POWER MINIMIZATION IN PARALLEL VECTOR BROADCAST CHANNELS WITH SEPARATE LINEAR PRECODING

Christoph Hellings, Wolfgang Utschick, and Michael Joham

Associate Institute for Signal Processing
Technische Universität München, 80290 München, Germany
phone: +49 89 289 28516, fax: +49 89 289 28522, email: {hellings|utschick|joham}@tum.de

ABSTRACT

The problem of power-efficient communication with guaranteed quality of service (expressed in terms of per-user rate constraints) is considered in a set of parallel vector broadcast channels with separate linear precoding on each subchannel. With the restriction to use linear precoding without time-sharing, the arising optimization problem is non-convex due to the non-concavity of the rate equations. Choosing a rate-space formulation, we derive an algorithm to compute the globally optimal solution, which is based on a branch-and-bound strategy to solve a difference-of-monotonic-functions (DM) reformulation of the problem. Despite its exponential complexity, the method is of theoretical interest as it can be used as a benchmark for heuristic algorithm. We also discuss a suboptimal algorithm, which is based on a recently proposed gradient-projection method for power minimization in multiple-input multiple-output (MIMO) broadcast channels. In numerical simulations, the gradient-based method turns out to perform close to the globally optimal solution if a good initialization is chosen.

1. INTRODUCTION

Many communication systems with a multi-antenna base station, e.g., fading or multi-carrier broadcast channels, can be modelled as parallel vector broadcast channels. If a certain quality of service (QoS) has to be guaranteed for the users of such a system, the question arises, how much power is needed to achieve the demanded QoS when optimally exploiting the diversity of the channel coefficients on the different carriers or during the different fading states.

A common approach is to model the quality of service requirements as minimum rate constraints. The resulting optimization problem will be introduced in Section 2 together with the mathematical description of the system model. If dirty-paper coding and time-sharing are possible in the considered communication system, the problem can be transformed to a convex optimization problem, and the globally optimal solution can be found by standard methods of convex optimization (e.g., [1, 2]). However, if the system is constrained to use linear transceivers, this is no longer possible.

Different heuristic approaches that are applicable to parallel vector broadcast channels with linear precoding have been proposed in the literature, e.g., in [3–6]. These algorithms make different assumptions about the types of strategies that are allowed in the system. For instance, [4–6] refrain from the application of time-sharing between different operation points, and [3–6] assume that precoding has to be performed separately on each of the parallel subchannels, i.e., no transmit symbol may be spread across several subchannels. The latter assumption was called carrier-noncooperative transmission in [7].

In our previous work [8], we considered the problem with time-sharing and with separate precoding on each subchannel. As time-sharing can be interpreted as allowing convex combinations between different transmit strategies, the rate constraints are convexified in this case, and the problem can be treated by means of a dual decomposition approach. The solution in [8] consists of solving a convex outer problem with non-convex inner problems for each subchannel, where the non-convex problems can be solved by means of monotonic optimization techniques.

In the paper at hand, we adopt both assumptions introduced above, i.e., we treat the case of separate precoding on each subchannel without time-sharing, as was done in [4–6]. However, different from the solutions in [4–6], we do not restrict ourselves to zero-forcing techniques. Since we do not allow time-sharing, a dual decomposition approach like in [8] is no longer possible since the non-convex problem without time-sharing exhibits a duality gap in general. Nevertheless, we will show that the globally optimal solution can be found. An algorithm computing this solution, which makes use of monotonicity properties of the rate-space reformulation introduced Section 3, is presented in Section 4.

In the literature, monotonic optimization frameworks have been used to optimize single-carrier wireless communication system with linear transceivers, e.g., in [9] for a vector broadcast channel with two users, in [10] for vector channels with an arbitrary number of users, in [11, 12] for a multiple-input single-output (MISO) interference channel with two transmitter-receiver pairs, and in [13] for a single-antenna interference channel with an arbitrary number of transmitter-receiver pairs. These works are based on the polyblock method [14], which is—as mentioned in [12]—a special case of the branch and bound framework (e.g. [15, 16]).

In this paper, monotonic optimization will be used to solve the power minimization problem in parallel vector broadcast channels with linear precoding: after classifying the optimization problem at hand as a difference-of-monotonic-functions (DM) problem, it can be solved with the branch-and-bound method from [16], which we briefly recapitulate in Section 4.

As will become clear at the end of that section, the approach is not meant to be implemented in a practical system since its computational complexity is prohibitively high. Therefore, we also discuss a less complex suboptimal solution (cf. Section 5), which is a special case of the gradient-projection algorithm from [17]. This method can be interpreted within the context of the same rate space formulation. The globally optimal solution can then be used as a benchmark to evaluate the performance of the gradient-projection method in numerical simulations.

