Carrier-Cooperative Zero-Forcing for Power Minimization in Parallel MIMO Broadcast Channels

Stephan Herrmann, Christoph Hellings, and Wolfgang Utschick

46th Asilomar Conference on Signals, Systems and Computers
(ACSSC 2012)

Pacific Grove, CA, USA, 4th - 7th November, 2012

Technische Universität München
Associate Institute for Signal Processing
http://www.msv.ei.tum.de

Carrier-Cooperative Zero-Forcing for Power Minimization in Parallel MIMO Broadcast Channels

Stephan Herrmann, Christoph Hellings, and Wolfgang Utschick
Associate Institute for Signal Processing, Technische Universität München, 80290 Munich, Germany
Telephone: +49 89 289-28516, Fax: +49 89 289-28504, Email: st.herrmann@mytum.de, {hellings,utschick}@tum.de

***Abstract*—We consider the power minimization problem with per-user minimum rate constraints for parallel multiple-input multiple-output (MIMO) broadcast channels employing zero-forcing beamforming. Recent results have shown that spreading data streams across several carriers—so called carrier-cooperative (CC) transmission—can lead to a reduction of the sum transmit power in such a scenario. However, using state-of-the-art power minimization algorithms based on zero-forcing, only carrier-noncooperative (CN) solutions can be obtained. In this paper, we derive a novel algorithm that is capable of finding CC transmit strategies and can achieve a significant decrease in sum transmit power compared to a conventional zero-forcing power minimization method. The key point of the algorithm is that it combines greedy allocation of data streams, which is a popular technique to optimize zero-forcing strategies, with a gradient-based update of the filter vectors, which is a way to ensure that CC solutions can be obtained. Numerical simulations show that the advantage of the new algorithm is most pronounced in an environment where users have spectrally similar channels.**

I. Introduction

In a communication system transmitting over a set of orthogonal resources (e.g., carriers), we can distinguish between *carrier-cooperative* (CC) and *carrier-noncooperative* (CN) transmission schemes [1], [2]. In the case of carrier-cooperative transmission, the encoding of the data streams is performed jointly on all carriers. This means that data streams can be spread across carriers and be recovered at the receiver by combining the signals received on all carriers. Contrarily, carrier-noncooperative transmission schemes perform an allocation of subsets of the data streams to certain carriers with separate encoding and decoding on each carrier.

Although CN transmission is known to be capacity achieving in parallel MIMO broadcast channels (e.g., [3]), it was shown to be potentially suboptimal if the transmit strategy is restricted to linear transceivers [4]. In our companion work [5], we show that this potential suboptimality also holds in the case of power minimization using linear zero-forcing (ZF) beamforming without time-sharing, i.e., the sum transmit power necessary to fulfill a set of quality of service (QoS) constraints (expressed in terms of minimum per-user rates) using carrier-noncooperative zero-forcing might be higher than the power needed with carrier-cooperative zero-forcing. Since zero-forcing beamforming without time-sharing is much easier to implement than the highly complex capacity achieving dirty paper coding (DPC, e.g., [3]) and, therefore, much more appropriate for use in practical systems, the result from [5] motivates further research on CC ZF.

The problem of power minimization with minimum rate constraints and zero-forcing constraints was studied for parallel MIMO broadcast channels in [6] and for parallel multiple-input single-output (MISO) broadcast channels in [7], [8]. However, the methods proposed in these papers treat all carriers separately, i.e., they are CN schemes. As exposed in [2], CC transmission schemes could be developed by applying algorithms designed for MIMO broadcast channels to an equivalent single-carrier channel (cf. Section II, where the system model is introduced). However, doing so with the algorithm from [6] still leads to carrier-noncooperative solutions because this algorithm computes filter vectors based on singular value decompositions of the channel matrices [2]. Therefore, the efficient design of linear CC ZF transmit strategies without time-sharing has been an open problem.

In this paper, we propose an efficient algorithm to find suboptimal solutions to this problem, and we show in numerical simulations (cf. Section V) that conventional CN strategies can be outperformed by these solutions. The key point of the algorithm is that it breaks with the paradigm that no iterative algorithms are needed to optimize zero-forcing strategies (e.g., [6]–[8] for power minimization and [9]–[11] for sum rate maximization). As was shown in [2], iterative updates of the transmit and receive filters are an appropriate way to optimize CC strategies as long as a CC initialization is chosen. To apply this result to the case of zero-forcing, where a stream selection is inevitable, we combine an iterative gradient-based algorithm (cf. Section IV) with the classical approach of successive stream allocation used in [6]–[11] (cf. Section III).

