Research Article

Feedback Reduction in Uplink MIMO OFDM Systems by Chunk Optimization

Eduard Jorswieck,[1] Aydin Sezgin,[2] Björn Ottersten,[1] and Arogyaswami Paulraj[2]

[1] *ACCESS Linnaeus Center, Electrical Engineering, KTH - Royal Institute of Technology, 100 44 Stockholm, Sweden*
[2] *Information Systems Laboratory, Stanford University, CA 94305, USA*

Correspondence should be addressed to Eduard Jorswieck, eduard.jorswieck@ee.kth.se

The performance of multiuser MIMO systems can be significantly increased by channel-aware scheduling and signal processing at the transmitters based on channel state information. In the multipleantenna uplink multicarrier scenario, the base station decides centrally on the optimal signal processing and spectral power allocation as well as scheduling. An interesting challenge is the reduction of the overhead in order to inform the mobiles about their transmit strategies. In this work, we propose to reduce the feedback by chunk processing and quantization. We maximize the weighted sum rate of a MIMO OFDM MAC under individual power constraints and chunk size constraints. An efficient iterative algorithm is developed and convergence is proved. The feedback overhead as a function of the chunk size is considered in the rate computation and the optimal chunk size is determined by numerical simulations for various channel models. Finally, the issues of finite modulation and coding schemes as well as quantization of the precoding matrices are addressed.

1. INTRODUCTION

The exploitation of channel state information (CSI) at the transmitter in wireless systems has been a highly active research area. This transmit CSI can significantly improve the performance and reliability of multiple antenna single-user as well as multiuser systems [1–3]. Utilizing this information effectively is one of the major challenges in future mobile communication systems like, for example, WiMAX. Multiple input multiple output (MIMO) multiple access channels (MAC) and broadcast channels (BC) utilizing cyclic prefix orthogonal frequency division multiplexing (CP-OFDM) are a central part of WiMAX. Thus optimal transmit strategies that optimize the performance of such systems were proposed in [4–7].

Until recently, a lot of attention was given to single-user MIMO systems, which is changing nowadays. The paradigm shift from single-user MIMO to multiuser MIMO is highlighted in [8, 9]. Most recent work on multiuser sytems concentrates on the BC, that is, the downlink. Recently, the weighted sum rate optimization is studied for flat-fading MIMO systems in [10] and the extension to MIMO OFDM

is developed in [11]. An overview of different linear precoding schemes for the MIMO BC is given in [12]. The question about the amount of feedback has been raised for the BC in [13]. Regarding the uplink channel, finite rate feedback is studied in [14] for the multiple-antenna case, and the average throughput is analyzed in [15].

Depending on the system under consideration, either perfect CSI or long-term CSI is assumed to be available at the transmitter in order to derive the optimal precoding strategy. Under perfect [4, 16, 17] and long-term CSI [18, 19], the optimal linear precoding matrices are found at the central base station by convex optimization. Then the linear precoding matrices can be applied to up- and downlink by the duality theory [20, 21]. With imperfect CSI at the transmitter, the duality theory does not hold any longer [22].

In the uplink scenario with centralized channel-aware scheduling at the base station, which is considered in this paper, one important issue is to inform the mobiles about their precoding strategies with limited amount of feedback. The more information is needed at the transmitter and the more this information has been exact, the more feedback is required.
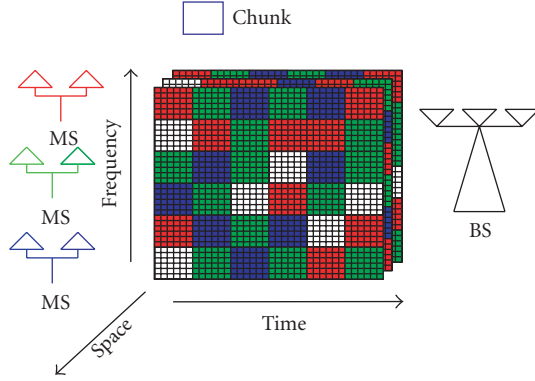
FIGURE 1: Multiuser OFDM MIMO MAC with chunk processing.

