Utility Maximization in the Multi-User MISO Downlink with Linear Precoding

Johannes Brehmer, Wolfgang Utschick
Associate Institute for Signal Processing
Technische Universität München
{brehmer, utschick}@tum.de

***Abstract*—The maximization of an increasing function over the set of achievable rates in a multi-user, multi-antenna downlink is addressed. In general, the set of rates achievable by linear precoding and treating interference as noise is nonconvex. As a result, the corresponding utility maximization problem is nonconvex. The rate region can be convexified by time sharing, and the utility maximization over the convexified region can be solved via Lagrange duality. Still, subproblems in the dual problem remain nonconvex. It is shown how all the aforementioned nonconvex problems can be solved to global optimality in the framework of monotonic optimization. Moreover, it is investigated to what extent utility is increased by time sharing. While all problems can be solved to global optimality, the resulting computational complexity is rather high, thus the proposed solution strategies mainly provide a benchmark for locally optimum, less complex methods. Numerical results demonstrate that a method which finds stationary points on the boundary of the rate region can provide close-to-optimum performance.**

I. Introduction

We consider a multi-user wireless communication system, where a central transmitter (base station, access point) sends data to K receivers. The transmitter is assumed to be equipped with multiple transmit antennas, while the receivers all have a single antenna. The transmitter performs linear precoding, but, due to complexity reasons, does not employ strategies to cancel known interference. Optimization of such multi-user MISO systems under different performance metrics and constraints has received wide attention in the last decade. Examples include power minimization under SINR constraints [1], minimization of sum MSE or MSE and SINR balancing under a power constraint [1], [2]. Treating interference as noise usually results in nonconvex optimization problems. For the power minimization and balancing problems, efficient algorithms exist that converge to the global optimum, see, e.g., [1]. For other problems, such as the maximization of the sum rate [3], no algorithms exist that have a practically feasible complexity for larger problem size and are globally optimum. In [4], sum rate maximization is solved to global optimality by using an algorithm for deterministic global optimization that exploits a monotonicity in the problem structure.

We assume that the physical layer is described by its achievable rate region, and the properties of the upper layers are modeled by a system utility function [5], [6], whose value depends only on the rates provisioned to the K users. Under this simple model, the design objective is to find the rate vector that maximizes utility. In the multi-user MISO downlink, the rate region is (in general) nonconvex [3], rendering the utility maximization problem a nonconvex problem. The rate region can be convexified by time sharing – but the algorithm needed to compute rate points on the convex region still corresponds to a nonconvex problem.

We consider two types of utilities, concave and nonconcave [7], and two types of rate regions, nonconvex and convex by time sharing, resulting in three types of optimization problems, each having a different structure. For each type, we develop a solution strategy. Based on the fundamental assumption that the utility function is increasing in the users' rates, the utility maximization problem always exhibits a monotonicity structure, which can be exploited to find the global optimum with (relatively) efficient deterministic algorithms [8]. As a result, by proper decomposition and exploitation of the monotonicity structure, all proposed strategies yield globally optimal solutions.

Despite the relative efficiency of the globally optimum strategies, the resulting computational complexity still grows quickly with the number of users. As a result, the proposed strategies mainly serve as a benchmark for sub-optimum, less complex approaches. As a simple, locally optimal strategy, we employ a gradient projection algorithm to find a stationary point on the boundary of the rate region.

II. System Model

A multi-user downlink is considered, with a central base station with N transmit antennas transmitting to K single-antenna receivers. The received signal at the k-th receiver is given by

$$y_k = \sum_{q=1}^{K} \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{w}_q s_q + \eta_k,$$

where $\boldsymbol{h}_k^{\mathrm{H}} \in \mathbb{C}^{1\times N}$, $\boldsymbol{w}_k \in \mathbb{C}^{N\times 1}$, and $s_k \in \mathbb{C}$ are the channel, precoder, and data symbol of user k, respectively, and η_k is circularly symmetric AWGN with zero mean and variance σ^2. The transmitted signal

