ON THE ASYMPTOTIC OPTIMALITY OF BLOCK-DIAGONALIZATION FOR THE MIMO BC UNDER LINEAR FILTERING

Raphael Hunger and Michael Joham

Associate Institute for Signal Processing, Technische Universität München, 80290 Munich, Germany
Telephone: +49 89 289-28508, Fax: +49 89 289-28504, Email: hunger@tum.de

ABSTRACT

In this paper, we address the high SNR regime of the MIMO broadcast channel under linear filtering. For systems where the base station is equipped with more antennas than the user terminals have in sum, we prove that block-diagonalization is the asymptotically optimum transmission strategy for maximizing the sum rate. For this type of transmission strategy, the asymptotically optimum transmit covariance matrix in the broadcast channel is derived in closed form. In addition, we present an expression for the asymptotic sum capacity for an instantaneous channel realization which only depends on this particular channel realization. No precoders or singular-value-decompositions arise as they used to do in hitherto existing sum rate expressions in the multi-antenna terminal case. All results are deduced from the dual multiple access channel in which the optimum transmit covariance matrices can easily be computed. Our recent rate duality for multi-antenna systems where the individual streams of a user are not treated as self-interference allows us then to convert the solution of the dual uplink back to the downlink and to find the optimum transmit and receive filters in the broadcast channel.

1. INTRODUCTION

Block-diagonalization (BD) is a multi-user MIMO channel decomposition technique that orthogonalizes the overall MIMO broadcast channel (BC) of the different users into parallel single-user MIMO channels without inter-user interference. It can thus be thought of as the generalization of zero-forcing channel inversion algorithms for the case when the receivers have multiple antennas. Since capacity achieving transmission strategies do not seem to exist for the multi-user MIMO case under linear filtering, block-diagonalization was introduced as a constraint which leads to closed-form expressions for the resulting sum rate [1]. Depending on the number of available antennas at the base station and the number of antennas at the terminals, standard BD (e.g. [1–3]), BD with user selection (e.g. [4]), or generalized BD where the receive filters are taken into account (e.g. [5]), can be applied. While the complete suppression of inter-user interference, that was heuristically introduced to end up with closed form expressions [1], fails to reach the sum capacity in general, we will show that in the high SNR regime, block-diagonalization is indeed the asymptotically optimum transmission strategy in the multi-user MIMO BC with linear filtering. In order to prove above statement, we utilize the recently introduced rate duality [6] for multi-antenna users under linear filtering and investigate the sum rate maximization problem in the dual uplink multiple access channel (MAC) instead. Despite the simpler structure of the dual MAC (aligned channel and precoder indices), the simplest multi-user setup with single antenna terminals already allows for the presumption that closed form expressions for the sum capacity will remain infeasible even in the multiple access channel irrespective of whether linear or nonlinear filtering is considered. Fortunately, the high signal-to-noise ratio regime is an exception to this deflating circumstance, since there, asymptotic results on the sum capacity have been discovered for dirty paper coding and partly for linear filtering. In [7, 8] for example, the single user point-to-point MIMO case was decomposed into a supremum capacity term, an instantaneous SNR effect term, and an instantaneous capacity degradation term due to the eigenvalue spread. Outage capacity and throughput of a fading point-to-point MIMO system are analyzed in [9], whereas a lower and an upper bound on the sum rate of block-diagonalization was derived in [10]. Nonetheless, precoder-free expressions for the asymptotic sum capacity of a point-to-multipoint broadcast channel for an *instantaneous* channel realization did not exist so far for linear filtering in the multi-antenna terminals case, only sum rate expressions still containing the precoders have been derived yet when the individual users are equipped with several antennas, see [11, Eq. (21)] and [12, Eq. (10)]. Instead, *ergodic* statements can be found in the literature, see for example [11–13]. Therein, the affine approximation of the sum capacity introduced in [14] and elaborately discussed in [15] was applied.

