A CONVEX APPROXIMATION METHOD FOR MULTIUSER MISO SUM RATE MAXIMIZATION UNDER DISCRETE RATE CONSTRAINTS

Hoi-To Wai, Qiang Li, Wing-Kin Ma

Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong
{htwai, qli}@ee.cuhk.edu.hk, wkma@ieee.org

ABSTRACT

This paper considers a discrete sum rate maximization (DSRM) problem for transmit optimization in multiuser MISO downlink. Unlike many existing sum rate maximization designs, DSRM focuses on a scenario where each user's achievable rate can only be chosen from a given discrete rate set. This discrete rate-based design is motivated by the fact that practical communication systems can support only a finite number of combinations of modulation and coding schemes. We tackle the DSRM problem first by deriving a novel reformulation of DSRM, in which the discrete rate variables are absorbed by the objective function. Then, from this reformulation, an approximation algorithm based on convex optimization and iterative solution refinement is developed. Simulations results are provided to demonstrate the performance of the proposed algorithm compared with some state-of-the-art algorithms.

***Index Terms*—** discrete sum rate maximization, transmit optimization, transmit beamforming

1. INTRODUCTION

In multiuser MIMO downlink with channel state information at the transmitter, sum rate maximization-based transmit optimization has been a very active research topic. While *dirty paper coding* (DPC) [1] has been proven to achieve the sum capacity under this scenario, the high complexity of DPC prohibits its practical implementation. As a compromise, a more popularly adopted alternative is to treat the inter-user interference as noise and optimize the achievable sum rate. Even so, the resulting sum rate maximization problem is still difficult, and in general NP-hard [2]. In light of this, suboptimal but pragmatic transmit designs have been extensively studied [3–7], and they can yield satisfactory rate performance as simulations have shown. However, the vast majority of the present studies have relied on an implicit assumption— that each user's achievable rate can be adapted to *any* non-negative real number. While it is theoretically possible to achieve a continuous-valued data rate, such scheme is often not realizable in practical communication systems, as the latter can afford to employ only a finite number of modulation and coding schemes (MCSs). For example, in the 3GPP LTE standard [8], there are at most 16 combinations of MCSs with the allowable data rate ranging from 0.15 bits to 5.55 bits per channel use. Simply ignoring this finite discrete rate limitation may lead to significant performance degradation in practical systems.

In this paper, we take the discrete rate constraints explicitly into our transmit design, and consider a discrete sum rate maximization (DSRM) problem for multiuser MISO downlink. While the sum rate maximization problem is already hard to solve, this is even more so with DSRM owing to the incorporation of combinatorial constraints. To the best of our knowledge, the idea of DSRM formulation for accommodating the practical finite rate constraints was first proposed in [9] for multiuser MISO downlink. There, the authors use a mixed-integer second order cone program (MI-SOCP) formulation, and then employ a branch-and-bound method to handle the problem. However, as solving the MI-SOCP entails a worst-case exponential complexity, the authors propose another method where a realizable low-complexity heuristic was suggested.

This paper also aims at developing a pragmatic approach for the DSRM problem. First, we derive a novel equivalent formulation of DSRM. The advantage of the new formulation is that it involves only continuous variables, and the discrete rate variables are absorbed into the formulation itself. Then, based on the new formulation, we derive a convex approximation to DSRM and an iterative procedure to refine the approximate solution. Numerical results demonstrate that the proposed algorithm outperforms the state-of-the-art algorithms in terms of both the sum rate and the transmit power.

Before delving into the details of our method, we should briefly discuss related or prior work. Sum rate maximization has attracted much interest not only in the MISO or MIMO multiuser downlink scenario, but also in related scenarios such as dynamic spectrum management [2] and power control in wireless networks [10]. As such, one can find numerous solution methods in the literature; e.g., see [2–7, 10] and the references therein. For the scenario considered here, we should mention zero-forcing beamforming with user selection [3], block diagonalization [7], uplink-downlink duality [4] and weighted MMSE (WMMSE) minimization [5,6], to name a few. However, as mentioned previously, we see much fewer work that incorporates finite discrete rate constraints. In dynamic spectrum management, DSRM was considered in [11,12]. The work closest to ours is that by Cheng *et. al* [9], which has been reviewed above. The two studies are nevertheless different— Cheng *et. al* choose to handle the discrete rate variables explicitly, while we intentionally circumvent them through a reformulation, as our endeavor to provide a practically efficient transmit design. Moreover, we should note that our method may be regarded as a conceptual generalization of our previous work on joint admission control and beamforming [13].