Notation: We write $\mathbf{I}_L$ for the identity matrix of size L, $\mathbf{0}$ for the zero vector, $\mathbf{1}$ for the all-ones vector, and $\boldsymbol{e}_i$ for the ith canonical unit vector (1 as ith entry and 0 elsewhere). We use $\bullet^{\mathrm{T}}$ to denote the transpose, $\bullet^*$ for the conjugate, $\bullet^{\mathrm{H}}$ for the conjugate transpose, $|\bullet|$ for the absolute value as well as for the cardinality of a set, and $\|\bullet\|$ for the Euclidean norm. We write $\forall(k,s)$ for $\forall k \in \{1,\ldots,K\}$, $\forall s \in \{1,\ldots,S_k\}$.

II. System Model and Problem Formulation

We consider a MIMO broadcast system with C orthogonal carriers, M transmit antennas, and K users, where user k has N_k receive antennas. All channels are assumed to be frequency-flat within a carrier c, and perfect channel knowledge is assumed. The channel between the base station and receiver k on carrier c is characterized by a channel matrix $\boldsymbol{H}_k^{(c)} \in \mathbb{C}^{N_k \times M}$ and a noise covariance matrix $\boldsymbol{C}_k^{(c)} \in$

$\mathbb{C}^{N_k \times N_k}$ describing the additive circularly symmetric Gaussian noise $\boldsymbol{\eta}_k^{(c)} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{C}_k^{(c)})$ at the receiver. We assume $\boldsymbol{C}_k^{(c)} = \mathbf{I}_{N_k}$ without loss of generality since in any other case, a whitening filter $\boldsymbol{C}_k^{(c),-\frac{1}{2}}$ could be applied and considered as part of a whitened channel $\boldsymbol{H}_{k,\text{eff}}^{(c)} = \boldsymbol{C}_k^{(c),-\frac{1}{2}} \boldsymbol{H}_k^{(c)}$.

The data for user k consists of $S_k \leq \min\{N_k C, MC\}$ streams of circularly symmetric Gaussian data symbols, i.e., $\boldsymbol{x}_k = [x_{k,1}, \ldots, x_{k,S_k}]^{\mathrm{T}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{S_k})$. These data streams are transmitted by the base station using linear beamforming in a way that data streams may be spread across several carriers, i.e., part of the signal $x_{k,s}$ might be found on any carrier c. Such a carrier-cooperative transmission can be written as

$$\begin{bmatrix} \boldsymbol{y}_k^{(1)} \\ \vdots \\ \boldsymbol{y}_k^{(C)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}_k^{(1)} & & \\ & \ddots & \\ & & \boldsymbol{H}_k^{(C)} \end{bmatrix} \sum_{j=1}^{K} \tilde{\boldsymbol{T}}_j \boldsymbol{x}_j + \begin{bmatrix} \boldsymbol{\eta}_k^{(1)} \\ \vdots \\ \boldsymbol{\eta}_k^{(C)} \end{bmatrix} \quad (1)$$

where the beamforming matrices $\tilde{\boldsymbol{T}}_j \in \mathbb{C}^{MC \times S_j}$ can be decomposed into

$$\tilde{\boldsymbol{T}}_j = \begin{bmatrix} \boldsymbol{t}_{j,1} & \ldots & \boldsymbol{t}_{j,S_j} \end{bmatrix} \operatorname{diag}(\sqrt{p_{j,1}}, \ldots, \sqrt{p_{j,S_j}}) \quad (2)$$

with $\|\boldsymbol{t}_{j,s}\| = 1$ and $p_{j,s} \geq 0$ to get a stream-wise perspective of power allocation and beamforming.

Note that (1) is equivalent to data transmission in a single-carrier broadcast channel (cf., e.g., [1], [2])

$$\boldsymbol{y}_k = \boldsymbol{H}_k \sum_{j=1}^{K} \tilde{\boldsymbol{T}}_j \boldsymbol{x}_j + \boldsymbol{\eta}_k \quad (3)$$

with the equivalent single-carrier channel matrix

$$\boldsymbol{H}_k = \operatorname{blockdiag}\left(\boldsymbol{H}_k^{(1)}, \ldots, \boldsymbol{H}_k^{(C)}\right) \in \mathbb{C}^{N_k C \times MC} \quad (4)$$

and the noise vector

$$\boldsymbol{\eta}_k = \begin{bmatrix} \boldsymbol{\eta}_k^{(1),\mathrm{T}} & \ldots & \boldsymbol{\eta}_k^{(C),\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \in \mathbb{C}^{N_k C}. \quad (5)$$