In MIMO OFDM systems, this control overhead and signal processing complexity are quite large, leading to the definition of the so-called time-frequency tiles or chunks [23]. To this end, the physical channel structure divides the available time-frequency resources into tiles. The tiles or chunks are considered two dimensional, and each chunk comprises a number of adjacent subcarriers in frequency domain and a number of consecutive OFDM symbols in time domain as illustrated in Figure 1. The application of chunks is wide spread and it is proposed, for example, in [23] for multiple antenna systems. For all subcarriers and all OFDM symbols within the chunk, the same spatial signal processing is applied, reducing the signal processing complexity and the feedback overhead considerably.

In the single input single output (SISO) case, the power and rate control in each chunk has to be optimized and the performance decreases as the chunk size increases. Numerical evidence of this fact has been provided in [24]. Whereas in the MIMO case, the spatial signal processing, that is, the linear precoding has to be optimized per chunk. If the channel is flat within one chunk, the original optimization for the single carrier case can be reused. However, it turns out that the MIMO channel matrices within one chunk vary even at small chunk sizes. This motivates the detailed analysis.

In this paper, the following contributions are made to the problem of resource allocation in OFDM-SDMAW(We call this technique OFDM-SDMA because multiple users can be allocated simultaneously to different chunks over time and frequency domains.) uplink systems under limited feedback.

(1) We formulate the weighted sum rate maximization under individual power constraints and under the assumption that only one linear precoding matrix per chunk is fedback.

(2) The programming problem is solved by an efficient iterative algorithm based on an inner fix-point algorithm and outer iterative water filling. The convergence of the proposed algorithm is proved.

(3) The tradeoff between performance and feedback overhead is analyzed by formulating an effective transmission rate that takes the amount of feedback directly into account.

(4) The effective transmission rate is illustrated for different channel models (ideal, IEEE 802.11n [25], WIM2 [26]).

(5) Finally, the framework is extended to cope with finite modulation and coding schemes as well as finite quantization of the linear precoding matrices.

The first part of the paper restricts itself to the weighted sum rate maximization under chunk constraints. The system model, the limited feedback model, and the problem statements are described in Section 2. In Section 3, the optimization theoretic framework is developed and convergence proved. This is done first for the single-user case and then for the multiuser scenario. The implication of the results on the MIMO-OFDM MAC system design is discussed with respect to limited feedback, limited modulation and coding schemes (MCS), and quantized linear precoding in Section 4. In Section 5, numerical results illustrate the performance. The paper is concluded in Section 6 and further application and open problems are discussed. The appendices contain the proofs.

1.1. Notation and symbols

Vectors are denoted by boldface small letters $\mathbf{a}, \mathbf{b}$, and matrices by boldface capital letters $\mathbf{A}, \mathbf{B}$. $\mathbf{A}^T$, $\mathbf{A}^H$, and $\mathbf{A}^{-1}$ are the transpose, the conjugate transpose, and the inverse matrix operation, respectively. The identity matrix is $\mathbf{I}$, and $\mathbf{1}$ is the vector with all ones. $\mathbf{A}^{1/2}$ is the square root matrix of $\mathbf{A}$ and $[\mathbf{A}]_{j,k}$ denotes the entry in the jth row and the kth column of $\mathbf{A}$. The expectation is denoted by $\mathbb{E}$.

We will use the following symbols: N is the number of carriers; B is the chunk size. Therefore, there are $M = N/B$ chunks. The transmit power constraint of user k is P_k. The channel matrix of user k on carrier n is given by $\mathbf{H}_{k,n}$. The transmit covariance matrix of user k on chunk m is given by $\mathbf{Q}_{k,m}$. The inverse noise power is ρ. The weight of user k to compute the weighted sum rate is given by w_k.

2. SYSTEM MODEL AND PRELIMINARIES

In this section, we introduce the MIMO MAC OFDM model. Since we operate in frequency-selective fading, there are two dimensions for resource allocation available, namely the spatial domain (multiple antennas) and the spectral domain (multiple carriers). To address the two dimensions, we apply linear (over space) precoders for each carrier. At the receiver, on each carrier, MMSE-successive interference cancellation (SIC) is applied.