2. PROBLEM STATEMENT

We consider a standard unicast multiuser MISO downlink scenario (see, e.g., [3]), where a multi-antenna base station (BS) simultaneously transmits K independent messages to K single-antenna users, each message for one user. Under this setting, the signal received by user k at time t may be described as

$$y_k(t) = \boldsymbol{h}_k^H \boldsymbol{x}_k(t) + \sum_{m \neq k} \boldsymbol{h}_k^H \boldsymbol{x}_m(t) + n_k(t),$$

This work was supported by a Direct Grant by the Chinese University of Hong Kong (Project ID 2050506).

where $(\cdot)^H$ denotes the Hermitian transpose, $\boldsymbol{h}_k \in \mathbb{C}^N$ is the channel response from the BS to the kth user; N is the number of antennas employed by the BS; $n_k(t) \sim \mathcal{CN}(0, \sigma_k^2)$ is additive complex Gaussian noise with mean zero and variance σ_k^2; and $\boldsymbol{x}_k(t) \in \mathbb{C}^N$ is the transmit signal intended for user k. Denote $\boldsymbol{W}_k = \mathrm{E}\{\boldsymbol{x}_k(t)\boldsymbol{x}_k(t)^H\}$ as the covariance matrix of the transmit signal $\boldsymbol{x}_k(t)$, and assume that the inter-user interference is treated as independently distributed Gaussian noise at the user. Then, the information rate for user k may be formulated as [14]

$$R_k = \log_2\Big(1 + \frac{\boldsymbol{h}_k^H \boldsymbol{W}_k \boldsymbol{h}_k}{\sigma_k^2 + \sum_{l\neq k} \boldsymbol{h}_k^H \boldsymbol{W}_l \boldsymbol{h}_k}\Big),\ k = 1, \ldots, K,$$

which is achievable when every input signal $\boldsymbol{x}_k(t)$ follows a complex Gaussian distribution, i.e., $\boldsymbol{x}_k(t) \sim \mathcal{CN}(0, \boldsymbol{W}_k)$. Notice that for the case where $\boldsymbol{W}_k = \boldsymbol{w}_k\boldsymbol{w}_k^H$ for some $\boldsymbol{w}_k \in \mathbb{C}^N$, or equivalently, $\mathrm{rank}(\boldsymbol{W}_k) \leq 1$, the physical transmit strategy reduces to *transmit beamforming*; viz., $\boldsymbol{x}_k(t) = \boldsymbol{w}_k s_k(t)$ where $\boldsymbol{w}_k$ is the beamforming vector and $s_k(t)$ is a data stream that carries information for user k. In this study, we consider a general transmit covariance structure, in which we do not put rank constraints on $\boldsymbol{W}_k$, or assume transmit beamforming. It is however interesting to note that in our ensuing development, we will show that our problem automatically leads to rank-one solutions with $\boldsymbol{W}_k$.

Our problem is to maximize the weighted sum rate of all the users, given that their achievable rates R_k are chosen from a predetermined *discrete rate set*. Mathematically, this discrete sum rate maximization (DSRM) problem may be formulated as [9]

$$\max_{\{\boldsymbol{W}_k, R_k\}_{k=1}^K} \sum_{k=1}^K \lambda_k R_k \tag{1a}$$

$$\text{s.t.} \sum_{k=1}^K \mathrm{Tr}(\boldsymbol{W}_k) \leq P, \boldsymbol{W}_k \succeq \boldsymbol{0},\ \forall k \in \mathcal{K}, \tag{1b}$$

$$R_k \leq \log_2\Big(1 + \frac{\boldsymbol{h}_k^H \boldsymbol{W}_k \boldsymbol{h}_k}{\sigma_k^2 + \sum_{l\neq k} \boldsymbol{h}_k^H \boldsymbol{W}_l \boldsymbol{h}_k}\Big), \tag{1c}$$

$$R_k \in \{R^0, R^1, \ldots, R^M\},\ \forall k \in \mathcal{K}, \tag{1d}$$

where $\mathcal{K} \triangleq \{1, \ldots, K\}$, $\lambda_k > 0$ is the priority weight for user k, $P > 0$ is the transmit power budget at the BS and $\{R^0, R^1, \ldots, R^M\}$ denotes the discrete rate set in which we assume $0 = R^0 < R^1 < \cdots < R^M$. The discrete rate set depends on modulation and coding schemes offered by the system, see [9] for more descriptions. As can be seen above, the DSRM formulation constraints the achievable rates R_k to lie in the discrete rate set $\{R^0, R^1, \ldots, R^M\}$.