Note that unlike for the problem discussed in [13], a geometric programming approximation as proposed in [18] is not applicable to the scenario considered in this paper. The reasons for this are twofold. First, due to the multi-antenna base-station, its beamforming vectors have to be chosen before the problem can be approximatively written as a geometric program, and an alternating optimization of the filter vectors and the power allocation is necessary. Thus, a series of geometric programs would have to be solved instead of a single geometric program. Second, the geometric programming approximation is based on the assumption of high signal-to-interference-and-noise ratio (SINR) for all users, which is not a reasonable assumption in a broadcast channel: unlike in an interference channel, intended signal and interference are transmitted over the same channel in a broadcast scenario. Consequently, unless the channels are orthogonalized by zero-forcing beamformers, the SINR can be very high for a user only if the transmit power of

This work was supported by Deutsche Forschungsgemeinschaft (DFG) under fund Jo724/1-1.

the signals intended for the other users is low, but in that case, the SINR becomes low for the other users. Due to this fact, a geometric program approximating the original optimization in the broadcast channel does not even have a feasible solution in general. Instead, a successive convex approximation [18], which involves computing a monomial approximation with an exponentially high number of terms, would be necessary. Therefore, the geometric programming framework from [18] is not considered in the remainder of this paper.

In this work, vectors are typeset in boldface letters, and we use $\mathbf{0}$ for the zero vector and $\mathbf{1}$ for the all-ones vector. The vector $\boldsymbol{e}_i$ is the i-th canonical unit vector, which has a one as the i-th entry and zeros elsewhere. The operations $\bullet^{\mathrm{T}}$ and $\bullet^{\mathrm{H}}$ denote the transpose and the conjugate transpose, respectively. We use the shorthand notation $(\bullet^{(n)})_{n=1}^{N}$ for $[\bullet^{(1),\mathrm{T}},\dots,\bullet^{(N),\mathrm{T}}]^{\mathrm{T}}$. The order relation $\boldsymbol{x}\geq\boldsymbol{y}$ has to be understood element-wise, and $\mathbb{R}^{n}_{0,+}$ is the closed positive orthant of the $\mathbb{R}^n$, i.e., $\mathbb{R}^{n}_{0,+}=\{\boldsymbol{x}\in\mathbb{R}^n\colon \boldsymbol{x}\geq\mathbf{0}\}$.

2. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a base station equipped with M antennas, which transmits independent data to K single-antenna receivers using N orthogonal subchannels (e.g., different carriers). The channels of all users k on all subchannels n are assumed to be frequency flat and perfectly known, so that they can be described by vectors $\boldsymbol{h}_k^{(n),\mathrm{H}}\in\mathbb{C}^{1\times M}$. The additive circularly symmetric complex Gaussian noise $\eta_k^{(n)}\sim\mathcal{CN}(0,\sigma_k^{(n),2})$ with variance $\sigma_k^{(n),2}$ is assumed to be independent across users and across subchannels and independent of the transmitted data.

We do not allow that a transmit symbol is spread across various subchannels. Thus, the signal that user k receives on subchannel n can be written as

$$y_k^{(n)}=\boldsymbol{h}_k^{(n),\mathrm{H}}\sum_{k'=1}^{K}\boldsymbol{t}_{k'}^{(n)}s_{k'}^{(n)}+\eta_k^{(n)}, \tag{1}$$

where $s_{k'}^{(n)}\sim\mathcal{CN}(0,1)$ is the data symbol of user k' on subchannel n, and $\boldsymbol{t}_{k'}^{(n)}\in\mathbb{C}^{M}$ is the corresponding beamforming vector.

It is well known that a set of rates can be achieved in a MISO broadcast channel with a certain sum transmit power if and only if the same rates can be achieved in the dual uplink [19] with the same sum transmit power. Since this duality can be applied on each subchannel, and since our aim is to minimize the sum of the per-subchannel transmit powers

$$P=\sum_{n=1}^{N}\sum_{k=1}^{K}\boldsymbol{t}_k^{(n),\mathrm{H}}\boldsymbol{t}_k^{(n)}, \tag{2}$$

we can also perform the optimization in the dual uplink. Afterwards, the downlink beamforming vectors can be obtained by transforming the results back to the downlink as described in [19].

In the dual uplink with uplink channel vectors $\boldsymbol{g}_k^{(n)}=\sigma_k^{(n),-1}\boldsymbol{h}_k^{(n)}$ and uplink noise covariance matrices $\boldsymbol{C}_{\boldsymbol{\eta}}^{(n)}=\mathbf{I}_M$, the signal received on subchannel n is given by

$$\boldsymbol{\xi}^{(n)}=\sum_{k=1}^{K}\boldsymbol{g}_k^{(n)}\sqrt{p_k^{(n)}}s_k^{(n)}+\boldsymbol{\eta}^{(n)}, \tag{3}$$

where $\boldsymbol{\eta}^{(n)}\sim\mathcal{CN}(\mathbf{0},\mathbf{I}_M)$ is the uplink noise on carrier n, and the sum transmit power can be expressed in terms of the uplink transmit powers $p_k^{(n)}$ as