Having derived the asymptotic sum capacity of the dual MAC, we can immediately conclude by means of the duality in [6] that the broadcast channel features the same high SNR sum capacity. Moreover, when we convert the simple solution for the precoders in the dual MAC that asymptotically achieve the sum capacity back to the broadcast channel, it turns out that the resulting BC system features a block-diagonal structure. Hence, we have a formal proof that block-diagonalization is the asymptotically optimum transmission strategy when linear filtering is considered what so far has not been shown, but often been used and investigated, see e.g. [11, 12]. Of course, the MAC to BC conversion also delivers the asymptotically optimum transmit and receive filters.

1.1. Contributions

The main contributions of this paper are summarized in the following list:

1. We prove that block diagonalization is asymptotically optimum for sum rate maximization in the broadcast channel.
2. Optimum precoding and transmit covariance matrices in the broadcast channel are derived by means of our rate duality in [6]. In contrast to existing block-diagonalization algorithms, singular-value-decompositions do not arise in our notation and therefore do not have to be computed.
3. We derive an analytic expression for the sum capacity of the multi-user MIMO BC achievable with linear filtering in the high SNR regime which is the first precoder-free closed form solution depending only on the channel matrices and the antenna configuration.
4. A closed form solution of the covariance matrices in the dual uplink asymptotically achieving this sum capacity.

1.2. Organization

In Section 2, the system model underlying the multi-user scenario is described. The optimum signaling strategy for the sum-rate maximization with linear filtering is derived in Section 3 in the dual multiple access channel and afterwards converted to the broadcast channel in Section 4. Section 5 concludes this paper.

2. SYSTEM MODEL

We consider the communication between an N antenna base station and K multi-antenna terminals, where user k multiplexes B_k data streams over his r_k antennas. For a short notation, we define r as the sum of all antennas at the terminals, i.e., $r=\sum_{k=1}^K r_k$, and b as the total number of transmitted streams, i.e., $b=\sum_{k=1}^K B_k$. Recent results [6] on the rate duality of the BC and the MAC under linear filtering and a user-wise joint stream decoding allow us to investigate the dual MAC (with its simpler structure) instead of the BC. Afterwards, the obtained solutions can conveniently be transformed back to the BC. In this MAC, user k applies a precoding matrix $\boldsymbol{T}_k \in \mathbb{C}^{r_k\times B_k}$ generating his $r_k \times r_k$ transmit covariance matrix $\boldsymbol{Q}_k=\boldsymbol{T}_k\boldsymbol{T}_k^{\mathrm{H}}$. The precoded symbol vector propagates over the channel described by the matrix $\boldsymbol{H}_k \in \mathbb{C}^{N\times r_k}$. At the receiver side, zero-mean noise $\boldsymbol{\eta}\in\mathbb{C}^N$ with identity covariance matrix is added and the receive filter for user k is denoted by $\boldsymbol{G}_k \in \mathbb{C}^{B_k\times N}$. Due to the reversed signal flow in the BC, we characterize the transmission from the base station to terminal k by the Hermitian channel $\boldsymbol{H}_k^{\mathrm{H}}$ in the BC, and the precoder dedicated to the B_k streams of user k is denoted by $\boldsymbol{P}_k \in \mathbb{C}^{N\times B_k}$. Throughout this paper, we assume that the base station has at least as many antennas as the terminals have in sum, i.e., $N\geq r$.

3. OPTIMUM SIGNALING IN THE DUAL MAC

Introducing the composite channel matrix $\boldsymbol{H}$ and the composite block-diagonal precoder matrix $\boldsymbol{T}$ of all K users via

$$\begin{aligned}\boldsymbol{H}&=[\boldsymbol{H}_1,\ldots,\boldsymbol{H}_K]\in\mathbb{C}^{N\times r},\\ \boldsymbol{T}&=\operatorname{blockdiag}\{\boldsymbol{T}_k\}_{k=1}^K\in\mathbb{C}^{r\times b},\end{aligned}\tag{1}$$

the rate of user k seeing interference from all other users can be expressed as (see [6])