The DSRM problem is difficult to solve as it involves a joint optimization of discrete and continuous variables. In [9], the authors handle essentially the same problem using a mixed-integer formulation. In the next section, we will describe our method which uses a different formulation and solution approach.

3. DSRM BY CONVEX APPROXIMATION

Our endeavor to attack DSRM is based on a convex approximation approach, which consists of an equivalent reformulation of the DSRM problem (1), a convex approximation formulation of the reformulated DSRM problem, and an iterative solution refinement procedure. They are presented in the following subsections.

3.1. Reformulation of the DSRM Problem (1)

Observe that in problem (1), the constraints (1c) specify whether the chosen data rates R_k are supportable for a given set of transmit covariances. The key of our reformulation is to appropriately formulate this supportability without explicitly choosing the discrete variables R_k. To this end, let us denote the following function

$$r_k(\{\boldsymbol{W}_l\}_{l=1}^K, R^i) \triangleq R^i\left(1 + \frac{1}{\sigma_k^2}\boldsymbol{h}_k^H\Big(\sum_{l\neq k}\boldsymbol{W}_l - \frac{1}{2^{R^i}-1}\boldsymbol{W}_k\Big)\boldsymbol{h}_k\right),$$

for $k = 1, \ldots, K$. It can be shown that $r_k(\{\boldsymbol{W}_l\}_{l=1}^K, R^i)$ is strictly increasing with respect to (w.r.t.) R^i. Moreover, the following equivalence holds

$$\text{(1c) holds} \quad \Longleftrightarrow \quad r_k(\{\boldsymbol{W}_l\}_{l=1}^K, R_k) \leq 0.$$

By the above equivalence, we have the following claim:

Claim 1 *The following problem is equivalent to the DSRM problem* (1)*:*

$$\max_{\{\boldsymbol{W}_k\}_{k=1}^K} \sum_{k=1}^K \lambda_k \sum_{i=1}^M (R^{i-1} - R^i)\ f_k^i(\{\boldsymbol{W}_l\}_{l=1}^K) \tag{2a}$$

$$\text{s.t.} \sum_{k=1}^K \mathrm{Tr}(\boldsymbol{W}_k) \leq P,\ \boldsymbol{W}_k \succeq \boldsymbol{0},\ \forall k \in \mathcal{K}, \tag{2b}$$

where

$$f_k^i(\{\boldsymbol{W}_l\}_{l=1}^K) \triangleq \mathrm{card}\left(\max\left\{0, r_k(\{\boldsymbol{W}_l\}_{l=1}^K, R^i)\right\}\right), \tag{3}$$

and $\mathrm{card}(x) = 0$ *if* $x = 0$*, and* $\mathrm{card}(x) = 1$*, otherwise.*

We should point out that the main advantage of the reformulation (2) is that it involves only the continuous decision variables $\{\boldsymbol{W}_k\}$, and the discrete rate selection is done automatically in the process of solving problem (2).

Proof of Claim 1: As $0 = R^0 < \ldots < R^M$ and $r_k(\{\boldsymbol{W}_l\}_{l=1}^K, R^i)$ is strictly increasing w.r.t. R^i, for every feasible solution $\{\boldsymbol{W}_l\}_{l=1}^K$ of problem (2), there exists an index $0 \leq i_k \leq M$ such that $r_k(\{\boldsymbol{W}_l\}_{l=1}^K, R^i) \leq 0$ for $i = 0, \ldots, i_k$, and $r_k(\{\boldsymbol{W}_l\}_{l=1}^K, R^i) > 0$ for all other i's. As a result, we have

$$f_k^i(\{\boldsymbol{W}_l\}_{l=1}^K) = \begin{cases} 0, & \text{if } i \leq i_k, \\ 1, & \text{otherwise.} \end{cases}$$