Applying a linear receive filter $\boldsymbol{G}_k^{\mathrm{H}} \in \mathbb{C}^{S_k \times N_k C}$ to obtain an estimate $\hat{\boldsymbol{x}}_k = \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{y}_k$, the rate of user k reads as [12]

$$r_k = \log_2 \det(\mathbf{I}_{S_k} + \tilde{\boldsymbol{T}}_k^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{G}_k \boldsymbol{R}_k^{-1} \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k \tilde{\boldsymbol{T}}_k) \quad (6)$$

with

$$\boldsymbol{R}_k = \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{G}_k + \sum_{j \neq k} \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k \tilde{\boldsymbol{T}}_j \tilde{\boldsymbol{T}}_j^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{G}_k. \quad (7)$$

In CC transmit schemes, the beamformers $\boldsymbol{T}_k$ and the receive filters $\boldsymbol{G}_k^{\mathrm{H}}$ do not have any special structure. However, when applying existing ZF-based optimization methods (e.g., [6], [9], [10]) to the equivalent single-carrier broadcast channel (3), we obtain solutions that can be brought to the form

$$\tilde{\boldsymbol{T}}_k = \operatorname{blockdiag}\left(\tilde{\boldsymbol{T}}_k^{(1)}, \ldots, \tilde{\boldsymbol{T}}_k^{(C)}\right) \quad (8)$$

$$\boldsymbol{G}_k^{\mathrm{H}} = \operatorname{blockdiag}\left(\boldsymbol{G}_k^{(1),\mathrm{H}}, \ldots, \boldsymbol{G}_k^{(C),\mathrm{H}}\right) \quad (9)$$

matching the block structure of the channel matrices [2]. The reason for this is that the abovementioned algorithms choose the receive filters based on singular value decompositions of the channel matrices. As the transmit filters in (8) allow us to partition the data symbol vectors as

$$\boldsymbol{x}_k = \begin{bmatrix} \boldsymbol{x}_k^{(1),\mathrm{T}} & \ldots & \boldsymbol{x}_k^{(C),\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \quad (10)$$

where $\boldsymbol{x}_k^{(c)}$ contains symbols which are transmitted only over carrier c, these filters correspond to a CN strategy.

As mentioned above, we consider the case of zero-forcing, i.e., an estimate $\hat{x}_{k,s}$ of the sth stream of user k may not contain interference of any other data stream, including streams of the same user. This is expressed by the constraints

$$\boldsymbol{g}_{k,s}^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{t}_{\ell,t} = 0 \quad \forall (k,s) \neq (\ell, t) \quad (11)$$

where $\boldsymbol{g}_{k,s}^{\mathrm{H}}$ is the sth row of $\boldsymbol{G}_k^{\mathrm{H}}$. As a consequence of the zero-forcing constraints, we obtain $S_{\text{tot}} = \sum_{k=1}^{K} S_k$ independent subchannels since no interference of users $j \neq k$ is present in (7) and since the per-user rate (6) decomposes into

$$r_k = \sum_{s=1}^{S_k} r_{k,s} \quad (12)$$

where the per-stream rates

$$r_{k,s} = \log_2 (1 + \gamma_{k,s} p_{k,s}) \quad (13)$$

can be expressed by means of the subchannel gains

$$\gamma_{k,s} = \frac{\left| \boldsymbol{g}_{k,s}^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{t}_{k,s} \right|^2}{\boldsymbol{g}_{k,s}^{\mathrm{H}} \boldsymbol{g}_{k,s}}. \quad (14)$$

The problem of minimizing the sum transmit power

$$P = \sum_{k=1}^{K} \sum_{s=1}^{S_k} p_{k,s} \quad (15)$$

under QoS and ZF constraints now reads as

$$\begin{aligned} \min_{\substack{(S_k)\,\forall k \\ (p_{k,s} \geq 0, \boldsymbol{g}_{k,s}, \boldsymbol{t}_{k,s})\,\forall (k,s)}} \quad & \sum_{k=1}^{K} \sum_{s=1}^{S_k} p_{k,s} \\ \text{s.t.}\ & r_k \geq \varrho_k \ \forall k \\ & \boldsymbol{g}_{k,s}^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{t}_{\ell,t} = 0 \ \forall (k,s) \neq (\ell,t) \end{aligned} \quad (16)$$

where $\varrho_k > 0$ is the minimum rate required by user k, and r_k is a function of the optimization variables as can be seen in (12) through (14). Note that the constraint $r_k \geq \varrho_k$ is always active in the optimal solution since the per-user rates r_k are strictly increasing in the per-stream powers $p_{k,s}$.

In order to fulfill both the ZF and the QoS constraints, it is necessary that $K \leq MC$, but there is no limit on the total number of receive antennas $\sum_{k=1}^{K} N_k$.