$$\begin{aligned}R_k&=\log_2\left|\mathbf{I}_N+\Big(\mathbf{I}_N+\sum_{\ell\neq k}\boldsymbol{H}_\ell\boldsymbol{Q}_\ell\boldsymbol{H}_\ell^{\mathrm{H}}\Big)^{-1}\boldsymbol{H}_k\boldsymbol{Q}_k\boldsymbol{H}_k^{\mathrm{H}}\right|\\ &=-\log_2\left|\mathbf{I}_{B_k}-\boldsymbol{T}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{X}^{-1}\boldsymbol{H}_k\boldsymbol{T}_k\right|,\end{aligned}\tag{2}$$

where the substitution $\boldsymbol{X}$ reads as

$$\boldsymbol{X}=\mathbf{I}_N+\sum_{\ell=1}^K\boldsymbol{H}_\ell\boldsymbol{Q}_\ell\boldsymbol{H}_\ell^{\mathrm{H}}=\mathbf{I}_N+\boldsymbol{H}\boldsymbol{T}\boldsymbol{T}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}.$$

Note that (2) in general requires all streams of a single user to be decoded jointly. Reformulating the rate expression (2), we get

$$\begin{aligned}R_k&=-\log_2\left|\boldsymbol{E}_k^{\mathrm{T}}\big(\mathbf{I}_b-\boldsymbol{T}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{X}^{-1}\boldsymbol{H}\boldsymbol{T}\big)\boldsymbol{E}_k\right|\\ &=-\log_2\left|\boldsymbol{E}_k^{\mathrm{T}}\big(\mathbf{I}_b+\boldsymbol{T}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{T}\big)^{-1}\boldsymbol{E}_k\right|,\end{aligned}\tag{3}$$

where the transposed kth block unit matrix is defined via

$$\boldsymbol{E}_k^{\mathrm{T}}=[\mathbf{0},\ldots,\mathbf{0},\mathbf{I}_{B_k},\mathbf{0},\ldots,\mathbf{0}]\in\{0,1\}^{B_k\times b}$$

with the identity matrix at the kth block and the ith block with $i\neq k$ corresponds to the zero matrix of dimension $B_k\times B_i$. Due to the assumption that the base station has more antennas than the terminals have in sum, all r streams can be activated leading to square precoders $\boldsymbol{T}_k$ with $B_k=r_k\ \forall k$. Raising P_{Tx}, all r streams become active, $\boldsymbol{T}$ becomes full rank, and all eigenvalues of $\boldsymbol{T}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{T}$ become much larger than one. In the asymptotic limit, we obtain

$$\begin{aligned}R_k&\cong-\log_2\left|\boldsymbol{T}_k^{-1}\boldsymbol{E}_k^{\mathrm{T}}\big(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\big)^{-1}\boldsymbol{E}_k\boldsymbol{T}_k^{-\mathrm{H}}\right|\\ &=\ \log_2\left|\boldsymbol{Q}_k\right|-\log_2\left|\boldsymbol{E}_k^{\mathrm{T}}\big(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\big)^{-1}\boldsymbol{E}_k\right|,\end{aligned}\tag{4}$$