$$\boldsymbol{x} = \sum_{k=1}^{K} \boldsymbol{w}_k s_k$$

is subject to a sum-power constraint $\mathrm{E}\left[\|\boldsymbol{x}\|_2^2\right] \leq P$, which, under the assumption of uncorrelated, unit power data symbols, translates into a constraint on the precoders $\boldsymbol{w}_k$:

$$\sum_{k=1}^{K} \|\boldsymbol{w}_k\|_2^2 \le P. \tag{1}$$

III. Rate Region

Treating the interference from other users as noise, for given precoders $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_K)$ an achievable rate vector $\boldsymbol{R} = (R_1, \ldots, R_K)$ is given by $\boldsymbol{R} \in \mathbb{R}_+^K : R_k < r_k(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_K)$, with

$$r_k(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_K) = \log_2 \left(1 + \frac{|\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{w}_k|^2}{\sigma^2 + \sum_{q \ne k} |\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{w}_q|^2} \right).$$

The set of achievable rate vectors is defined as the closure of all such vectors for a given transmit power constraint, i.e.,

$$\mathcal{R} = \left\{ \boldsymbol{r}(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_K) : \sum_{k=1}^{K} \|\boldsymbol{w}_k\|_2^2 \le P \right\}. \tag{2}$$

Using the duality relation between uplink and downlink [1], a more compact parameterization of the rate region $\mathcal{R}$ can be given as follows:

$$\mathcal{R} = \{ \boldsymbol{r}(\boldsymbol{p}) : \boldsymbol{p} \in \mathcal{P} \}, \tag{3}$$

with the set of feasible transmit powers in the dual uplink

$$\mathcal{P} = \left\{ \boldsymbol{p} \in \mathbb{R}_+^K : \|\boldsymbol{p}\|_1 \le P \right\}, \tag{4}$$

and the uplink rates

$$r_k(\boldsymbol{p}) = \log_2 \det \left(\boldsymbol{I} + \left(\sigma^2 \boldsymbol{I} + \sum_{q \ne k} \boldsymbol{h}_q \boldsymbol{h}_q^{\mathrm{H}} p_q \right)^{-1} \boldsymbol{h}_k \boldsymbol{h}_k^{\mathrm{H}} p_k \right), \tag{5}$$

As discussed in, e.g., [3], the rate region $\mathcal{R}$ may be nonconvex. By allowing for time-sharing between vectors in $\mathcal{R}$, any vector in the convex hull of $\mathcal{R}$ is also achievable, resulting in a second set of achievable rates, denoted as $\mathcal{C}$:

$$\mathcal{C} = \mathrm{co}(\mathcal{R}).$$

IV. Utility Maximization

The sets $\mathcal{R}$ an $\mathcal{C}$ define the sets of achievable rates for the system model under consideration. In the following, the problem of determining a rate vector in $\mathcal{R}$ or $\mathcal{C}$ that maximizes a utility function u is considered. In this work, a very general notion of utility is employed: A utility function u is simply a function $u : \mathbb{R}_+^K \to \mathbb{R}$ that is increasing, i.e.,

$$\boldsymbol{r} \le \boldsymbol{r}' \Rightarrow u(\boldsymbol{r}) \le u(\boldsymbol{r}'). \tag{6}$$

Moreover, it is assumed that u is continuous. An additional assumption that is frequently made in the context of utility maximization is that u is concave. As argued in [7], nonconcave utilities represent an important class of applications, therefore this work does not require u to be concave.

If no time-sharing mode is provisioned, the utility maximization problem is given by

$$\max_{\boldsymbol{r}} u(\boldsymbol{r}) \quad \text{s.t.} \quad \boldsymbol{r} \in \mathcal{R}. \tag{7}$$

Problem (7) corresponds to the maximization of an increasing function over a compact set. Even if u is assumed to be concave, problem (7) can in general not be assumed to be convex, due to the potential nonconvexity of $\mathcal{R}$.