III. Greedy Stream Allocation

As we perform zero-forcing beamforming, the total number of streams that the base station can transmit is limited to $\sum_{k=1}^{K} S_k \leq MC$, but the number of channel outputs $\sum_{k=1}^{K} N_k C$ might be larger than MC. For this reason, a stream allocation procedure is necessary.

To avoid the exponential complexity of an exhaustive search, stream allocation is commonly performed by a successive scheme such as greedy allocation [6], [10], [13], where the base station allocates streams one after another in a way that the value of the cost function is decreased as much as possible in each allocation step. For the allocated streams, filter vectors have to be chosen, which is typically done based on a singular value decomposition (SVD) for the receive filters and based on channel inversion for the transmit filters (e.g., [6], [10]). However, it was shown in [2] that such schemes can only find carrier-noncooperative transmit strategies. Therefore, we combine greedy allocation with an iterative filter update method with random initialization, which—according to [2]—is capable of finding carrier-cooperative solutions.

Let $P(\boldsymbol{s})$ with $\boldsymbol{s} = \begin{bmatrix} S_1 & S_2 & \dots & S_K \end{bmatrix}^T \in \mathbb{N}^K$ be the sum power computed for a certain stream allocation $\boldsymbol{s}$ using the iterative filter update method, which is presented in detail in Section IV. In each step, the greedy algorithm, which is summarized in Algorithm 1, computes $P(\boldsymbol{s}+\boldsymbol{e}_j)$ for all users j with $S_j < N_j C$ and eventually allocates the stream such that the sum power decreases most, i.e.,

$$\boldsymbol{s} \leftarrow \boldsymbol{s} + \boldsymbol{e}_{j^\star} \text{ with } j^\star = \underset{j}{\operatorname{argmin}}\, P(\boldsymbol{s}+\boldsymbol{e}_j). \tag{17}$$

The procedure is stopped if MC streams have been allocated or no further decrease in sum power is achieved, i.e., if $P(\boldsymbol{s}+\boldsymbol{e}_{j^\star}) > P(\boldsymbol{s})$.

Note that choosing the filters by means of an iterative algorithm allows us to directly start with one data stream for each user, i.e., $S_k = 1\ \forall k$, while other successive allocation methods usually start with no data streams, i.e., $S_k = 0\ \forall k$. Consequently, we do not face the problem discussed, e.g., in [6], [8], that the QoS constraints of the optimization (16) cannot be fulfilled during the first allocation steps.

Algorithm 1 Greedy Power Minimization for CC ZF

set $\boldsymbol{s} \leftarrow \boldsymbol{s}_{\text{tmp}} \leftarrow \mathbf{1}$, compute $P(\boldsymbol{s}_{\text{tmp}})$ using Algorithm 2
while $\mathbf{1}^{\mathrm{T}}\boldsymbol{s} < MC$ **do**
    for all j with $\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{s} < N_j C$ **do**
        compute $P(\boldsymbol{s}+\boldsymbol{e}_j)$ using Algorithm 2
        if $P(\boldsymbol{s}+\boldsymbol{e}_j) < P(\boldsymbol{s}_{\text{tmp}})$ **then**
            $\boldsymbol{s}_{\text{tmp}} \leftarrow \boldsymbol{s}+\boldsymbol{e}_j$
        end if
    end for
    if $P(\boldsymbol{s}_{\text{tmp}}) > P(\boldsymbol{s})$ **then**
        break
    end if
    $\boldsymbol{s} \leftarrow \boldsymbol{s}_{\text{tmp}}$
end while

IV. Optimization of Filters and Per-Stream Rates

In order to reduce the transmit power, we perform an alternating optimization that consists of gradient descent steps for the receive filters $\boldsymbol{g}_{k,s}$ and a waterfilling-like optimization of new variables $\rho_{k,s}$ representing the per-stream rates.

Solving (13) for the per-stream powers, we can express

$$p_{k,s} = \frac{q_{k,s}}{\gamma_{k,s}} \text{ with } q_{k,s} = 2^{\rho_{k,s}} - 1 \tag{18}$$

explicitly as functions of the per-stream rates $\rho_{k,s}$ and the subchannel gains $\gamma_{k,s}$.