since $\boldsymbol{E}_k^{\mathrm{T}}\boldsymbol{T}^{-1}=\boldsymbol{T}_k^{-1}\boldsymbol{E}_k^{\mathrm{T}}$. The notation $x\cong y$ means that the difference $x-y$ vanishes when the sum power P_{Tx} goes to infinity. Interestingly, the rate of user k depends only on the determinant of his own transmit covariance matrix $\boldsymbol{Q}_k$, and not on the covariance matrices of the other users! Consequently, the eigenbases of all transmit covariance matrices do not influence the rates of the users, only the powers of the eigenmodes are relevant. Let the eigenvalue decomposition of $\boldsymbol{Q}_k$ read as $\boldsymbol{Q}_k = \boldsymbol{V}_k \boldsymbol{\Lambda}_k \boldsymbol{V}_k^{\mathrm{H}}$ with unitary $\boldsymbol{V}_k$ and the diagonal nonnegative power allocation $\boldsymbol{\Lambda}_k$. Due to the determinant operator, $\boldsymbol{V}_k$ can be chosen arbitrarily and therefore, we set $\boldsymbol{V}_k = \mathbf{I}_{r_k}\ \forall k$ without loss of generality. Let the power allocation matrix be composed by the entries $\boldsymbol{\Lambda}_k = \mathrm{diag}\{\lambda_k^{(i)}\}_{i=1}^{r_k}$. As only the traces of the covariance matrices are involved in the sum power constraint $\sum_{k=1}^{K} \mathrm{tr}(\boldsymbol{Q}_k) \leq P_{\mathrm{Tx}}$, the determinant $|\boldsymbol{Q}_k| = |\boldsymbol{\Lambda}_k|$ is maximized by setting

$$\lambda_k^{(1)} = \ldots = \lambda_k^{(r_k)} := \lambda_k, \tag{5}$$

i.e., by evenly distributing the power allocated to that user onto his individual modes, so $\boldsymbol{Q}_k = \lambda_k \mathbf{I}_{r_k}\ \forall k$. Inserting (5) into the asymptotic rate equation of user k in (4) leads to the asymptotic sum rate expression

$$\sum_{k=1}^{K} R_k \cong \sum_{k=1}^{K} \left(r_k \log_2 \lambda_k - \log_2 \left| \boldsymbol{E}_k^{\mathrm{T}} \left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{H} \right)^{-1} \boldsymbol{E}_k \right| \right). \tag{6}$$

Subject to the sum power constraint $\sum_{k=1}^{K} r_k \lambda_k \leq P_{\mathrm{Tx}}$, the sum rate in (6) is maximized for

$$\lambda_k = \frac{P_{\mathrm{Tx}}}{r}, \tag{7}$$

so power is evenly allocated to the users (similar to the single-antenna case proven in [11]), and every user evenly distributes his fraction of power onto his modes. Finally, the asymptotic sum rate at high SNR reads as

$$\sum_{k=1}^{K} R_k \cong r \log_2 P_{\mathrm{Tx}} - r \log_2 r - \sum_{k=1}^{K} \log_2 \left| \boldsymbol{E}_k^{\mathrm{T}} \left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{H} \right)^{-1} \boldsymbol{E}_k \right|, \tag{8}$$

and is interestingly achieved with the transmit covariance matrices $\boldsymbol{Q}_k = P_{\mathrm{Tx}}/r \cdot \mathbf{I}_{r_k}\ \forall k$. Using (8), we are able to quantify the asymptotic sum rate that can be achieved by means of linear filtering for every single channel realization and antenna/user profile in terms of the transmit power P_{Tx} and the channel itself as long as $N \geq r$ holds. Note that no precoders arise in (8) in contrast to [12, Eq. (10)] and [11, Eq. (21)]. In principle, the ergodic rate can be obtained by averaging corresponding to *any* distribution of the channel. In [12], results on the *ergodic* rate were presented for the specific case of Rayleigh fading only, where the channel entries of $\boldsymbol{H}_1, \ldots, \boldsymbol{H}_K$ all have the same distribution. More complicated fading models cannot be captured due to this restricting assumption. Moreover, the instantaneous rate expression is given by means of bases representing null spaces of shortened channel matrices taken from [1] and not as a function of the channel purely as we do in (8). Concerning the asymptotic rate expression, we have created a smooth transition from the r single-antenna-users system configuration in [13] where no cooperation exists between the antenna elements at the terminals, to the single-user point-to-point MIMO link where all r antennas fully cooperate, see [16] for example. In between, we can now specify any antenna/user profile we want and compute the feasible rate in the asymptotic limit under linear filtering.