The feedback limitation introduces blocks of carriers which are precoded with identical linear precoding matrices. We will call those blocks chunks. This additional constraint reduces the feedback overhead and signal processing complexity.

The problem statements are described at the end of this section. It will turn out that the overall multiuser problem can be deconstructed into an iterative solution of single-user problems. Therefore, we present both problem statements.
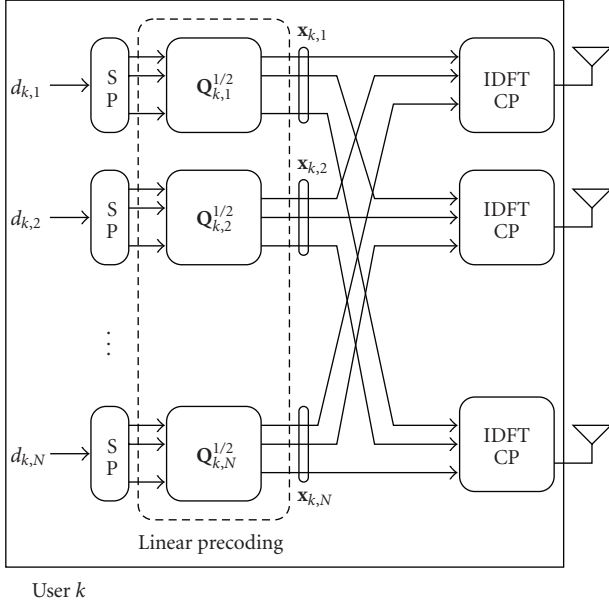
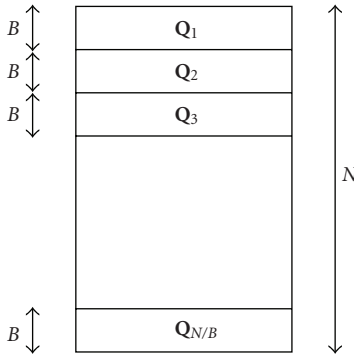FIGURE 2: Transmitter processing for uplink MIMO OFDM system of user k.

FIGURE 3

2.1. Uplink MIMO OFDM system

Consider an ideal multiuser MIMO CP-OFDM system with K users, N carriers, L taps, n_T transmit antennas, n_R receive antennas. Let us focus on the multiple access scenario (uplink). The transmit processing at the kth user is shown in Figure 2.

The N data streams $d_{k,1}, \ldots, d_{k,N}$ of user k are serial parallel converted and linear precoded by $\mathbf{Q}_{k,1}^{1/2}, \ldots, \mathbf{Q}_{k,N}^{1/2}$. Note that the number of parallel data streams depend on the rank of the transmit covariance matrix $\mathbf{Q}_{k,n}^{1/2}$. Next, the N times n_T outputs of the linear precoder are processed in front of each transmit antenna by an IDFT and a cyclic prefix is added (CP-OFDM). Then one OFDM symbol per antenna is transmitted simultaneously.

The received signal on carrier $1 \leq n \leq N$ at the base station is given by

$$\mathbf{y}_n = \sum_{k=1}^{K} \mathbf{H}_{k,n} \mathbf{x}_{k,n} + \mathbf{n}_n, \qquad (1)$$

where $\mathbf{H}_{k,n}$ is the flat-fading channel matrix of user $1 \leq k \leq K$ on carrier n, $\mathbf{x}_{k,n}$ is the transmit vector of user k on carrier n, and $\mathbf{n}_n$ is the white Gaussian noise with variance $\sigma_n^2 = 1/\rho$ on carrier n. The individual transmit power constraint for user k is $\sum_{n=1}^{N} \mathbb{E}[\mathbf{x}_{k,n}\mathbf{x}_{k,n}^H] \leq P_k$. The base station is assumed to apply an MMSE frontend per spatial stream combined with SIC. This receiver architecture is shown to be information lossless in [27, Section 8.3.4].