Furthermore, by introducing the projection matrix

$$\boldsymbol{P}_{k,s}^{\perp} = \mathbf{I}_{MC} - \boldsymbol{Q}_{k,s}\left(\boldsymbol{Q}_{k,s}^{\mathrm{H}}\boldsymbol{Q}_{k,s}\right)^{-1}\boldsymbol{Q}_{k,s}^{\mathrm{H}} \tag{19}$$

with

$$\boldsymbol{Q}_{k,s} = \big[\boldsymbol{H}_1^{\mathrm{H}}\boldsymbol{g}_{1,1}\ \boldsymbol{H}_1^{\mathrm{H}}\boldsymbol{g}_{1,2}\ \dots \\ \dots\ \boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}_{k,s-1}\ \boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}_{k,s+1}\ \dots\big] \tag{20}$$

we can write the optimal zero-forcing transmit filters as functions of all receive filters (see, e.g., [14]–[16]):

$$\boldsymbol{t}_{k,s} = \frac{\boldsymbol{P}_{k,s}^{\perp}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}_{k,s}}{\|\boldsymbol{P}_{k,s}^{\perp}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}_{k,s}\|}. \tag{21}$$

Since $\boldsymbol{P}_{k,s}^{\perp}$ projects onto the nullspace of all effective channels $\boldsymbol{g}_{\ell,t}^{\mathrm{H}}\boldsymbol{H}_\ell$ with $(\ell,t) \neq (k,s)$, it is not a function of $\boldsymbol{g}_{k,s}$, and it is invariant to permutations of the columns of $\boldsymbol{Q}_{k,s}$.

Combining (14) and (21), we get

$$\gamma_{k,s} = \frac{\boldsymbol{g}_{k,s}^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{P}_{k,s}^{\perp}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}_{k,s}}{\boldsymbol{g}_{k,s}^{\mathrm{H}}\boldsymbol{g}_{k,s}} \tag{22}$$

which only depends on the receive filters and is invariant to the scaling of these filters. Together with (18), this allows us to rewrite the optimization problem (16) as

$$\min_{(\rho_{k,s}\geq 0,\, \boldsymbol{g}_{k,s})\,\forall(k,s)} \sum_{k=1}^{K}\sum_{s=1}^{S_k}\frac{q_{k,s}}{\gamma_{k,s}} \quad \text{s.t.} \sum_{s=1}^{S_k}\rho_{k,s} = \varrho_k\ \forall k. \tag{23}$$

Note that we could also use the beamforming vectors $\boldsymbol{t}_{k,s}$ as optimization variables and express the receive filters as functions of the beamformers, i.e., perform zero-forcing at the receivers. However, in most practical systems, the base station has more degrees of freedom than the receivers, i.e., $MC > N_k C$. Therefore, when we aim at supporting up to MC data streams, at least part of the zero-forcing constraints have to be fulfilled by means of an appropriate choice of the transmit filters. A straightforward approach is to fulfill not only some, but all of the ZF constraints by choosing the transmit filters (21), which allows for arbitrary receive filters.

For fixed filters, the subchannel gains $\gamma_{k,s}$ are constants, and (23) reduces to a waterfilling-like optimization [17], which is solved by

$$\rho_{k,s} = \max\left\{0,\ \log_2\left(\frac{\lambda_k\gamma_{k,s}}{\ln 2}\right)\right\} \tag{24}$$

where the optimal water level is obtained from

$$\lambda_k = \ln 2\left(\frac{2^{\varrho_k}}{\prod_{s\in\mathcal{S}_{k,\mathrm{a}}}\gamma_{k,s}}\right)^{\frac{1}{|\mathcal{S}_{k,\mathrm{a}}|}} \tag{25}$$

with $\mathcal{S}_{k,\mathrm{a}}$ being the set of streams of user k with $\rho_{k,s} > 0$.

For fixed per-stream rates, the scalars $q_{k,s}$ are constants, and (23) reduces to an unconstrained optimization. Thus, the sum power can be decreased by a gradient descent step

$$\boldsymbol{g}_{k,s}^{\text{new}} \leftarrow \boldsymbol{g}_{k,s} - d\frac{\partial P}{\partial \boldsymbol{g}_{k,s}^{*}}, \quad \boldsymbol{g}_{k,s}^{\text{new}} \leftarrow \frac{\boldsymbol{g}_{k,s}^{\text{new}}}{\|\boldsymbol{g}_{k,s}^{\text{new}}\|} \quad \forall (k,s) \tag{26}$$

where d is the step size, and

$$\frac{\partial P}{\partial \boldsymbol{g}_{k,s}^{*}} = -\frac{q_{k,s}}{\gamma_{k,s}^2}\frac{\partial \gamma_{k,s}}{\partial \boldsymbol{g}_{k,s}^{*}} - \sum_{\substack{\ell=1 \\ (\ell,t)\neq(k,s)}}^{K}\sum_{t=1}^{S_\ell} \frac{q_{\ell,t}}{\gamma_{\ell,t}^2}\frac{\partial \gamma_{\ell,t}}{\partial \boldsymbol{g}_{k,s}^{*}} \tag{27}$$