4. OPTIMUM SIGNALING IN THE BC

Using our recent rate duality in [6], we can convert the simple solution for the covariance matrices $\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_K$ in the dual MAC to covariance matrices $\boldsymbol{S}_1, \ldots, \boldsymbol{S}_K$ in the BC, where the Hermitian channels are applied. Since this duality explicitly uses the receive filters in the MAC as scaled transmit matrices in the BC, we first compute the MMSE receivers in the dual MAC, as they are optimum and generate sufficient statistics. This MMSE receive filter $\boldsymbol{G}_k$ for user k in the dual MAC reads as

$$\boldsymbol{G}_k = \boldsymbol{E}_k^{\mathrm{T}} \boldsymbol{T}^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}} \left(\mathbf{I}_N + \boldsymbol{H} \boldsymbol{T} \boldsymbol{T}^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}} \right)^{-1}.$$

Using asymptotically optimum precoders $\boldsymbol{T}_k = \sqrt{P_{\mathrm{Tx}}/r}\,\mathbf{I}_{r_k}$, above expression asymptotically converges to

$$\boldsymbol{G}_k \cong \sqrt{r/P_{\mathrm{Tx}}} \cdot \boldsymbol{E}_k^{\mathrm{T}} \left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{H} \right)^{-1} \boldsymbol{H}^{\mathrm{H}}. \tag{9}$$

Let $\boldsymbol{P}_k$ denote the precoder of user k in the BC, then the ith column $\boldsymbol{p}_{k,i}$ of $\boldsymbol{P}_k$ follows from the conjugate ith row $\boldsymbol{g}_{k,i}'^{\mathrm{T}}$ of the matrix $\boldsymbol{G}_k' = \boldsymbol{W}_k^{\mathrm{H}} \boldsymbol{G}_k$ via (see [6])

$$\boldsymbol{p}_{k,i} = \alpha_{k,i} \boldsymbol{g}_{k,i}'^{*} = \frac{\alpha_{k,i}}{\sqrt{P_{\mathrm{Tx}}/r}} \cdot \boldsymbol{H} \left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{H} \right)^{-1} \boldsymbol{E}_k \boldsymbol{W}_k \boldsymbol{e}_i, \tag{10}$$

where the scaling factor $\alpha_{k,i}$ is obtained by the duality transformation and $\boldsymbol{W}_k$ is a unitary decorrelation matrix. This matrix $\boldsymbol{W}_k$ finally ensures that the individual streams of every user can be decoded separately instead of jointly without having to face the rate loss that usually has to be taken into account when streams are decoded separately rather than jointly, see [6]. Since we convert only the asymptotically optimum transmit precoders and receive filters, the duality transformation from the MAC to the BC in [6] drastically simplifies and can even be computed in closed form. In particular, the matrices $\boldsymbol{M}_{a,b}$ in [6, Eq. (23)] vanish for $a \neq b$ yielding a *diagonal* matrix $\boldsymbol{M}$ from which the scaling factors are derived. Those scalars now compute to

$$\alpha_{k,i} = \frac{\sqrt{P_{\mathrm{Tx}}/r}}{\|\boldsymbol{g}_{k,i}'\|_2}. \tag{11}$$

In combination with (10), the ith column of the precoder associated to user k reads as

$$\boldsymbol{p}_{k,i} = \sqrt{P_{\mathrm{Tx}}/r} \cdot \frac{\boldsymbol{H} \left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{H} \right)^{-1} \boldsymbol{E}_k \boldsymbol{W}_k \boldsymbol{e}_i}{\left\| \boldsymbol{H} \left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{H} \right)^{-1} \boldsymbol{E}_k \boldsymbol{W}_k \boldsymbol{e}_i \right\|_2},$$

generating the precoder matrix

$$\boldsymbol{P}_k = \sqrt{P_{\mathrm{Tx}}/r} \cdot \boldsymbol{H}\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\right)^{-1}\boldsymbol{E}_k\boldsymbol{W}_k\boldsymbol{D}_k^{-1}, \qquad (12)$$

where the ith diagonal element of the diagonal matrix $\boldsymbol{D}_k$ is

$$[\boldsymbol{D}_k]_{i,i} = \sqrt{\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{W}_k^{\mathrm{H}}\boldsymbol{E}_k^{\mathrm{T}}\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\right)^{-1}\boldsymbol{E}_k\boldsymbol{W}_k\boldsymbol{e}_i}. \qquad (13)$$