$$P=\sum_{n=1}^{N}\sum_{k=1}^{K}p_k^{(n)}. \tag{4}$$

The achievable rate of user k on subchannel n is then given by

$$r_k^{(n)}(\boldsymbol{p}^{(n)})= \tag{5}$$
$$\log_2\left(1+p_k^{(n)}\boldsymbol{g}_k^{(n),\mathrm{H}}\left(\mathbf{I}_M+\sum_{k'\neq k}p_{k'}^{(n)}\boldsymbol{g}_{k'}^{(n)}\boldsymbol{g}_{k'}^{(n),\mathrm{H}}\right)^{-1}\boldsymbol{g}_k^{(n)}\right),$$

which is a function of the vector $\boldsymbol{p}^{(n)}=[p_1^{(n)},\dots,p_K^{(n)}]^{\mathrm{T}}$, and the per-user sum rate is $r_k=\sum_{n=1}^{N}r_k^{(n)}(\boldsymbol{p}^{(n)})$.

As already stated, we want to find the minimal transmit power that has to be spent in order to serve all users with a desired quality of service, where the QoS constraints are expressed in terms of demanded minimum rates $\boldsymbol{\rho}=[\rho_1,\dots,\rho_K]^{\mathrm{T}}$. The optimization problem reads

$$\min_{(\boldsymbol{p}^{(n)})_{n=1}^{N}\in\mathbb{R}^{KN}_{0,+}}\ \sum_{n=1}^{N}\mathbf{1}^{\mathrm{T}}\boldsymbol{p}^{(n)}\quad \text{s.t.}\ \sum_{n=1}^{N}\boldsymbol{r}^{(n)}(\boldsymbol{p}^{(n)})\geq\boldsymbol{\rho} \tag{6}$$

with $\boldsymbol{r}^{(n)}(\boldsymbol{p}^{(n)})=[r_1^{(n)}(\boldsymbol{p}^{(n)}),\dots,r_K^{(n)}(\boldsymbol{p}^{(n)})]^{\mathrm{T}}$. In this paper, we assume that the rate requirement vector $\boldsymbol{\rho}$ is chosen such that a feasible solution to (6) exists. To verify if this condition is fulfilled, the sufficient feasibility test for multi-carrier broadcast channels proposed in [20] can be used.

3. RATE SPACE FORMULATION

In this section, we will introduce a rate space formulation of problem (6). To this end, we introduce per-subchannel rate targets $\rho_k^{(n)}$ for all users as new variables, group them in vectors $\boldsymbol{\rho}^{(n)}=[\rho_1^{(n)},\dots,\rho_K^{(n)}]^{\mathrm{T}}$, and replace the rate constraints of problem (6) with

$$\boldsymbol{r}^{(n)}(\boldsymbol{p}^{(n)})\geq\boldsymbol{\rho}^{(n)}\geq\mathbf{0}\quad\forall n\quad\text{and}\quad\sum_{n}\boldsymbol{\rho}^{(n)}\geq\boldsymbol{\rho}. \tag{7}$$

The task is now to find the minimal transmit power for given per-subchannel rate targets and the optimal division of the per-user rate requirements in per-subchannel rate targets. A similiar approach was pursued in our previous work [17] on MIMO broadcast channels, where per-stream rate targets were introduced. The optimization with respect to these per-stream rate targets was performed in a suboptimal manner by a gradient-projection approach. In this paper, we will not only discuss the application of this method to the special case of parallel vector broadcast channels, but also propose an algorithm that is capable of finding the globally optimal per-subchannel rate targets.

For given values of all $\rho_k^{(n)}$, the power minimization problem can be solved separately on each subchannel as the coupling between the different subchannels is removed. Thus, the problem

$$\min_{\boldsymbol{p}^{(n)}\in\mathbb{R}^{K}_{0,+}}\ \mathbf{1}^{\mathrm{T}}\boldsymbol{p}^{(n)}\quad\text{s.t.}\ \boldsymbol{r}^{(n)}(\boldsymbol{p}^{(n)})\geq\boldsymbol{\rho}^{(n)} \tag{8}$$

can be solved independently on each subchannel n. Problem (8) is the well-investigated power minimization problem in single-carrier vector broadcast channels, which can be solved in a globally optimal manner, e.g., using the iterative method from [21].

Similar as in [8], we define the function

$$\boldsymbol{q}^{(n)}(\boldsymbol{\rho}^{(n)})=\begin{cases}\text{optimal }\boldsymbol{p}^{(n)}\text{ of (8) if (8) is feasible}\\ [\infty,\dots,\infty]^{\mathrm{T}}\text{otherwise,}\end{cases} \tag{9}$$

where infeasibility can result from too demanding rate targets $\rho_k^{(n)}$. The feasibility test can be performed by means of [22]. Problem (6) is then equivalent to

$$\min_{(\boldsymbol{\rho}^{(n)})_{n=1}^N \in \mathbb{R}_{0,+}^{KN}} \sum_{n=1}^N \mathbf{1}^{\mathrm{T}} \boldsymbol{q}^{(n)}(\boldsymbol{\rho}^{(n)}) \quad \text{s.t.} \sum_{n=1}^N \boldsymbol{\rho}^{(n)} \geq \boldsymbol{\rho}. \tag{10}$$