2.2. Limited feedback

The control unit at the base station takes queueing information as well as physical layer information into account and provides a set of linear precoding strategies for all users. Different scheduling strategies are possible ranging from throughput-oriented scheduling, which is also a subject of the current paper, to stability-based approaches [28]. Since the CSI of all users is necessary for the decision, the centralized approach leads to a base station that informs the users about their transmit strategies by feedback. We assume that the coherence time T in channel uses is large enough to inform the mobiles on transmit strategies for the current channel state.

In order to reduce the amount of feedback required to inform every user on every carrier about the linear precoding matrix, a number of B carriers is assigned the same linear precoding matrix $\mathbf{Q}_b$ (see Figure 3). The total number of carriers N is divided into chunks of size B. Each chunk of size B is processed with the same precoding matrix $\mathbf{Q}_b$. Thus the number of precoding matrices is reduced by a factor of B from N to $M = N/B$. Note that B does not necessarily correspond to the coherence bandwidth of the channel.

Obviously, there is a tradeoff between the amount of feedback and the system performance. The larger B is the less feedback information is required the poorer the system performance will be. The smaller B is the more feedback information is required and the better is the nominal system performance.

2.3. Problem statements

The main question that is answered in this paper is motivated in the previous section: *what is the optimal transmit strategy and what is the optimal chunk size that maximizes the net throughput?* The detailed questions about the impact of the load (number of user K, number of antennas n_T), the impact of the fairness (maximum throughput scheduler, weighted sum rate), and the impact of the channel model, and the user distribution follow immediately.

To answer the main questions and the followup questions, we need to develop an algorithm that finds the optimal linear precoding matrices for a given parameter set. The flat fading case and chunk size of one $N = B = 1$ are solved in

[16]. For $N \geq B > 1$, even the single-user case leads to an optimization problem that cannot be solved in closed form. The following single-user single-chunk problem is the building block that is needed to develop the solution for the multiuser multiple-chunk optimization (In this work, the objective function is always the mutual information with Gaussian code books, except in Section 4.3 in which finite MCS are studied.),

$$\max_{\mathbf{Q}} \sum_{b=1}^{B} c_b \big(\log \det (\mathbf{Z}_b + \mathbf{H}_b \mathbf{Q} \mathbf{H}_b^H) - \log \det (\mathbf{Z}_b) \big) \tag{2}$$

$$\text{s.t. } \mathbf{Q} \succeq \mathbf{0}, \ \text{tr} (\mathbf{Q}) \leq P.$$

The coefficients $c_1, \ldots, c_B$ are nonnegative real numbers and they will be defined below. The operational meaning of the positive definite matrix $\mathbf{Z}_b$ will be the spatial noise plus interference covariance matrix, $\mathbf{H}_b$ will be identified with the MIMO channels within one chunk, $\mathbf{Q}$ is the transmit covariance matrix, and P is the sum transmit power constraint.

Next, the spectral power allocation and the multiuser weighted-sum rate problem is incorporated. Let the weights $w_1, \ldots, w_K$ be ordered in decreasing order, that is, $w_1 \geq w_2 \geq \cdots \geq w_K \geq 0$. We arrive at the following optimization problem: maximize the weighted-sum rate of the K-user MIMO N-carrier OFDM uplink with chunk size B and weights $w_1, \ldots, w_K$,

$$\max_{\mathbf{Q}_{1,1}, \ldots, \mathbf{Q}_{K,M}} \sum_{m=1}^{M} \sum_{b=1}^{B} \sum_{k=1}^{K} \underbrace{(w_k - w_{k+1})}_{c_k}$$

$$\times \log \det \left(\mathbf{I} + \rho \sum_{j=1}^{k} \mathbf{H}_{j,m,b} \mathbf{Q}_{j,m} \mathbf{H}_{j,m,b}^H \right)$$

$$\text{s.t. } \mathbf{Q}_{k,m} \succeq \mathbf{0}, 1 \leq m \leq M, \quad \sum_{m=1}^{M} \text{tr} (\mathbf{Q}_{k,m}) \leq P_k, 1 \leq k \leq K, \tag{3}$$

where $\mathbf{H}_{j,m,b}$ denotes the channel matrix of user j in the bth carrier of chunk m. The optimal SIC orders were used [29, Proposition 2]. The individual power constraint of user k is P_k. ρ is the inverse noise variance defined in Section 2.1. The coefficients c_k in (2) are defined as $c_k = w_k - w_{k+1}$ and thus are nonnegative. We set $w_{K+1} = 0$.