If time-sharing is allowed, the utility maximization problem reads as

$$\max_{\boldsymbol{r}} u(\boldsymbol{r}) \quad \text{s.t.} \quad \boldsymbol{r} \in \mathcal{C}. \tag{8}$$

Problem (8) corresponds to the maximization of an increasing function over a compact convex set. If u is concave, problem (8) is convex.

V. Monotonic Optimization

The utility maximization problems (7) and (8) correspond to the maximization of an increasing function over a compact set in $\mathbb{R}_+^K$. Optimization problems that exhibit this structure can be solved in the framework of monotonic optimization [8]. Monotonic optimization provides efficient deterministic algorithms for solving monotonic problems to global optimality.

A basic problem of monotonic optimization is the maximization of an increasing function over a compact normal set [8]. A subset $\mathcal{S}$ of $\mathbb{R}_+^D$ is said to be *normal* in $\mathbb{R}_+^D$ (or briefly, normal), if $\boldsymbol{x} \in \mathcal{S}, \boldsymbol{0} \le \boldsymbol{y} \le \boldsymbol{x} \Rightarrow \boldsymbol{y} \in \mathcal{S}$. The rate regions $\mathcal{C}$ and $\mathcal{R}$ are normal: any rate vector $\boldsymbol{r}'$ that is smaller than an achievable rate vector $\boldsymbol{r}$ is also achievable.

A. Polyblock Algorithm

Let $f : \mathbb{R}_+^D \to \mathbb{R}$ be a continuous, increasing function and $\mathcal{S} \subset \mathbb{R}_+^D$ a compact normal set. Then

$$\max_{\boldsymbol{x}} f(\boldsymbol{x}) \quad \text{s.t.} \quad \boldsymbol{x} \in \mathcal{S} \tag{9}$$

constitutes a monotonic optimization problem. The basic algorithm for solving monotonic optimization problems is the so-called *polyblock algorithm*. A polyblock is simply the union of a finite number of hyper-rectangles in $\mathbb{R}_+^D$: Given a discrete set $\mathcal{V} \subset \mathbb{R}_+^D$, a polyblock $\mathcal{P}(\mathcal{V})$ is defined as

$$\mathcal{P}(\mathcal{V}) = \bigcup_{\boldsymbol{v} \in \mathcal{V}} \{ \boldsymbol{x} \in \mathbb{R}_+^D, \boldsymbol{x} \le \boldsymbol{v} \}.$$

The set $\mathcal{V}$ contains the vertices of the polyblock $\mathcal{P}(\mathcal{V})$.

Due to the fact that $\mathcal{S}$ is a compact normal subset of $\mathbb{R}_+^D$ there exists a set $\mathcal{V}^{(0)}$ such that $\mathcal{S} \subseteq \mathcal{P}(\mathcal{V}^{(0)})$. Moreover, starting with $n = 0$, either $\mathcal{S} = \mathcal{P}(\mathcal{V}^{(n)})$ or there exists a discrete set $\mathcal{V}^{(n+1)} \subset \mathbb{R}_+^K$ such that

$$\mathcal{S} \subseteq \mathcal{P}(\mathcal{V}^{(n+1)}) \subset \mathcal{P}(\mathcal{V}^{(n)}). \tag{10}$$

In other words, the polyblocks $\mathcal{P}(\mathcal{V}^{(n)})$ represent an iteratively refined outer approximation of the set $\mathcal{S}$.

Consider the problem of maximizing an increasing function f over the polyblock $\mathcal{P}(\mathcal{V}^{(n)})$:

$$\max_{\boldsymbol{x} \in \mathcal{P}(\mathcal{V}^{(n)})} f(\boldsymbol{x}). \tag{11}$$

Let $\check{\boldsymbol{x}}^{(n)}$ denote a maximizer of problem (11). Due to the monotonicity of f, there exists a maximizer such that $\check{\boldsymbol{x}}^{(n)} \in \mathcal{V}^{(n)}$, i.e., the maximum of an increasing function over a polyblock is attained on one of the vertices [8]. Due to the fact that the vertex set of a polyblock is discrete, problem (11) can be solved to global optimality by searching over all $\boldsymbol{v} \in \mathcal{V}^{(n)}$.