The equivalence between (6) and (10) follows from the fact that (8) is solved in a globally optimal manner. Thus, for the optimizer $(\boldsymbol{p}^{(n)})_{n=1}^N$ of (6), it holds that

$$\sum_{n=1}^N \mathbf{1}^{\mathrm{T}} \boldsymbol{q}^{(n)} \left(\boldsymbol{r}^{(n)}(\boldsymbol{p}^{(n)})\right) = \sum_{n=1}^N \mathbf{1}^{\mathrm{T}} \boldsymbol{p}^{(n)}, \tag{11}$$

and the optimal sum power of (6) is also achievable in (10). On the other hand, every feasible solution of (10) corresponds to a sum power sufficient to fulfill the constraints of the original problem (6), so that no sum power lower than the optimum of (6) is achievable in (10).

4. GLOBALLY OPTIMAL SOLUTION

In order to derive a globally optimal solution, we will now study the monotonicity properties of problem (10). The value of $\mathbf{1}^{\mathrm{T}} \boldsymbol{q}^{(n)}(\boldsymbol{\rho}^{(n)})$ is non-decreasing in $\boldsymbol{\rho}^{(n)}$, i.e., $\mathbf{1}^{\mathrm{T}} \boldsymbol{q}^{(n)}(\boldsymbol{\rho}'^{(n)}) \leq \mathbf{1}^{\mathrm{T}} \boldsymbol{q}^{(n)}(\boldsymbol{\rho}^{(n)})$ for $\boldsymbol{\rho}'^{(n)} \leq \boldsymbol{\rho}^{(n)}$. This can be seen from the fact that all rate vectors $\boldsymbol{\rho}'^{(n)} \leq \boldsymbol{\rho}^{(n)}$ are elements of the rate region with sum power $\mathbf{1}^{\mathrm{T}} \boldsymbol{q}^{(n)}(\boldsymbol{\rho}^{(n)})$ if $\boldsymbol{\rho}^{(n)}$ is achievable with finite sum power. The same reasoning was already used in [8]. If $\boldsymbol{\rho}^{(n)}$ cannot be achieved with finite sum power, the inequality is obvious. Consequently, problem (10) has a non-decreasing cost function.

The problem could be solved by an outer approximation method with NK variables similar to the polyblock approach [14]. Details of this method are omitted because we will concentrate on a different solution approach, which is based on the reformulation

$$\min_{(\boldsymbol{\rho}^{(n)})_{n=1}^{N-1} \in \mathbb{R}_{0,+}^{K(N-1)}} f_1\left((\boldsymbol{\rho}^{(n)})_{n=1}^{N-1}\right) - f_2\left((\boldsymbol{\rho}^{(n)})_{n=1}^{N-1}\right) \tag{12}$$

$$\text{with} \quad f_1\left((\boldsymbol{\rho}^{(n)})_{n=1}^{N-1}\right) = \sum_{n=1}^{N-1} \mathbf{1}^{\mathrm{T}} \boldsymbol{q}^{(n)}(\boldsymbol{\rho}^{(n)})$$

$$\text{and} \quad f_2\left((\boldsymbol{\rho}^{(n)})_{n=1}^{N-1}\right) = -\mathbf{1}^{\mathrm{T}} \boldsymbol{q}^{(n)} \left(\max\left\{\mathbf{0}, \boldsymbol{\rho} - \sum\nolimits_{n=1}^{N-1} \boldsymbol{\rho}^{(n)}\right\}\right),$$

where the maximum has to be understood element-wise. Rewriting the constraints that (10) imposes on $\boldsymbol{\rho}^{(N)}$ as

$$\boldsymbol{\rho}^{(N)} \geq \boldsymbol{\rho} - \sum\nolimits_{n=1}^{N-1} \boldsymbol{\rho}^{(n)} \quad \text{and} \quad \boldsymbol{\rho}^{(N)} \geq \mathbf{0} \tag{13}$$

and taking into account that at least one of the inequalities in (13) is fulfillled with equality in the optimum due to the monotonicity of the cost function of (10), the equivalence of (10) and (12) is evident.

In (12), the number of variables is $K(N-1)$, and the only constraint is the non-negativity constraint of the rate targets. The cost function is the difference of monotonically non-decreasing functions. Therefore, the problem can be solved with the branch-reduce-and-bound method from [16].

In the following, we will briefly recapitulate this method, adapting it to problem (12). Since the branch-reduce-and-bound (BRB) algorithm for difference-of-monotonic-functions (DM) problems from [16] also allows for DM constraints, but problem (12) only has a DM cost function, the step which is called reduction is not necessary for our application and will be omitted.