We can immediately see, that the precoding filters in (12) lead to a block diagonalization of the transmission, since $\boldsymbol{H}_\ell^{\mathrm{H}}\boldsymbol{P}_k = \boldsymbol{0}$ holds for $k \neq \ell$. Next, the decorrelation matrix $\boldsymbol{W}_k$ which enables the duality is usually chosen as the eigenbasis of $\boldsymbol{G}_k\boldsymbol{H}_k\boldsymbol{T}_k \cong \mathbf{I}_{r_k}$, which asymptotically coincides with the identity matrix due to (9). Since all eigenvalues are identical to one, the decorrelation matrices $\boldsymbol{W}_k$ are not given a priori, but can easily be computed such that the BC features the same sum rate as the dual MAC. By means of (12) and the block diagonalization property of the precoders, we obtain for user k's receive signal

$$\boldsymbol{y}_k = \boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{P}_k\boldsymbol{s}_k + \boldsymbol{\eta}_k = \sqrt{P_{\mathrm{Tx}}/r} \cdot \boldsymbol{W}_k\boldsymbol{D}_k^{-1}\boldsymbol{s}_k + \boldsymbol{\eta}_k, \quad (14)$$

where $\boldsymbol{\eta}_k \in \mathbb{C}^{r_k}$ is the noise and $\boldsymbol{s}_k$ the symbol vector of user k both having an identity covariance matrix. From (14), the rate of user k achieved in the BC reads as

$$R_k = \log_2\left|\mathbf{I}_{r_k} + P_{\mathrm{Tx}}/r \cdot \boldsymbol{W}_k\boldsymbol{D}_k^{-2}\boldsymbol{W}_k^{\mathrm{H}}\right|, \qquad (15)$$

which asymptotically converges to

$$R_k \cong r_k \log_2 P_{\mathrm{Tx}} - r_k \log_2 r - \log_2|\boldsymbol{D}_k^2|. \qquad (16)$$

For the asymptotic result in (16), the identity matrix $\mathbf{I}_{r_k}$ in (15) was omitted, so the true rate R_k will always converge to the right hand side from (16) *from above*. It remains to minimize the determinant of $\boldsymbol{D}_k^2$ by the choice of the unitary decorrelation filter $\boldsymbol{W}_k$, such that the asymptotic limit of R_k is maximized, see (16). From the definition of $\boldsymbol{D}_k$ in (13), *Hadamard's* inequality [17] tells us that $\boldsymbol{W}_k$ has to be chosen as the unitary eigenbasis of $\boldsymbol{E}_k^{\mathrm{T}}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{-1}\boldsymbol{E}_k$. Choosing $\boldsymbol{W}_k$ this way, $\boldsymbol{D}_k^2$ contains the eigenvalues of $\boldsymbol{E}_k^{\mathrm{T}}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{-1}\boldsymbol{E}_k$, and therefore, the elements of $\boldsymbol{D}_k^2$ are as different as possible since the eigenvalues of any positive definite matrix majorize its diagonal elements according to *Schur's* theorem [18]:

$$\boldsymbol{E}_k^{\mathrm{T}}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{-1}\boldsymbol{E}_k = \boldsymbol{W}_k\boldsymbol{D}_k^2\boldsymbol{W}_k^{\mathrm{H}}. \qquad (17)$$

The optimum transmit covariance matrix $\boldsymbol{S}_k = \boldsymbol{P}_k\boldsymbol{P}_k^{\mathrm{H}}$ of user k in the BC which asymptotically achieves the same sum rate as the dual MAC counterpart reads by means of (12) and (17) as

$$\boldsymbol{S}_k = \frac{P_{\mathrm{Tx}}}{r} \cdot \boldsymbol{H}^{+\,\mathrm{H}}\boldsymbol{E}_k\left(\boldsymbol{E}_k^{\mathrm{T}}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{-1}\boldsymbol{E}_k\right)^{-1}\boldsymbol{E}_k^{\mathrm{T}}\boldsymbol{H}^{+}. \quad (18)$$