If $\check{\boldsymbol{x}}^{(n)} \in \mathcal{S}$, a global maximizer is found. In general, however, $\check{\boldsymbol{x}}^{(n)}$ will lie outside of $\mathcal{S}$, due to the fact that the polyblock represents an outer approximation.

The upper right boundary $\partial\mathcal{S}$ of $\mathcal{S}$ contains the weakly Pareto efficient points of $\mathcal{S}$:

$$\partial\mathcal{S} = \{\boldsymbol{x} \in \mathcal{S} : \nexists \boldsymbol{x}' \in \mathcal{S} : x'_k > x_k, \forall k\}.$$

Denote by $\boldsymbol{y}^{(n)} \in \partial\mathcal{S}$ the intersection between $\partial\mathcal{S}$ and the line segment connecting the origin with $\check{\boldsymbol{x}}^{(n)}$. Let $\hat{\boldsymbol{x}}^{(n)}$ denote the best intersection point computed so far, i.e.,

$$\hat{\boldsymbol{x}}^{(n)} = \boldsymbol{y}^{(\ell^*)}, \ell^* = \underset{\ell \in \{1,\ldots,n\}}{\operatorname{argmax}} f(\boldsymbol{y}^{(\ell)}).$$

Moreover, let f^* denote the global maximum of (9). From $\hat{\boldsymbol{x}}^{(n)} \in \mathcal{S}$ and $\mathcal{S} \subseteq \mathcal{P}(\mathcal{V}^{(n)})$ it follows that

$$f(\hat{\boldsymbol{x}}^{(n)}) \leq f^* \leq f(\check{\boldsymbol{x}}^{(n)}). \qquad (12)$$

Intuitively, as the outer approximation of $\mathcal{S}$ by a polyblock is refined at each step, $f(\check{\boldsymbol{x}}^{(n)})$ eventually converges to f^*. Due to the continuity of f, this convergence also holds for $\hat{\boldsymbol{x}}^{(n)}$, i.e., $\hat{\boldsymbol{x}}^{(n)}$ converges to a global maximizer of f over $\mathcal{S}$. See [8] for a rigorous proof. According to Eq. (12), an ϵ-optimal solution is found if $f(\hat{\boldsymbol{x}}^{(n)}) \geq f(\check{\boldsymbol{x}}^{(n)}) - \epsilon$.

One possible method to construct a sequence of polyblocks $\mathcal{P}(\mathcal{V}^{(n)})$ that satisfies (10) is as follows [8]: Define

$$\mathcal{K}(\boldsymbol{x}) = \left\{\boldsymbol{y} \in \mathbb{R}_+^D : y_k > x_k, k \in \mathcal{I}(\boldsymbol{x})\right\},$$

with

$$\mathcal{I}(\boldsymbol{x}) = \{k : x_k > 0\}. \qquad (13)$$

The desired rule for constructing a sequence of polyblocks that satisfies (10) is

$$\mathcal{P}(\mathcal{V}^{(n+1)}) = \mathcal{P}(\mathcal{V}^{(n)}) \setminus \mathcal{K}(\hat{\boldsymbol{x}}^{(n)}).$$

The rules for computing the corresponding vertex set $\mathcal{V}^{(n+1)}$ are provided in [8].

Note that the polyblock method relies on a method for computing the intersection points $\boldsymbol{y}^{(n)}$ on the boundary $\partial\mathcal{S}$.