For a function f with

$$f(\boldsymbol{x}) = f_1(\boldsymbol{x}) - f_2(\boldsymbol{x}), \tag{14}$$

where f_1 and f_2 are non-decreasing in $\boldsymbol{x}$, the following bounds are valid:

$$f(\boldsymbol{x}) \geq L_{\mathcal{B}} = f_1(\boldsymbol{a}) - f_2(\boldsymbol{b}) \quad \forall \boldsymbol{x} \in \mathcal{B} \tag{15}$$

$$\min_{\boldsymbol{x} \in \mathcal{B}} f(\boldsymbol{x}) \leq U_{\mathcal{B}} = f_1(\boldsymbol{a}) - f_2(\boldsymbol{a}) \tag{16}$$

with $\mathcal{B} = \{\boldsymbol{x}\colon \boldsymbol{a} \leq \boldsymbol{x} \leq \boldsymbol{b}\}$. The first bound is due to the monotonicity of f_1 and f_2 and the second bound is an achievability bound. Given a set of boxes $\mathbb{B}$, the smallest lower bound $L = \min_{\mathcal{B} \in \mathbb{B}} L_{\mathcal{B}}$ is a lower bound to $f(\boldsymbol{x})$ for $\boldsymbol{x} \in \bigcup_{\mathcal{B} \in \mathbb{B}} \mathcal{B}$. By always replacing the box $\mathcal{B}^* = \operatorname{argmin}_{\mathcal{B} \in \mathbb{B}} L_{\mathcal{B}}$ by two disjoint subboxes $\mathcal{B}'$ and $\mathcal{B}''$ with $\mathcal{B}' \cup \mathcal{B}'' = \mathcal{B}^*$ (using a bisectional rectangular subdivision, cf. [16]), L can be increased until $|L - \min_{\boldsymbol{x} \in \bigcup_{\mathcal{B} \in \mathbb{B}} \mathcal{B}} f(\boldsymbol{x})| < \epsilon$ if f_1 and f_2 are continuous [16]. The reason for this is that for continuous f_1, f_2, the bound (15) is consistent, i.e., $|L_{\mathcal{B}} - \min_{\boldsymbol{x} \in \mathcal{B}} f(\boldsymbol{x})| \to 0$ for $\boldsymbol{b} - \boldsymbol{a} \to \mathbf{0}$.

As the minimum $\min_{\boldsymbol{x} \in \bigcup_{\mathcal{B} \in \mathbb{B}} \mathcal{B}} f(\boldsymbol{x})$ is unknown, convergence is instead detected using the criterion $|L - U| < \epsilon$, where $U = \min_{\mathcal{B} \in \mathbb{B}} U_{\mathcal{B}}$ is the smallest upper bound, i.e., the current best value. For continuous functions f_1, f_2, (16) is consistent, too, so that it is ensured that U eventually converges to the actual minimum, and thus $|L - U| \to 0$.

For the problem at hand, we have $\boldsymbol{x} = (\boldsymbol{\rho}^{(n)})_{n=1}^{N-1}$, and we start with a box $\mathcal{B}_0 = \{\boldsymbol{x}\colon \mathbf{0} \leq \boldsymbol{x} \leq (\boldsymbol{y}^{(n)})_{n=1}^{N-1}\}$ with $\boldsymbol{y}^{(n)} = \boldsymbol{\rho}$. This choice of $\mathcal{B}_0$ surely contains the optimizer of problem (12) and all $\boldsymbol{x} \in \mathcal{B}_0$ fulfill the non-negativity constraint. For convergence, it suffices that f_1 and f_2 are continuous at all points yielding a finite value of the function, and this continuity was shown in [8, Lemma 1]. The intuition behind this property is that for achievable rate targets $\boldsymbol{\rho}^{(n)}$, $\boldsymbol{q}^{(n)}(\boldsymbol{\rho}^{(n)})$ can be shown to be the inverse function of the continuous uplink rate function $\boldsymbol{r}^{(n)}(\boldsymbol{p}^{(n)})$.

In each iteration, the two bounds have to be calculated for two new subboxes $\mathcal{B}'$ and $\mathcal{B}''$, where the lower corner of $\mathcal{B}'$ coincides with the lower corner of $\mathcal{B}^*$ while the upper corner of $\mathcal{B}''$ coincides with the upper corner of $\mathcal{B}^*$. Reusing previously computed values of f_1 and f_2, this can be done with one evaluation of f_1, namely at the lower corner $\boldsymbol{a}''$ of $\mathcal{B}''$, and two evaluations of f_2, namely at $\boldsymbol{a}''$ and at the upper corner $\boldsymbol{b}'$ of $\mathcal{B}'$. Thus, the per-subchannel problem (8), which can be solved in polynomial time, has to be solved $N+1$ times per iteration of the branch-and-bound method. However, it follows from [23, Theorem 4] that in the worst case, the number of branch-and-bound iterations needed to find an ϵ-optimal solution is $\mathcal{O}((\frac{c_1}{\epsilon})^{\frac{K(N-1)}{c_2}})$, where c_1 and c_2 are constants that depend on properties of the objective function. Therefore, the worst-case complexity of the algorithm to solve (12) is exponential in K and N.