In (18), we make use of the channel pseudo-inverse $\boldsymbol{H}^+$ which is defined via $\boldsymbol{H}^+ = (\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{H}}$. Note that r_k eigenvalues of $\boldsymbol{S}_k$ are P_{Tx}/r whereas the remaining $N - r_k$ ones are zero. Thus, $\boldsymbol{S}_k$ is a weighted orthogonal projector. Furthermore, $\mathrm{tr}(\boldsymbol{S}_k) = P_{\mathrm{Tx}} \cdot r_k/r\ \forall k$, so the power is uniformly allocated to the individual users in the broadcast channel as well. Comparing (18) with the simple solution of the transmit covariance matrix $\boldsymbol{Q}_k = P_{\mathrm{Tx}}/r \cdot \mathbf{I}_{r_k}$ in the dual MAC, it becomes obvious that the optimum covariance matrices are much more difficult to find directly in the BC without using the rate duality, than in the dual MAC. Plugging the optimum $\boldsymbol{D}_k^2$ containing the eigenvalues of $\boldsymbol{E}_k^{\mathrm{T}}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{-1}\boldsymbol{E}_k$ into (16) finally yields

$$R_k \cong r_k \log_2 P_{\mathrm{Tx}} - r_k \log_2 r - \log_2\left|\boldsymbol{E}_k^{\mathrm{T}}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{-1}\boldsymbol{E}_k\right|.$$

Hence, the maximum sum rate (8) in the dual MAC is also achieved in the BC. While the rate expression in (15) at first glance seems to require the joint detection of all the streams belonging to user k, the application of the simple matched filter receiver

$$\boldsymbol{B}_k = \boldsymbol{D}_k^{-1}\boldsymbol{W}_k^{\mathrm{H}}$$

to the receive signal $\boldsymbol{y}_k$ of every user k completely decorrelates the individual streams of user k. As a consequence, the individual streams of every user can be detected separately without any loss in rate, and the SINR of the ith stream belonging to user k reads as

$$\mathrm{SINR}_{k,i} = \frac{P_{\mathrm{Tx}}}{r} \cdot \frac{1}{[\boldsymbol{D}_k^2]_{i,i}}$$

entailing the rate $R_{k,i} = \log_2(1+\mathrm{SINR}_{k,i})$. Summing up the rates of all r_k streams of user k, we obtain

$$R_k = \sum_{i=1}^{r_k} R_{k,i} = \log_2\prod_{i=1}^{r_k}\left(1 + \frac{P_{\mathrm{Tx}}}{r} \cdot \frac{1}{[\boldsymbol{D}_k^2]_{i,i}}\right)$$

in accordance to (15), and the rate R_k of user k asymptotically converges to

$$\begin{aligned} R_k &\cong r_k \log_2 P_{\mathrm{Tx}} - r_k \log_2 r - \log_2|\boldsymbol{D}_k^2| \\ &= r_k \log_2 P_{\mathrm{Tx}} - r_k \log_2 r - \log_2\left|\boldsymbol{E}_k^{\mathrm{T}}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{-1}\boldsymbol{E}_k\right|. \end{aligned}$$

Therefore, stream-wise detection achieves the same user rates for our special choice of the decorrelation matrices $\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K$. While the transmission chain consisting of the precoder and the channel leads to a block-diagonalization, the inclusion of the matched filter receiver even leads to a total diagonalization of the transmission chain. Therefore, a joint diagonalization of the transmission chain by both the transmit filter and the matched filter receiver with appropriately chosen decorrelation matrices is asymptotically optimum in the same way as the block-diagonalization of the channel by the sender alone is.