VI. Solution Strategies

A. Concave Utility, Convex Region

Under the assumption that the utility function is concave, the problem

$$\max_{\boldsymbol{r}} u(\boldsymbol{r}) \quad \text{s.t.} \quad \boldsymbol{r} \in \mathcal{C}$$

is not only a monotonic, but also a convex optimization problem. The proposed solution strategy is based on Lagrangian duality [9]. First, introduce additional variables $\boldsymbol{s} \in \mathbb{R}_+^K$, and define the equivalent problem

$$\max_{\boldsymbol{r},\boldsymbol{s}} u(\boldsymbol{s}) \quad \text{s.t.} \quad \boldsymbol{0} \leq \boldsymbol{s} \leq \boldsymbol{r}, \boldsymbol{r} \in \mathcal{C}. \qquad (14)$$

By dualizing the inequality constraint $\boldsymbol{s} \leq \boldsymbol{r}$, the dual function is given by

$$g(\boldsymbol{\lambda}) = \max_{\boldsymbol{s} \geq \boldsymbol{0}} u(\boldsymbol{s}) - \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s} + \max_{\boldsymbol{r} \in \mathcal{C}} \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{r}.$$

Note that evaluating g at $\boldsymbol{\lambda}$ involves solving a *weighted sum rate maximization* (WsrMax) problem. Solving WsrMax problems is discussed in Section VI-D.

The optimum dual variable $\boldsymbol{\lambda}^*$ is found using an outer linearization method [9]. Assuming that $\operatorname{int}\mathcal{C} \neq \emptyset$, strong duality holds and the optimum rate $\boldsymbol{r}^*$ can be recovered from the primal iterates in the outer linearization method [9].

B. Nonconcave Utility, Convex Region

If the utility function u is not concave, the problem

$$\max_{\boldsymbol{r}} u(\boldsymbol{r}) \quad \text{s.t.} \quad \boldsymbol{r} \in \mathcal{C}$$

no longer represents a convex problem. Still, as long as u is increasing, the utility maximization problem represents a monotonic optimization problem and can be solved using the polyblock algorithm.

According to Section V-A, the polyblock algorithm requires a means to compute the intersection between a line segment $\left\{x\check{\boldsymbol{r}}^{(n)} : 0 \leq x \leq 1\right\}$ and $\partial\mathcal{C}$. Computing the intersection point can be formulated as the following optimization problem:

$$\max_{x} \ x \quad \text{s.t.} \quad x\check{\boldsymbol{r}}^{(n)} \in \mathcal{C}. \qquad (15)$$

Lacking a suitable parameterization of $\mathcal{C}$, however, a formulation that is more amendable to a numerical solution approach is as follows:

$$\max_{x,\boldsymbol{r}} \ x \quad \text{s.t.} \quad x\check{\boldsymbol{r}}^{(n)} \leq \boldsymbol{r}, \ \boldsymbol{r} \in \mathcal{C}. \qquad (16)$$

Problem (16) can be solved via Lagrange duality: after dualizing the constraint $x\check{\boldsymbol{r}}^{(n)} \leq \boldsymbol{r}$, all subproblems that involve $\mathcal{C}$ are again weighted sum rate maximization problems, which can be solved using the results from Section VI-D. In particular, (16) is a convex problem, and strong duality holds. Thus, the desired intersection point on the boundary of $\mathcal{C}$ is given by

$$\boldsymbol{y}^{(n)} = x^*\check{\boldsymbol{r}}^{(n)} = d^*\check{\boldsymbol{r}}^{(n)}, \qquad (17)$$

where d^* denotes the dual solution.

C. Nonconvex Region

Due to the fact that $\mathcal{R}$ cannot be assumed to be convex, the problem

$$\max_{\boldsymbol{r}} u(\boldsymbol{r}) \quad \text{s.t.} \quad \boldsymbol{r} \in \mathcal{R}$$

is in general a nonconvex problem, regardless of the properties of u. Accordingly, it is solved with the polyblock algorithm presented in Section V-A.