By substituting $f_k^i(\{\boldsymbol{W}_l\}_{l=1}^K)$ into (2a), the objective function is simplified to

$$\sum_{k=1}^K \lambda_k R^{i_k} - \sum_{k=1}^K \lambda_k R^M. \tag{4}$$

The latter term $\sum_{k=1}^K \lambda_k R^M$ can be ignored as it is a constant. From (4), one can see that every feasible solution $\{\boldsymbol{W}_l\}_{l=1}^K$ of problem (2) corresponds to a supportable weighted sum rate. Hence, an optimal solution of (2) is the one that yields the maximal weighted sum rate, which is exactly the goal of the DSRM problem (1). ■

From this point on, we will focus on the equivalent DSRM formulation in (2). Upon careful examination of problem (2), it can be shown that the DSRM problem may have multiple optimal solutions, where some are allowed to use higher transmit powers. As minimal transmit power under the same sum rate performance is desired, we consider a slightly modified version of problem (2) which is stated below:

Observation 1 *Let $0 < \epsilon < \min_{k,i}(\lambda_k(R^i - R^{i-1})/P)$. Consider the following problem*

$$\begin{aligned} \max_{\{\boldsymbol{W}_k\}_{k=1}^K} \quad & \sum_{k=1}^K \lambda_k \sum_{i=1}^M (R^{i-1} - R^i)\, f_k^i(\{\boldsymbol{W}_l\}_{l=1}^K) - \epsilon \sum_{k=1}^K \mathrm{Tr}(\boldsymbol{W}_k) \\ \text{s.t.} \quad & \sum_{k=1}^K \mathrm{Tr}(\boldsymbol{W}_k) \le P,\ \boldsymbol{W}_k \succeq \boldsymbol{0},\ \forall k \in \mathcal{K}. \end{aligned} \quad (5)$$

Then, any optimal solution of problem (5) *is an optimal solution of* (2) *with the minimum transmit power.*

The proof of Observation 1 is given in Appendix 6.1.

3.2. Convex Approximation to the DSRM Problem (5)

The equivalent DSRM problem in (5) is difficult to solve. In particular, the difficulty comes from the (nonconvex) cardinality function in (3). Our approach is to ignore the cardinality function, which is reminiscent of the widely-adopted ℓ_1 approximation approach in compressive sensing [15]. To be more specific, (3) is replaced by

$$\tilde{f}_k^i(\{\boldsymbol{W}_l\}_{l=1}^K) \triangleq \max\left\{0, r_k(\{\boldsymbol{W}_l\}_{l=1}^K, R^i)\right\}. \quad (6)$$

This leads to the following approximation to problem (5):

$$\begin{aligned} \max_{\{\boldsymbol{W}_k\}_{k=1}^K} & \sum_{k=1}^K \lambda_k \sum_{i=1}^M (R^{i-1} - R^i)\, \tilde{f}_k^i(\{\boldsymbol{W}_l\}_{l=1}^K) - \tilde{\epsilon} \sum_{k=1}^K \mathrm{Tr}(\boldsymbol{W}_k) \\ \text{s.t.} & \sum_{k=1}^K \mathrm{Tr}(\boldsymbol{W}_k) \le P, \boldsymbol{W}_k \succeq \boldsymbol{0},\ \forall k \in \mathcal{K}, \end{aligned} \quad (7)$$

where $\tilde{\epsilon}$ is a scaled penalty parameter to be described below. One can verify that problem (7) is convex (recalling that $R^{i-1} - R^i < 0$). Furthermore, by introducing slack variables μ_k^i to replace $\tilde{f}_k^i(\{\boldsymbol{W}_l\}_{l=1}^K)$ and changing the sign of the objective function, (7) can be rewritten as the following semidefinite program (SDP):

$$\begin{aligned} \min_{\{\boldsymbol{W}_k, \{\mu_k^i\}\}_{k=1}^K} \quad & \sum_{k=1}^K \lambda_k \sum_{i=1}^M (R^i - R^{i-1})\, \mu_k^i + \tilde{\epsilon} \sum_{k=1}^K \mathrm{Tr}(\boldsymbol{W}_k) \\ \text{s.t.} \quad & \sum_{k=1}^K \mathrm{Tr}(\boldsymbol{W}_k) \le P, \boldsymbol{W}_k \succeq \boldsymbol{0},\ \forall k \in \mathcal{K}, \\ & \mu_k^i \ge 0,\ \mu_k^i \ge r_k(\{\boldsymbol{W}_l\}_{l=1}^K, R^i),\ \forall\, i,\ k, \end{aligned} \quad (8)$$

which is efficiently solvable using off-the-shelf optimization softwares, e.g., CVX [16].