with

$$\frac{\partial \gamma_{k,s}}{\partial \boldsymbol{g}_{k,s}^{*}} = \frac{1}{\|\boldsymbol{g}_{k,s}\|^2}\left(\boldsymbol{H}_k \boldsymbol{P}_{k,s}^{\perp} \boldsymbol{H}_k^{\mathrm{H}} - \gamma_{k,s}\mathbf{I}_{N_k C}\right)\boldsymbol{g}_{k,s} \tag{28}$$

$$\frac{\partial \gamma_{\ell,t}}{\partial \boldsymbol{g}_{k,s}^{*}} = \frac{1}{\|\boldsymbol{g}_{\ell,t}\|^2}\boldsymbol{H}_k \boldsymbol{P}_{\ell,t}^{\perp} \boldsymbol{H}_\ell^{\mathrm{H}} \boldsymbol{g}_{\ell,t}\boldsymbol{g}_{\ell,t}^{\mathrm{H}}\boldsymbol{H}_\ell \boldsymbol{Q}_{\ell,t}\left(\boldsymbol{Q}_{\ell,t}^{\mathrm{H}}\boldsymbol{Q}_{\ell,t}\right)^{-1}\boldsymbol{e}_j \tag{29}$$

where $(\ell, t) \neq (k, s)$, and j is the index of the column of $\boldsymbol{Q}_{\ell,t}$ containing the vector $\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}_{k,s}$. The scaling of $\boldsymbol{g}_{k,s}^{\text{new}}$ to unit norm in (26) is for numerical robustness and does not change the subchannel gains [cf. (22)].

We choose the step size d as

$$d \leftarrow \frac{d_{\text{init}}}{\max_{(k,s)}\left\|\frac{\partial P}{\partial \boldsymbol{g}_{k,s}^{*}}\right\|} \tag{30}$$

where the scaling operation is again for numerical robustness. As the gradient step can guarantee an improvement only locally, we implement a simple step size control: if the receive filters $\boldsymbol{g}_{k,s}^{\text{new}}$ obtained from (26) lead to an increased sum power, we divide the step size d by a factor of 2 and retry the gradient step until a decrease is eventually achieved. If even a certain minimal step size d_{min} does not lead to an improvement or the relative improvement is smaller than a given value ϵ, the algorithm terminates. In our simulations, we used $d_{\text{init}} = \frac{1}{2}$, $d_{\text{min}} = 10^{-14}$, and $\epsilon = 10^{-3}$. Due to the step size control, the method is robust to the choice of d_{init}.

As shown in Algorithm 2, we repeat the gradient step and the waterfilling procedure in an alternating manner. Convergence is guaranteed as the transmit power is bounded by the optimal value and both steps can only decrease the power.

In order to obtain a CC solution, the initial filter vectors have to correspond to a CC strategy [2]. Therefore, we choose the initial receive filters of new streams by applying a Gram-Schmidt orthogonalization to a set of random vectors. For existing streams, we keep the solution obtained in the previous outer iteration as initialization.

Algorithm 2 Power Minimization for Fixed Stream Allocation

choose CC initialization $\boldsymbol{g}_{k,s}$, compute $\gamma_{k,s}$ using (22)
while true do
  compute $\rho_{k,s}$ using (24), (25) and P using (15), (18)
  compute $\frac{\partial}{\partial \boldsymbol{g}_{k,s}^{*}}P$ using (27)
  compute d using (30)
  while true do
    compute $\boldsymbol{g}_{k,s}^{\text{new}}$ using (26)
    compute P^{new} from (15), (18) using $\boldsymbol{g}_{k,s}^{\text{new}}$ in (22)
    if $P^{\text{new}} < P$ **or** $d < d_{\text{min}}$ **then**
      break
    else
      $d \leftarrow \frac{d}{2}$
    end if
  end while
  if $d < d_{\text{min}}$ **or** $\frac{P - P_{\text{new}}}{P} < \epsilon$ **then**
    break
  end if
  $\boldsymbol{g}_{k,s} \leftarrow \boldsymbol{g}_{k,s}^{\text{new}}$
end while

V. Numerical Analysis

In this section, we compare the average sum transmit power of the proposed carrier-cooperative zero-forcing (CC ZF) algorithm to the power achieved by the greedy allocation (GA) discussed in [6], which is based on carrier-noncooperative transmission. Moreover, we include curves for the optimal sum transmit power achievable with DPC, i.e., with a nonlinear technique with time-sharing and without zero-forcing constraints. Though it is carrier-noncooperative, the DPC solution computed by means of the algorithm in [3] is the global optimum of the power minimization problem in parallel MIMO broadcast channels and, thus, a lower bound to any suboptimal CC or CN transmission scheme. All simulations are averaged in the dB-domain for 1000 channel realizations.