Analogous to Section VI-B, computing the intersection point between a line segment $\left\{x\check{\boldsymbol{r}}^{(n)} : 0 \leq x \leq 1\right\}$ and $\partial\mathcal{R}$ corresponds to the optimization problem

$$\max_{x} \ x \quad \text{s.t.} \quad x\check{\boldsymbol{r}}^{(n)} \in \mathcal{R}. \qquad (18)$$

In contrast to Section VI-B, however, the problem cannot be solved by Lagrangian duality, due to the possible nonconvexity of $\mathcal{R}$. To each rate vector $\boldsymbol{r} \in \mathcal{R}$ corresponds an SINR vector $\boldsymbol{\gamma}(\boldsymbol{r})$, with $\gamma_k(\boldsymbol{r}) = 2^{r_k} - 1$. In [1], an SINR balancing problem is defined to test whether an SINR vector $\boldsymbol{\gamma}$ is feasible. A rate vector $\boldsymbol{r}$ is achievable if the SINR vector $\boldsymbol{\gamma}(\boldsymbol{r})$ is feasible. As a result, the SINR balancing algorithm from [1] can be used to test for membership in $\mathcal{R}$. Based on this test, (18) is solved by bisection.

D. Weighted Sum Rate Maximization

One method to compute points on $\partial\mathcal{C}$ and $\partial\mathcal{R}$ is the maximization of a weighted sum of rates. As shown in Section VI, WsrMax plays a key role in solving the utility maximization problem (8). Note that

$$\max_{\boldsymbol{r}\in\mathcal{C}} \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{r} = \max_{\boldsymbol{r}\in\mathcal{R}} \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{r}.$$

In other words, for maximization of the weighted sum rate, time sharing is not required. Accordingly, the weighted sum rate problem itself is a problem of type (7) with

$$u(\boldsymbol{r}) = \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{r},$$

and can be solved with the methods discussed in Section VI-C.

Alternatively, the optimum rate vector can be found by determining an optimum power allocation in the dual MAC:

$$\max_{\boldsymbol{p}\in\mathcal{P}} \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{r}(\boldsymbol{p}). \tag{19}$$

The objective function in (19) can be written as a difference of two concave, increasing functions [3]. Accordingly, problem (19) can be solved in either the framework of difference of convex functions optimization [10], or the framework of monotonic optimization.

VII. A Local Method

The solution strategies discussed in Section VI solve the utility maximization problems to global optimality. In all three cases, the polyblock algorithm is used to solve monotonic problems (note that even in the case of the seemingly convex problem (8), the weighted sum rate problems which have to be solved to evaluate the dual function are nonconvex). The computational complexity of the polyblock algorithm grows quickly with the problem dimension – as a result, solving the utility maximization problem to global optimality is practically feasible only for a small number of users. As a result, the proposed global strategies mainly provide a benchmark for local methods, which do not guarantee global optimality, but provide convergence to local optima at much lower computational complexity.

Consider problems of type (7). From the monotonicity of u, it follows that the maximum utility is attained on the boundary $\partial\mathcal{R}$. Based on this observation, it is obvious to search for a local optimum on the boundary $\partial\mathcal{R}$. A special feature of the MISO downlink is the availability of a differentiable parameterization of the boundary $\partial\mathcal{R}$ in terms of the uplink powers $\boldsymbol{p}$:

$$\partial\mathcal{R} = \left\{\boldsymbol{r}(\boldsymbol{p}) : \boldsymbol{p}\in\mathcal{P}, \mathbf{1}^{\mathrm{T}}\boldsymbol{p} = P\right\},$$

i.e., the boundary $\partial\mathcal{R}$ is fully characterized by the rate points that correspond to using full power in the dual uplink [1].