It should be noted that problem (7) may not be a tight approximation to problem (5). Hence, after solving (7), a further solution refinement is necessary. We will elaborate on this in the next subsection. At this point, some important observations on problem (8) are described as follows.

Observation 2 With $\tilde{\epsilon} = L\epsilon$ and $L \triangleq R^M[1 + P\max_k(\boldsymbol{h}_k^H \boldsymbol{h}_k)/\sigma_k^2]$, problem (8) is a convex relaxation of the DSRM problem (5).

The proof of Observation 2 is omitted here due to space limitation. Essentially, we observe that (5) can be modeled as a $\{0,1\}$-mixed integer program as the cardinality function $f_k^i(\cdot)$ takes value from either 0 or 1. It follows that we can apply a continuous relaxation to the $\{0,1\}$ constraint and obtain a convex-relaxed problem (see, e.g., [17]), which can further be shown to be equivalent to (8). It should be noted that for the conventional continuous sum rate maximization problem, no convex relaxation has been reported in the literature to our best knowledge. Curiously, with DSRM, we are able to derive a convex relaxation.

Proposition 1 *For $\tilde{\epsilon} > 0$, any optimal solution $\{\boldsymbol{W}_k^\star\}_{k=1}^K$ of problem* (8) *must satisfy* $\mathrm{rank}(\boldsymbol{W}_k^\star) \le 1$ *for all $k \in \mathcal{K}$.*

The proof is given in Appendix 6.2. This result implies that the optimal transmit strategy found by (8) must be *transmit beamforming*.

3.3. An Iterative Refinement for DSRM

In this subsection, we propose a DSRM solution generation procedure for the convex approximation concept derived above. The key insight behind is that if a given rate profile $(R^{I_1}, R^{I_2}, \ldots, R^{I_K})$ is supportable, then with a sufficiently small $\tilde{\epsilon}$, the optimal solution to the following truncated problem of (8) (notice the change with the inner summation)

$$\begin{aligned} \min_{\{\boldsymbol{W}_k, \{\mu_k^i\}\}_{k=1}^K} \quad & \sum_{k=1}^K \lambda_k \sum_{i=1}^{I_k} (R^i - R^{i-1})\, \mu_k^i + \tilde{\epsilon} \sum_{k=1}^K \mathrm{Tr}(\boldsymbol{W}_k) \\ \text{s.t.} \quad & \textstyle\sum_{k=1}^K \mathrm{Tr}(\boldsymbol{W}_k) \le P, \boldsymbol{W}_k \succeq \boldsymbol{0},\ \forall k \in \mathcal{K}, \\ & \mu_k^i \ge 0,\ \mu_k^i \ge r_k(\{\boldsymbol{W}_l\}_{l=1}^K, R^i),\ \forall\, i,\ k \end{aligned} \quad (9)$$

is precisely a set of transmit covariance matrices that supports the same rate profile $(R^{I_1}, R^{I_2}, \ldots, R^{I_K})$. This observation suggests us to refine the solution of (7) by gradually lowering the target rate profile and solving (9) iteratively. The iterative refinement procedure is summarized below. Note that the iterative refinement procedure involves solving at most KM SDPs throughout the process.

- **Initialize**: A target rate profile: $R^{I_1} = \ldots = R^{I_K} = R^M$.

1. **(Transmit optimization)** Solve problem (9) under the profile $(R^{I_1}, R^{I_2}, \ldots, R^{I_K})$. If $r_k(\{\boldsymbol{W}_l\}_{l=1}^K; R^{I_k}) \le 0$ for all $k \in \mathcal{K}$, save $\{\boldsymbol{W}_k\}$ and terminate; otherwise go to Step 2.
2. **(Iterative refinement)** Select user k with the maximum $r_k(\{\boldsymbol{W}_l\}_{l=1}^K; R^{I_k})/R^{I_k}$ and decrease its maximum supportable rate by $I_k = I_k - 1$. Return to Step 1.