To get an impression about the complexity of the algorithm, the number of branch-and-bound iterations needed for the numerical simulations in Section 6 is plotted in Fig. 1. The histogram shows how often the number of iterations needed to compute the optimal solution lies in the indicated intervals. All details on the system considered in the numerical simulations are given in Section 6. We observe a very high number of iterations for some channel realizations while in most cases, a significantly lower number is sufficient. This complies with the fact that the algorithm has a high worst-case complexity.

5. GRADIENT-BASED SUBOPTIMAL SOLUTION

To solve problem (10) by means of a gradient-projection approach, we have to calculate the partial derivative of $P =$

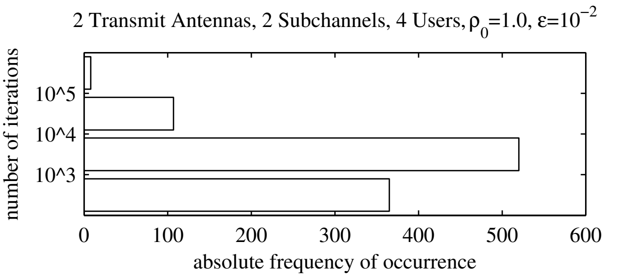

Figure 1: Histogram of the Number of Iterations of the Branch-and-Bound Algorithm.

$\sum_{s=1}^{N} \mathbf{1}^{\mathrm{T}} \boldsymbol{q}^{(s)}(\boldsymbol{\rho}^{(s)})$ with respect to $\rho_k^{(n)}$. To do so, we make use of the equivalence of the considered parallel vector broadcast channels with a MIMO broadcast channel with block-diagonal channel matrices

$$\boldsymbol{H}_k^{\mathrm{H}} = \mathrm{blkdiag}\left(\boldsymbol{h}_k^{(1),\mathrm{H}}, \ldots, \boldsymbol{h}_k^{(N),\mathrm{H}}\right) \in \mathbb{C}^{N \times MN}, \tag{17}$$

and diagonal noise covariance matrices

$$\boldsymbol{C}_{\boldsymbol{\eta}_k} = \mathrm{diag}\left(\sigma_k^{(1),2}, \ldots, \sigma_k^{(N),2}\right) \in \mathbb{C}^{N \times N} \tag{18}$$

(cf., e.g., [7]). The assumption of separate coding on each subchannel is equivalent to using diagonal matrices as transmit filters and block-diagonal matrices as receive filters in the dual uplink of this equivalent MIMO system.

Applying the results from [17] to this equivalent MIMO system taking into account the diagonal structure of the uplink precoding matrices, we get

$$\frac{\partial P}{\partial \rho_k^{(n)}} = \mathbf{1}^{\mathrm{T}} \begin{bmatrix} \frac{\partial r_1^{(n)}}{\partial p_1^{(n)}} & \cdots & \frac{\partial r_1^{(n)}}{\partial p_K^{(n)}} \\ \vdots & \ddots & \vdots \\ \frac{\partial r_K^{(n)}}{\partial p_1^{(n)}} & \cdots & \frac{\partial r_K^{(n)}}{\partial p_K^{(n)}} \end{bmatrix}^{-1} \boldsymbol{e}_k. \tag{19}$$

The partial derivatives inside the matrix are obtained by taking the derivative of (5), and they have to be evaluated at $\boldsymbol{p}^{(n)} = \boldsymbol{q}^{(n)}(\boldsymbol{\rho}^{(n)})$.

In general, after a gradient step

$$\tilde{\rho}_k^{(n)} \leftarrow \rho_k^{(n)} - d \frac{\partial P}{\partial \rho_k^{(n)}} \quad \forall k, \forall n \tag{20}$$

with step size d, the constraint of problem (10) is no longer fulfilled. A projection to the set of feasible rate targets in the sense of minimal Euclidian norm can be performed by solving a waterfilling-like optimization problem as was discussed, e.g., in [24].

If the value of the cost function of (10) is increased by the gradient update and the projection, a too large step size has been used. In this case, the gradient-projection step has to be performed with a decreased step size instead. Since the function $\boldsymbol{q}^{(n)}(\boldsymbol{\rho}^{(n)})$ maps infeasible rate targets to infinite transmit power, they are only a special case of an increased sum transmit power. In the subsequent iteration, the algorithm starts again with the initial step size and again decreases the step size until a decrease in sum power is attained. Repeating the gradient step and the projection until convergence, we obtain a suboptimal solution to (10). The step size adaption ensures that the sum transmit power is monotonically decreasing. Since the power is, in addition, bounded from below by the optimal solution, convergence of the algorithm is guaranteed.