After defining $\mu(\boldsymbol{p}) = u(\boldsymbol{r}(\boldsymbol{p}))$, the utility maximization problem (7) can be equivalently stated as

$$\max \mu(\boldsymbol{p}) \quad \text{s.t.} \quad \boldsymbol{p} \geq \mathbf{0}, \mathbf{1}^{\mathrm{T}}\boldsymbol{p} = P. \tag{20}$$

Under the assumption that u is differentiable, a stationary point of this optimization problem is obtained using Rosen's projected gradient algorithm [9]. The algorithm is initialized with $\boldsymbol{p}^{(0)} = \frac{P}{K}\mathbf{1}$. Due to the nonconvexity of problem (20), there is no guarantee that the corresponding rate vector is globally optimal. On the other hand, the computational complexity of the gradient approach is significantly lower than that of the utility-optimal strategies.

Finally, comparing Equation (20) with (7) shows the importance of finding the right representation of the utility maximization problem. While (7) is a monotonic optimization problem, (20) does not exhibit this structure – the function μ is in general not monotone in $\boldsymbol{p}$, and, except for the special case of weighted sum rate maximization, there is also no obvious representation as a difference of increasing functions. Thus, for an efficient globally optimum solution, it is important to recognize the monotonicity on the rate level, instead of trying to directly optimize the power allocation.

VIII. Simulation Results

For the numerical results, we consider a system utility u which is simply an average of the users' utilities u_k, $u(\boldsymbol{r}) = \frac{1}{K}\sum_{k=1}^{K} u_k(r_k)$. Two different types of utilities u_k are investigated, concave and nonconcave.

A. Concave utility

First, the case of concave utilities u_k is considered. For each user, a logarithmic utility is used:

$$u_k(r_k) = \ln(1 + r_k)$$

Figure 1 shows the average utility achieved by a system with $N = 4$ antennas, $K = 3$ users, and different values of $\mathrm{SNR} = \frac{P}{\sigma^2}$. The channels $\boldsymbol{h}_k$ are drawn from a circularly symmetric complex Gaussian distribution, the channel coefficients are iid. with unit power, and the users' channels are uncorrelated. Figure 1 shows almost no difference in average utility between maximization over $\mathcal{R}$ and $\mathcal{C}$, i.e., on average, time sharing is not necessary. Moreover, the local gradient projection strategy provides the same performance as the globally optimum strategies. Finally, Figure 1 also shows the average utility that is achieved if utility is maximized over the convex hull of the single user points (aka. TDMA with variable slot length).

In Figure 2, the number of users is increased to $K = N = 4$, while all other parameters remain unchanged. The results show a slight gain achieved by time sharing in the fully

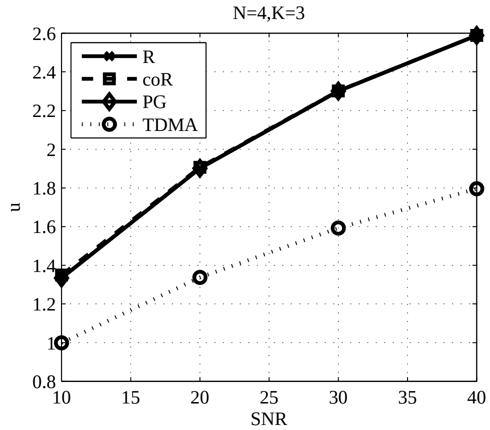

Figure 1. Average utility, concave u, $N = 4$, $K = 3$

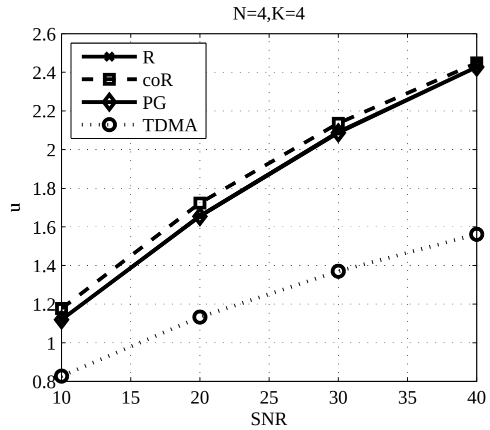

Figure 2. Average utility, concave u, $N = 4$, $K = 4$

loaded system. Moreover, the gradient projection strategy again provides optimum performance.