- **Return**: $\{\boldsymbol{W}_k\}$, $(R^{I_1}, R^{I_2}, \ldots, R^{I_K})$ as a feasible DSRM solution.

4. NUMERICAL EXAMPLES

In this section, we test the performance of the proposed DSRM algorithm and compare it to some state-of-the-art algorithms, namely Algorithm 1 in [9] and the WMMSE algorithm [5]. The former is a low-complexity method for DSRM, and the latter is originally designed for the conventional weighted sum rate maximization problem with unconstrained rate profile. As a benchmark, we will also present the result of WMMSE after quantizing the continuous rates to the discrete rate set with floor quantization. The simulation settings are as follows: The number of antennas at the base station is $N = 6$; there are $K = 8$ users in the cell. Following [9], we choose the discrete rate set to be that specified in 3GPP LTE standard [8], i.e., $\{0, 0.15, 0.23, 0.38, 0.6, 0.88, 1.18, 1.48, 1.91, 2.41, 2.73, 3.32, 3.90, 4.52, 5.11, 5.55\}$ (in bits). The scaled penalty parameter $\tilde{\epsilon}$ is set as $10^{-4}/P$. For simplicity, we set $\sigma_k^2 = 1$ and $\lambda_k = 1$ for all k. All the channel coefficients are randomly generated following an i.i.d. complex Gaussian distribution with zero mean and unit variance, and all results were averaged over 1,000 channel trials. We remark that exact solution methods for problem (1), e.g., the branch-and-bound method in [9], are computationally too slow to run in this scenario. Here, the unquantized WMMSE algorithm may serve as an upper bound to the sum rate performance.

In the first example, we investigate the sum rate performance of the various methods when changing the normalized transmit power budget P. The results are plotted in Fig. 1. As seen, the performance

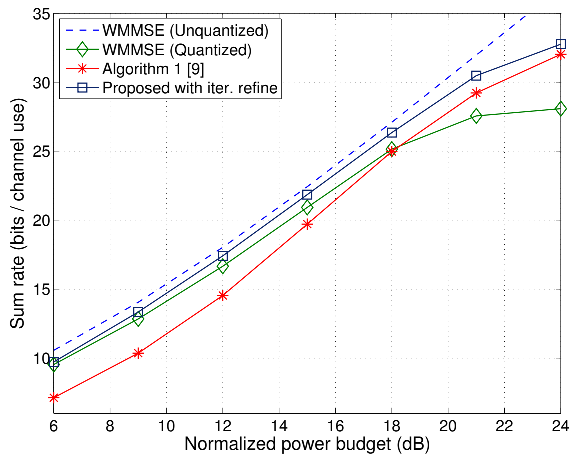

Fig. 1. Achievable sum rate against the total power budget.

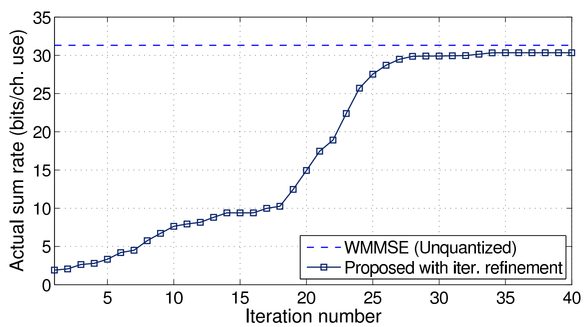

Fig. 2. Achievable sum rate against iteration no. for $P = 21$ dB.

of the proposed algorithm with iterative refinement is very close to that of the unquantized WMMSE, and outperforms the quantized WMMSE and Algorithm 1 of [9] over the whole range of powers tested. However, we should mention that the proposed algorithm has a slightly higher computational complexity than Algorithm 1 in [9]. There is a trade-off between complexity and sum rate performance.