The advantage of the optimization problem in (3) is that the objective function is jointly concave with respect to the tuple $(\mathbf{Q}_{1,1}, \ldots, \mathbf{Q}_{K,M})$, the constraint set is convex, and therefore the programming problem itself is convex. Due to the large number of optimization variables, the direct solution using standard convex optimization tools [30, 31] is not practically feasible. It is also not possible to solve (3) in closed form, however, we will develop an iterative algorithm that solves the problem efficiently even for high numbers of users, carriers, and antennas.

3. OPTIMIZATION THEORETIC RESULTS AND ALGORITHM DEVELOPMENT

In this section, we solve the theoretical problem statements from the last section. We will show that the multiuser problem (3) can be solved by iteratively solving single-user problems. Therefore, we start with the single-user problem first and develop an iterative algorithm. The convergence proof can be found in the appendix.

For the multiuser problem, the SIC decoding order is important. Fortunately, the optimal order depends only on the weights of the users (as in the nonchunk single-carrier case). Based on the single-user algorithm, we develop the multiuser algorithm.

3.1. Optimal single-user chunk processing

The single-user single-chunk case is the basic element of the iterative algorithm that is developed later for the overall multiuser problem solution. Therefore, we study this problem first.

Consider the following simple setup. The B parallel data stream vectors $\mathbf{d}_1, \ldots, \mathbf{d}_B$ of one chunk are linearly precoded by the same linear precoding matrix $\mathbf{Q}^{1/2}$ and then multiplied by different MIMO flat-fading channel matrices $\mathbf{H}_1, \ldots, \mathbf{H}_B$ to obtain

$$\mathbf{y}_b = \mathbf{H}_b \mathbf{Q}^{1/2} \mathbf{d}_b + \mathbf{n}_b, \quad \text{for } 1 \leq b \leq B. \tag{4}$$

The same positive semidefinite transmit covariance matrix $\mathbf{Q}$ has to be used for all channels within one chunk.

Let the input vectors be independently zero-mean complex Gaussian distributed with identity covariance, that is, $\mathbf{d}_k \sim \mathcal{CN}(0, \mathbf{I})$ and the noise vectors are independently zero-mean complex Gaussian distributed with covariance $\mathbf{Z}_b$, that is, $\mathbf{n}_b \sim \mathcal{CN}(0, \mathbf{Z}_b)$. The weighted mutual information between input and output of the system is given by

$$\Psi(\mathbf{Q}) = \sum_{b=1}^{B} c_b I(\mathbf{d}_b; \mathbf{y}_b)$$

$$= \sum_{b=1}^{B} c_b \log \det (\mathbf{Z}_b + \mathbf{H}_b \mathbf{Q} \mathbf{H}_b^H) - c_b \log \det (\mathbf{Z}_b). \tag{5}$$

If $B = 1$, the optimal choice of $\mathbf{Q} \succeq \mathbf{0}$ under trace constraint diagonalizes the channel matrix and the optimal power allocation is given by water filling [32]. This strategy is not applicable for $B > 1$ because $\mathbf{Q}$ cannot diagonalize jointly all channel matrices $\mathbf{H}_1, \ldots, \mathbf{H}_B$ except for the unlikely case that they all commute. The case $c_1 = c_2 = \cdots = c_B = 1$ and $\mathbf{Z}_1 = \mathbf{Z}_2 = \cdots = \mathbf{Z}_B = \sigma_n^2 \mathbf{I}$ is solved in [33]. Note that in [34] a similar but different iterative approach based on the Cholesky decomposition of $\mathbf{Q}$ was developed. Our approach has the important advantage that the optimization problem stays convex and global convergence can be proved.