B. Nonconcave Utility

For the case of nonconcave u_k, we adopt the sigmoidal model from [7]:

$$u_k(r_k) = c_k \left(\frac{1}{1 + \exp(-a_k(r_k - b_k))} + d_k \right).$$

The constants c_k and d_k are chosen such that $u(0) = 0, u(\infty) = 1$, while $a_k = 1(\text{bit/s/Hz})^{-1}$ and $b_k = 3\text{bit/s/Hz}$.

For the results shown in Figure 3, the simulation parameters are the same as in Figure 2, only the utility model was changed. For sigmoidal utilities, the gain achievable by time-sharing is slightly larger. This behaviour can be explained by the fact that the sigmoidal utility makes a steep transition between low and high utility value around the inflection point d_k. As a result, there is a higher sensitivity to the enlargement of the rate region that is provided by time sharing. Even for sigmoidal u_k, the gradient projection method performs as good as the optimum method. Not shown are results for $K = 3$ users, in this case all three methods show almost identical performance.

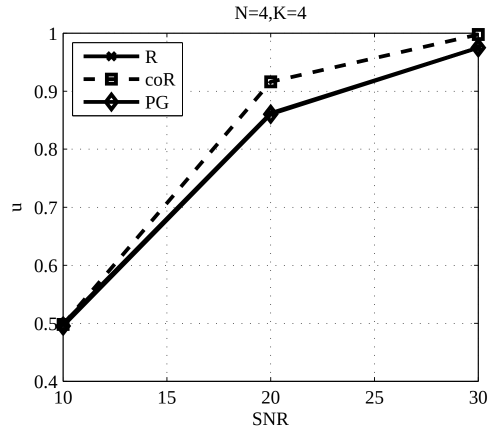

Figure 3. Average utility, sigmoidal u, $N = 4$, $K = 4$

IX. Conclusions

We provided strategies to solve different variations of the utility maximization problem in a multi-user MISO downlink with linear precoding. Most of the optimization problems in the linearly precoded MISO downlink are nonconvex. Still, the utility maximization problem shows a monotonicity structure on the rate level, which we exploited to derive globally optimum solution strategies.

Numerical results show only modest gains by time sharing. Moreover, in the scenarios under consideration, a low complexity locally optimum gradient strategy can provide almost optimum performance. This last conclusion clearly relies on the availability of globally optimum benchmark strategies, which are the main subject of this work.

References

[1] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 1, pp. 18–28, January 2004.

[2] A. Mezghani, M. Joham, R. Hunger, and W. Utschick, "Transceiver design for multi-user MIMO systems," in *Proc. International IEEE/ITG Workshop on Smart Antennas*, March 2006.

[3] M. Schubert and H. Boche, "Throughput maximization for uplink and downlink beamforming with independent coding," in *Proc. Conference on Information Sciences and Systems (CISS)*, March 2003.

[4] E. Jorswieck and E. Larsson, "Linear precoding in multiple antenna broadcast channels: efficient computation of the achievable rate region," in *Proc. International IEEE/ITG Workshop on Smart Antennas*, February 2008.

[5] S. Shenker, "Fundamental design issues for the future internet," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1176–1188, September 1995.

[6] X. Lin, N. B. Shroff, and R. Srikant, "A tutorial on cross-layer optimization in wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, pp. 1452–1463, 2006.

[7] J.-W. Lee, R. Mazumdar, and N. Shroff, "Non-convex optimization and rate control for multi-class services in the internet," *IEEE/ACM Transactions on Networking*, vol. 13, no. 4, pp. 827–840, Aug. 2005.

[8] H. Tuy, "Monotonic optimization: Problems and solution approaches," *SIAM Journal on Optimization*, vol. 11, no. 2, pp. 464–494, 2000.

[9] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming*. Wiley-Interscience, 2006.

[10] H. Tuy, *D.C. optimization: Theory, methods and algorithms*. Kluwer Academic Publishers, 1995, pp. 149–216.