5. CONCLUSION

We have shown that block-diagonalization is the asymptotically optimum transmission strategy in the broadcast channel and derived the asymptotic sum capacity when linear filtering is applied instead of dirty paper coding. Starting in the dual multiple access channel, we have found a very simple closed form solution for the optimum transmit covariance matrix which has then been converted back to the downlink by our recent rate duality. To the best of our knowledge, we have presented the first precoder-free asymptotic sum capacity expression of linear filtering which is only a function of the individual users' channel matrices, the available transmit power, and the antenna profile of the users.

6. REFERENCES

[1] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-Forcing Methods for Downlink Spatial Multiplexing in Multiuser MIMO Channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461–471, February 2004.

[2] L.-U. Choi and R. D. Murch, "A Transmit Preprocessing Technique for Multiuser MIMO Systems Using a Decomposition Approach," *IEEE Transactions on Wireless Communications*, vol. 3, no. 1, pp. 20–24, January 2004.

[3] P. Komulainen, M. Latva-aho, and M. Juntti, "Block Diagonalization for Mulituser MIMO TDD Downlink and Uplink in Time-Varying Channel," in *Proc. ITG/IEEE WSA 2008*, February 2008, pp. 74–81.

[4] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath, and B. L. Evans, "Low Complexity User Selection Algorithms for Multiuser MIMO Systems With Block Diagonalization," *IEEE Trans. on Signal Processing*, vol. 54, no. 9, pp. 3658–3663, September 2006.

[5] Z. Pan, K.-K. Wong, and T.-S. Ng, "Generalized Multiuser Orthogonal Space Division Multiplexing," *IEEE Trans. on Wireless Communications*, vol. 3, no. 6, pp. 1969–1973, November 2004.

[6] R. Hunger and M. Joham, "A General Rate Duality of the MIMO Multiple Access Channel and the MIMO Broadcast Channel," in *Global Telecommunications Conference (Globecom '08)*, November 2008.

[7] J. Salo, P. Suvikunnas, H. M. El-Sallabi, and P. Vainikainen, "Some results on MIMO mutual information: the high SNR case," in *Global Telecommunications Conference (Globecom '04)*, vol. 2, December 2004, pp. 943–947.

[8] J. Salo and P. Suvikunnas and H. M. El-Sallabi and P. Vainikainen, "Some Insights into MIMO Mutual Information: The High SNR Case," *IEEE Transactions on Wireless Communications*, vol. 5, no. 11, pp. 2997–3001, November 2006.

[9] N. Prasad and M. K. Varanasi, "Throughput analysis for MIMO systems in the high SNR regime," in *International Symposium on Information Theory (ISIT)*, July 2006, pp. 1954–1958.

[10] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath, and B. L. Evans, "Sum Capacity of Multiuser MIMO Broadcast Channels with Block Diagonalization," in *International Symposium on Information Theory (ISIT)*, July 2006, pp. 886–890.

[11] J. Lee and N. Jindal, "Dirty Paper Coding vs. Linear Precoding for MIMO Broadcast Channels," in *40th Asilomar Conference on Signals, Systems, and Computers (Asilomar 2006)*, October 2006, pp. 779–783.

[12] ——, "High SNR Analysis for MIMO Broadcast Channels: Dirty Paper Coding Versus Linear Precoding," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4787–4792, December 2007.

[13] N. Jindal, "High SNR Analysis of MIMO Broadcast Channels," in *International Symposium on Information Theory (ISIT 2005)*, September 2005, pp. 2310–2314.

[14] S. Shamai and S. Verdú, "The Impact of Frequency-Flat Fading on the Spectral Efficiency of CDMA," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1302–1327, May 2001.

[15] A. Lozano, A. M. Tulino, and S. Verdú, "High-SNR Power Offset in Multiantenna Communication," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4134–4151, December 2005.

[16] E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–596, November/December 1999.

[17] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1990.

[18] A.W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, ser. Mathematics in Science and Engineering, R. Bellman, Ed. Academic Press, 1979, vol. 143.