In the second example, we fix the power budget as $P = 21$ dB and plot the sum rate obtained by the proposed iterative refinement against the iteration number, for one random channel realization. For each iteration, the resulting discrete sum rate is obtained by quantizing the rates returned by solving (9). The results are shown in Fig. 2. It can be seen that without the iterative refinement (i.e., the iteration number being zero), the discrete sum rate obtained from the solution of (8) is far less than that of the unquantized WMMSE. However, as the iteration number increases, iterative refinement improves the sum rate gradually, and approaches the unquantized WMMSE.

The last example compares the transmit powers of the proposed algorithm and Algorithm 1 under the same setting as in Fig. 1. The results are shown in Table 1. We can see that the proposed algorithm requires less transmit power, and meanwhile attains higher sum rates than Algorithm 1 (cf. Fig. 1 for the rate comparison). The proposed algorithm is thus more power efficient under the tested scenario.

5. CONCLUSION

In this paper we have considered a discrete weighted sum rate maximization (DSRM) problem, and proposed a pragmatic approach to handling this challenging problem. In particular, we have developed a novel reformulation of DSRM. Based on this formulation, a new convex approximation-based algorithm was developed. Simulation results confirmed the efficacy of the proposed approach.

Table 1. Minimum transmit power for different power budgets.

Normalized power budget	Proposed	Algorithm 1 [9]
6 dB	**5.72 dB**	5.97 dB
15 dB	**14.78 dB**	14.93 dB
27 dB	**23.11 dB**	24.77 dB

6. APPENDIX

6.1. Proof of Observation 1

Our proof follows an idea similar to the proof of Claim 1 in [18]. First, it is obvious that any optimal solution to (5) is power minimal as the penalty term $\epsilon \sum_{k=1}^{K} \mathrm{Tr}(\boldsymbol{W}_k)$ corresponds to the transmit power. It remains to prove that any optimal solution to (5) is also optimal to (2).

In the following, we omit the dependence of $f_k^i(\cdot)$ on $\{\boldsymbol{W}_k\}_{k=1}^K$ for notational convenience. Let $\upsilon(\{f_k^i\}) \triangleq \sum_{k=1}^{K} \lambda_k \sum_{i=1}^{M} (R^{i-1} - R^i) f_k^i$ and $\{\boldsymbol{W}_k^\star, (f_k^i)^\star\}$ denote the objective and an optimal solution of (2), respectively. A key observation on (2) is that the objective function $\upsilon(\{f_k^i\})$ takes on discrete values as f_k^i is the cardinality function such that $f_k^i \in \{0, 1\}$. The minimum 'step size' is $\min_{k,i} \lambda_k (R^i - R^{i-1})$ (cf. (3)). Now, suppose on the contrary that there exists an optimal solution to (5), say $\{\overline{\boldsymbol{W}}_k, \overline{f}_k^i\}$, that is not optimal to (2). Then we have:

$$\upsilon(\{\bar{f}_k^i\}) \leq \upsilon(\{(f_k^i)^\star\}) - \min_{k,i} \lambda_k (R^i - R^{i-1}). \tag{10}$$

Furthermore, for $0 < \epsilon < \min_{k,i} (\lambda_k (R^i - R^{i-1})/P)$, we have

$$\epsilon \textstyle\sum_{k=1}^{K} \mathrm{Tr}(\boldsymbol{W}_k) < \min_{k,i} \lambda_k (R^i - R^{i-1}). \tag{11}$$

Combining (10) with (11), we get:

$$\upsilon(\{\bar{f}_k^i\}) - \epsilon \sum_{k=1}^{K} \mathrm{Tr}(\overline{\boldsymbol{W}}_k) < \upsilon(\{(f_k^i)^\star\}) - \epsilon \sum_{k=1}^{K} \mathrm{Tr}(\boldsymbol{W}_k^\star), \tag{12}$$

where we have exploited the fact that $\upsilon(\{\bar{f}_k^i\}) - \epsilon \sum_{k=1}^{K} \mathrm{Tr}(\overline{\boldsymbol{W}}_k) \leq \upsilon(\{\bar{f}_k^i\})$ to obtain a lower bound on the left hand side of (10). However, inequality (12) contradicts the optimality of $\{\overline{\boldsymbol{W}}_k, \overline{f}_k^i\}$.

