vanderberghe, *Convex optimization*, Course reader for EE364 (Stanford) and EE236 (UCLA).

[21] A. Erdelyi, *Higher Transcendental Functions*, vol. 1 and 2, McGraw-Hill, 1953.