We consider a set of $C = 5$ parallel broadcast channels with $M = 2$ transmit antennas, $K = 5$ users, and $N_k = 2$ antennas at each receiver.[1] Fig. 1 shows simulation results for i.i.d. Gaussian channel coefficients with zero mean and unit variance. The simulations are performed with the same minimum rate requirement $\varrho_k = \varrho$ for all users k, and various values of ϱ are considered.

We observe that the three curves lie close to each other for small rate requirements ϱ, i.e., CC ZF does not achieve a notable gain compared to GA, and both suboptimal zero-forcing schemes perform close to the DPC lower bound. As the base station has more degrees of freedom than there are users in the system, relatively low rate requirements can be fulfilled without big effort, and neither the additional freedom obtained by performing CC transmission nor the interference cancellation by means of DPC leads to a significant improvement. However, for higher ϱ, serving all users at their required rate gets more challenging, and there is a significant gap between the GA method and the optimal DPC solution. The proposed CC ZF algorithm is able to close at least part of this gap.

The gain of CC transmission gets more pronounced in a different channel model. Let us assume that the receivers encounter interference (e.g., from other cells) whose power

[1]Note that scenarios with such a small number of carriers C can occur in a practical system, e.g., if groups of carriers within the frequency coherence interval are combined and considered as one logical carrier.

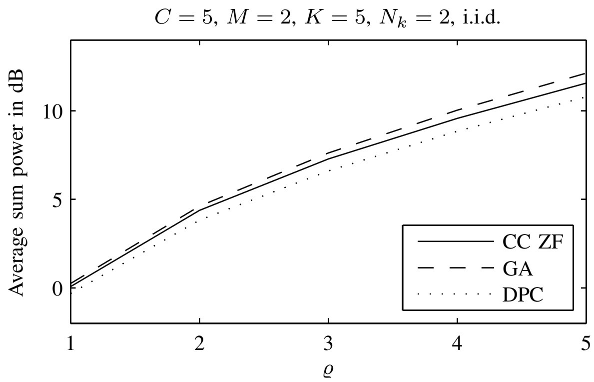

Fig. 1. Average power in dB for rate requirements $\varrho_k = \varrho\ \forall\ k$: simulation results for $M = 2,\ N = 2,\ K = 5,\ C = 5$, i.i.d. channel realizations.

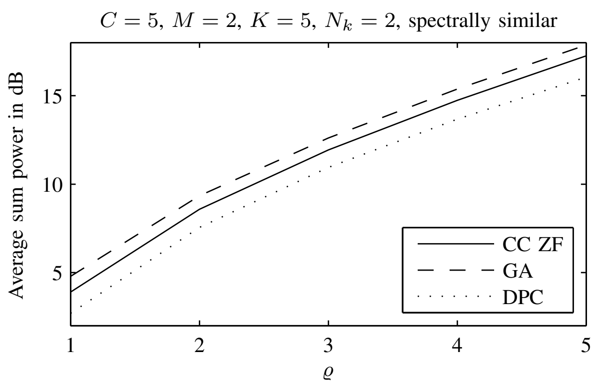

Fig. 2. Average power in dB for rate requirements $\varrho_k = \varrho\ \forall\ k$: simulation results for $M = 2,\ N = 2,\ K = 5,\ C = 5$, spectrally similar channels.

varies as a function of the carrier index c. Treating this interference as additional noise, the average quality of the whitened channels $\boldsymbol{H}_{k,\text{eff}}^{(c)} = \boldsymbol{C}_k^{(c),-\frac{1}{2}} \boldsymbol{H}_k^{(c)}$ (cf. Section II) depends on the carrier c. An easy way to model such a scenario is using scaled versions $\boldsymbol{H}_{k,\text{eff}}^{(c)} = w^{(c)} \boldsymbol{H}_k^{(c)}$ of the i.i.d. channel matrices $\boldsymbol{H}_k^{(c)}$ used above. For our simulations, we have chosen $w^{(c)} = 0.1 + 0.9\frac{c-1}{C-1}$. Adopting the nomenclature of [5], we say that the users have *spectrally similar* channels.