6.2. Proof of Proposition 1

Let $\{\boldsymbol{W}_k^\star\}$ be an optimal solution to (8), and denote $(z_k^i)^\star \geq 0$, $\upsilon^\star \geq 0$ and $\boldsymbol{X}_k^\star \succeq \boldsymbol{0}$ as the dual variables associated with the constraint $\mu_k^i \geq r_k(\{\boldsymbol{W}_l\}_{l=1}^K, R^i)$, the power budget constraint and the positive semidefinite constraint, respectively. The first order necessary optimality condition of (7) can be shown to be

$$\boldsymbol{X}_k^\star = \tilde{\epsilon}\boldsymbol{I} + \sum_{m \neq k} \sum_{i=1}^{M} (z_m^i)^\star R^i \boldsymbol{h}_m \boldsymbol{h}_m^H - \Big(\sum_{i=1}^{M} \frac{(z_k^i)^\star R^i}{2^{R^i} - 1} \Big) \boldsymbol{h}_k \boldsymbol{h}_k^H, \tag{13}$$

for all k, and the complementary slackness condition is:

$$\boldsymbol{X}_k^\star \boldsymbol{W}_k^\star = \boldsymbol{0}, \ \forall\, k. \tag{14}$$

Condition (13) and (14) suffice to prove our claim in Proposition 1. Observe that for $\tilde{\epsilon} > 0$, $\tilde{\epsilon}\boldsymbol{I}$ is of full-rank and $\boldsymbol{h}_k \boldsymbol{h}_k^H$ is of rank one. Hence, from (13), the rank of $\boldsymbol{X}_k^\star$ must be greater than or equal to $N - 1$. It then follows from (14) that $\mathrm{rank}(\boldsymbol{W}_k^\star) \leq 1$ for all k.

7. REFERENCES

[1] H. Weingarten, Y. Steinberg, and S. Shamai, “The capacity region of the Gaussian multiple-input multiple-output broadcast channel,” *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[2] Z.-Q. Luo and S. Zhang, “Dynamic spectrum management: complexity and duality,” *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–73, Feb. 2008.

[3] 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.

[4] S. Shi, M. Schubert, and H. Boche, “Rate optimization for multiuser MIMO systems with linear processing,” *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 4020–4030, Aug. 2008.

[5] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, “An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel,” *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[6] S. S. Christensen, R. Agarwal, E. Carvalho, and J. Cioffi, “Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.

[7] Q. Spencer, A. Swindlehurst, and M. Haardt, “Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels,” *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[8] S. Sesia, M. Baker, and I. Toufik, *LTE - The UMTS Long Term Evolution: From Theory to Practice*. John Wiley & Sons, 2011.

[9] Y. Cheng, A. Philipp, and M. Pesavento, “Dynamic rate adaptation and multiuser downlink beamforming using mixed integer conic programming,” in *Proc. EUSIPCO’ 2012*, Aug. 2012, pp. 824 –828.

[10] 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.

[11] C. Y. Wong, R. S. Cheng, K. B. Lataief, and R. D. Murch, “Multiuser OFDM with adaptive subcarrier, bit, and power allocation,” *IEEE J. Sel. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.

[12] M. Wolkerstorfer, J. Jaldén, and T. Nordstrom, “Column generation for discrete-rate multi-user and multi-carrier power control,” *IEEE Trans. Commun.*, vol. 60, no. 9, pp. 2712–2722, Sep. 2012.

[13] H.-T. Wai and W.-K. Ma, “A decentralized method for joint admission control and beamforming in coordinated multicell downlink,” in *Proc. Asilomar Conf. 2012*, Nov. 2012. [Online]. Available: http://arxiv.org/abs/1211.4665

[14] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley-Interscience, 1991.

[15] E. J. Candes and T. Tao, “Decoding by linear programming,” *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[16] I. CVX Research, “CVX: Matlab software for disciplined convex programming, version 2.0 beta,” http://cvxr.com/cvx, Sep. 2012.

[17] A. Aspremont and S. Boyd, *Relaxations and Randomized Methods for Nonconvex QCQPs*. Stanford, CA: Stanford Univ., 2003.

[18] E. Matskani, N. Sidiropoulos, L. Tassiulas, and Z.-Q. Luo, “Convex approximation techniques for joint multiuser downlink beamforming and admission control,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2682–2693, Jul. 2008.