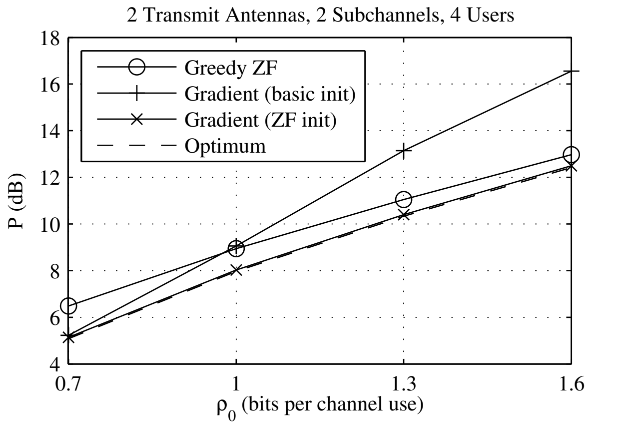

Figure 2: System with 4 Users: Transmit Power for Different Per-User Rate Requirements.

The algorithm described in this section is, in fact, a special case of the algorithm for MIMO broadcast channels from [17]. The difference is, however, that the update of the uplink transmit filters is eventually nothing more than a power allocation procedure due to the assumption of diagonal filter matrices. This power allocation is implicitly optimized within each evaluation of $\boldsymbol{q}^{(n)}(\boldsymbol{\rho}^{(n)})$. Furthermore, no uplink equalizers need to be computed during the execution of the algorithm since the formulation based on the rate equation (5) does not depend on these equalizers. Instead, they only need to be computed as part of the uplink-to-downlink transformation after convergence of the algorithm. Therefore, the algorithm from [17] is reduced to a series of gradient-projection steps for the system considered in this paper.

To compute the gradient in the first iteration, the initial rate targets $(\boldsymbol{\rho}^{(n)})_{n=1}^{N}$ need to be feasible. A feasible initialization could be found using the algorithm proposed in [20]. Another possibility is to use the per-stream rates resulting from any other heuristic power minimization as initial per-stream rate targets. For instance, the greedy zero-forcing scheme discussed in [6] could be used.

6. NUMERICAL RESULTS AND DISCUSSION

Due to its exponential complexity, the algorithm computing the globally optimal solution is not feasible for practical implementation. However, it is useful as a benchmark for less complex suboptimal algorithms. In Fig. 2, we consider a system with $M = 2$ transmit antennas, $N = 2$ subchannels, and $K = 4$ users. The small dimensionality of the problem enables us to compute the globally optimal solution by means of the proposed algorithm. The rate requirements are $\rho_1 = \rho_2 = \rho_0$ and $\rho_3 = \rho_4 = 2\rho_0$, all channel coefficients are i.i.d. circularly symmetric complex Gaussian with zero mean and unit variance, and the noise power is assumed to be $\sigma_k^{(n),2} = 1 \ \forall k, \forall n$. To average over 1,000 realizations, we use the arithmetic mean in the dB domain.

As discussed in Section 5, different initializations of the per-subchannel rate targets are possible for the suboptimal gradient-based algorithm. For the curve called "basic init", we used the algorithm from [20] to obtain a feasible initialization, while the rate allocation of the greedy zero-forcing scheme from [6] was used as initialization for the curve called "ZF init."

As can be seen, different choices of the initial per-stream rate targets can have a strong impact on the outcome of the algorithm as the algorithm may converge to different stationary points for different initializations. Especially for high rate requirements, the gradient-based method with the basic initialization computes solutions with high sum transmit power and is even outperformed by the greedy zero-forcing (ZF) scheme from [6]. However, using the per-stream rates computed by this zero-forcing method as initial

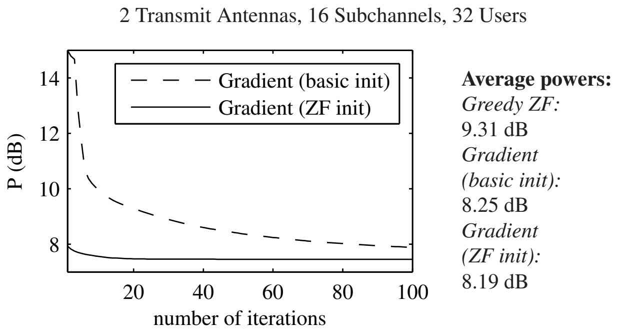

Figure 3: System with 32 Users and Random Rate Requirements: Convergence for a Certain Channel/Requirement Realization and Average Powers for 1,000 Realizations

per-stream rate targets, the gradient method can achieve a close-to-optimum performance on average. The reason for this is that it is more likely that the algorithm converges to the globally optimal solution if an initialization is used that is in the neighborhood of this solution. It is known that zero-forcing can perform close to the globally optimal solution (especially for high rate requirements) while the basic initialization might correspond to very high transmit powers as it is based only on feasibility considerations.

The high initial transmit power of the basic initialization can also be observed in Fig. 3, where we have plotted the development of the sum transmit power during the execution of the algorithm. For this simulation, we chose a larger system ($M = 2$ transmit antennas, $N = 16$ subchannels, and $K = 32$ users) and random per-user rate requirements, which are the absolute values of i.i.d. real Gaussian random variables with zero mean and unit variance. As this system does not operate in the high rate regime, the gradient method with the basic initialization is also able to achieve low transmit powers, but the number of necessary iterations is significantly lower if the zero-forcing initialization is used.