As can be seen in Fig. 2, CC transmission is beneficial even for low rate requirements in the case of spectrally similar channels. This can be explained as follows. By deciding which users to serve over the strongest carrier and which only over weaker ones, the transmit power needed for some users is traded off against the power needed for other users. While CN ZF can only make extremal decisions, namely serving or not serving a user on a carrier, CC transmission can implement a middle course, which might be a better trade-off [5].

VI. Conclusion

An algorithm to optimize carrier-cooperative zero-forcing transmission has to deal with the particularities of both concepts. As zero-forcing limits the number of possible data streams, a stream allocation has to be performed, and to optimize carrier-cooperative transmit strategies, iterative filter update methods are suitable. To derive a CC ZF power minimization algorithm for parallel MIMO broadcast channels, we have, therefore, combined greedy stream allocation with an iterative procedure based on gradient steps and waterfilling.

Numerical simulations show that the proposed CC ZF method can provide power savings compared to a state-of-the-art zero-forcing method without carrier cooperation. In particular, if the users in the system have spectrally similar channels, CC ZF leads to a significant improvement compared to conventional ZF. These results show that the suboptimality of CN ZF studied in [5] is not only of theoretical nature, but can indeed be observed in numerical simulations with random channels (and not only in a constructed example as in [5]).

References

[1] D. P. Palomar, M. A. Lagunas, and J. M. Cioffi, “Optimum linear joint transmit-receive processing for MIMO channels with QoS constraints,” *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1179–1197, May 2004.

[2] C. Hellings and W. Utschick, “Carrier-cooperative transmission in parallel MIMO broadcast channels: Potential gains and algorithms,” in *Proc. 8th Int. Symp. on Wireless Commun. Syst. (ISWCS) 2011*, Nov. 2011, pp. 774–778.

[3] M. Mohseni, R. Zhang, and J. Cioffi, “Optimized transmission for fading multiple-access and broadcast channels with multiple antennas,” *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1627–1639, Aug. 2006.

[4] C. Hellings and W. Utschick, “On the inseparability of parallel MIMO broadcast channels with linear transceivers,” *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6273–6278, Dec. 2011.

[5] C. Hellings, S. Herrmann, and W. Utschick, “Carrier cooperation can reduce the transmit power in parallel MIMO broadcast channels with zero-forcing,” 2012, submitted to IEEE Trans. Signal Process.

[6] C. Guthy, W. Utschick, and G. Dietl, “Spatial resource allocation for the multiuser multicarrier MIMO broadcast channel - a QoS optimization perspective,” in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP) 2010*, Mar. 2010, pp. 3166–3169.

[7] F. She, W. Chen, H. Luo, T. Huang, and X. Wang, “Joint power allocation and scheduling of multi-antenna OFDM system in broadcast channel,” presented at the Int. Conf. Commun. (ICC) 2009, Dresden, Germany, Jun. 14–18, 2009.

[8] C. Hellings, M. Joham, and W. Utschick, “Power minimization in parallel vector broadcast channels with zero-forcing beamforming,” presented at the IEEE GLOBECOM 2010, Miami, FL, Dec. 6–10, 2010.

[9] T. Yoo and A. Goldsmith, “On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming,” *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.

[10] F. Boccardi and H. Huang, “A near-optimum technique using linear precoding for the MIMO broadcast channel,” in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP) 2007*, vol. 3, Apr. 2007, pp. III-17–III-20.

[11] C. Guthy, W. Utschick, G. Dietl, and P. Tejera, “Efficient linear successive allocation for the MIMO broadcast channel,” in *Conf. Record 42nd Asilomar Conf. Signals, Syst., Comput. (ACSSC) 2008*, Oct. 2008, pp. 336–340.

[12] I. E. Telatar, “Capacity of multi-antenna Gaussian channels,” in *European Trans. Telecommun.*, vol. 10, Nov./Dec. 1999, pp. 585–595.

[13] G. Dimic and N. D. Sidiropoulos, “On downlink beamforming with greedy user selection: performance analysis and a simple new algorithm,” *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3857–3868, Oct. 2005.

[14] A. Wiesel, Y. Eldar, and S. Shamai, “Zero-forcing precoding and generalized inverses,” *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4409–4418, Sep. 2008.

[15] D. A. Schmidt, A. Gründinger, W. Utschick, and M. L. Honig, “Distributed precoder optimization for interfering MISO channels,” in *Proc. Int. ITG Workshop on Smart Antennas (ITG WSA) 2008*, Feb. 2008, pp. 285–290.

[16] Z. Zhou and B. Vucetic, “An orthogonal projection optimization algorithm for multi-user MIMO channels,” presented at the 71st Veh. Technol. Conf. (VTC) 2010-Spring, Taipei, Taiwan, May 16–19, 2010.

[17] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ: Wiley-Interscience, 2006.