In summary, the gradient-projection algorithm proposed in this paper, quickly converges to solutions that are close to the global optimum if a good initialization is used. Furthermore, as can be easily verified, the steps necessary in each iteration can be performed in polynomial time, so that the complexity of the method is significantly reduced when compared to the globally optimal solution with exponential complexity.

REFERENCES

[1] 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.

[2] G. Wunder and T. Michel, “Minimum Rates Scheduling for MIMO OFDM Broadcast Channels,” in *Proc. ISSSTA 2006*, Aug. 2006, pp. 510–514.

[3] N. Hassan and M. Assaad, “Low complexity margin adaptive resource allocation in downlink MIMO-OFDMA system,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3365–3371, Jul. 2009.

[4] C. Guthy, W. Utschick, and G. Dietl, “Spatial Resource Allocation for the Multiuser Multicarrier MIMO Broadcast Channel – A QoS Optimization Perspective,” in *Proc. ICASSP 2010*, Mar. 2010.

[5] F. She, W. Chen, H. Luo, T. Huang, and X. Wang, “Joint Power Allocation and Scheduling of Multi-Antenna OFDM System in Broadcast Channel,” in *Proc. ICC 2009*, Jun. 2009.

[6] C. Hellings, M. Joham, and W. Utschick, “Power Minimization in Parallel Vector Broadcast Channels with Zero-Forcing Beamforming,” in *Proc. IEEE GLOBECOM 2010*, Dec. 2010.

[7] 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.

[8] C. Hellings, M. Joham, M. Riemensberger, and W. Utschick, “Minimal Transmit Power in Parallel Vector Broadcast Channels with Linear Precoding,” 2010, *submitted to IEEE Trans. Signal Process.*

[9] E. A. Jorswieck and E. G. Larsson, “Linear precoding in multiple antenna broadcast channels: Efficient computation of the achievable rate region,” in *Proc. ITG WSA 2008*, Feb. 2008, pp. 21–28.

[10] J. Brehmer and W. Utschick, “Utility Maximization in the Multi-User MISO Downlink with Linear Precoding,” in *Proc. ICC 2009*, Jun. 2009.

[11] E. A. Jorswieck and E. G. Larsson, “Monotonic optimization framework for the MISO IFC,” in *Proc. ICASSP 2009*, Apr. 2009, pp. 3633–3636.

[12] ——, “Monotonic Optimization Framework for the Two-User MISO Interference Channel,” *IEEE Trans. Commun.*, vol. 58, no. 7, pp. 2159–2168, Jul. 2010.

[13] L. P. Qian, Y. J. Zhang, and J. Huang, “MAPEL: Achieving global optimality for a non-convex wireless power control problem,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1553–1563, Mar. 2009.

[14] H. Tuy, “Monotonic Optimization: Problems and Solution Approaches,” *SIAM Journal on Optimization*, vol. 11, no. 2, pp. 464–494, 2000.

[15] R. Horst and H. Tuy, *Global Optimization: Deterministic Approaches*, 3rd ed. Berlin, Germany: Springer, 1996.

[16] H. Tuy, F. Al-Khayyal, and P. Thach, “Monotonic Optimization: Branch and Cut Methods,” in *Essays and Surveys in Global Optimization*. New York, NY, USA: Springer, 2005, ch. 2, pp. 39–78.

[17] C. Hellings, M. Joham, and W. Utschick, “Gradient-Based Power Minimization in MIMO Broadcast Channels with Linear Precoding,” 2011, *submitted to IEEE Trans. Signal Process.*

[18] M. Chiang, C. W. Tan, D. Palomar, D. O’Neill, and D. Julian, “Power Control By Geometric Programming,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640–2651, Jul. 2007.

[19] P. Viswanath and D. N. C. Tse, “Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality,” *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.

[20] M. Joham, C. Hellings, and R. Hunger, “QoS Feasibility for the MIMO Broadcast Channel: Robust Formulation and Multi-Carrier Systems,” in *Proc. WiOpt 2010*, May 2010, pp. 610–614.

[21] M. Schubert and H. Boche, “Solution of the multiuser downlink beamforming problem with individual SINR constraints,” *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[22] R. Hunger and M. Joham, “A Complete Description of the QoS Feasibility Region in the Vector Broadcast Channel,” *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3870–3878, Jul. 2010.

[23] S. A. Vavasis, “Complexity Issues in Global Optimization: A Survey,” Cornell University, Ithaca, NY, USA, Tech. Rep., Jan. 1993.

[24] G. Scutari, D. P. Palomar, and S. Barbarossa, “Optimal Linear Precoding Strategies for Wideband Noncooperative Systems Based on Game Theory—Part II: Algorithms,” *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1250–1267, Mar. 2008.