Result: Solve optimization problem (7)
Input: Channel realization $\mathbf{H}_{1,1}, \ldots, \mathbf{H}_{M,B}$ and power constraint $P > 0$
initialization: for all $1 \le m \le M : \mathbf{Q}_m^0 = \mathbf{0}, \mathbf{Q}_m^1 = (P/M \cdot n_T)\mathbf{I}$, and set $\ell = 1$;
While $\|(\sum_{m=1}^{M} \Psi(\mathbf{Q}_m^\ell)) - (\sum_{m=1}^{M} \Psi(\mathbf{Q}_m^{\ell-1}))\| > \epsilon$ **do**
    $\ell = \ell + 1$;
    $\mathbf{Q}_m^\ell = \mathbf{Q}_m^{\ell-1(1/2)}\Psi'(\mathbf{Q}_m^{\ell-1})\mathbf{Q}_m^{\ell-1(1/2)}$ for all $1 \le m \le M$;
    $\mu = \sum_{m=1}^{M} \mathrm{tr}\,(\mathbf{Q}_m^\ell)$;
    $\mathbf{Q}_m^\ell = (P/Mn_T\mu)\mathbf{Q}_m^\ell$ for all $1 \le m \le M$;
end
output: Optimal set of transmit covariance matrices $\mathbf{Q}_1, \ldots, \mathbf{Q}_M$

ALGORITHM 1: Single-user optimal MIMO OFDM chunk processing.

Theorem 1. *Let the start point be* $\mathbf{Q}^0 = (P/n_T)\mathbf{I}$. *The update rule*

$$\mathbf{Q}^{\ell+1} = \frac{P\sum_{b=1}^{B} c_b (\mathbf{I} - \mathbf{Z}_b^{-1/2}[\mathbf{Z}_b + \mathbf{H}_b\mathbf{Q}^\ell\mathbf{H}_b^H]^{-1}\mathbf{Z}_b^{-1/2})}{\mathrm{tr}\,(\sum_{b=1}^{B} c_b (\mathbf{I} - \mathbf{Z}_b^{-1/2}[\mathbf{Z}_b + \mathbf{H}_b\mathbf{Q}^\ell\mathbf{H}_b^H]^{-1}\mathbf{Z}_b^{-1/2}))} \quad (6)$$

converges to the optimal $\mathbf{Q}^*$ *and solves optimization problem* (2).

The proof can be found in Appendix A. Note that the fixed point iteration in (6) has only linear convergence [35] and any Newton style algorithm has local super-linear convergence. However, the update rule in (6) is further refined to include spectral power allocation. If a Newton style algorithm is used, this extension is not directly possible.

Before the complete multiuser algorithm is developed, we consider the case in which multiple $M = N/B$ chunks are jointly optimized under a sum power constraint $\sum_{m=1}^{M} \mathrm{tr}\,(\mathbf{Q}_m) \le P$. This corresponds to the single-user MIMO OFDM case. The optimization problem reads

$$\max_{\mathbf{Q}_1, \ldots, \mathbf{Q}_M} \sum_{m=1}^{M} \sum_{b=1}^{B} c_b \big(\log \det (\mathbf{Z}_{m,b} + \mathbf{H}_{m,b}\mathbf{Q}_m\mathbf{H}_{m,b}^H)$$
$$- \log \det (\mathbf{Z}_{m,b}) \big) \quad (7)$$
$$\mathbf{Q}_m \succeq \mathbf{0}, 1 \le m \le M, \quad \sum_{m=1}^{M} \mathrm{tr}\,(\mathbf{Q}_m) \le P$$

and the spectral power allocation corresponds to water filling. The naive approach is to alternate between covariance matrix optimization and spectral power allocation because the problem is jointly concave in the chunk powers and the chunk covariance matrices. However, this approach converges usually very slow.

For the case in which $B > 1$ we develop an efficient algorithm that merges the spectral power allocation in the update rule from Theorem 1. The algorithm was presented for $c_1 = c_2 = \cdots = c_B = 1$ and $\mathbf{Z}_1 = \mathbf{Z}_2 = \cdots = \mathbf{Z}_B = \sigma_n^2\mathbf{I}$ in [33]. In the following, lemma an iterative algorithm is proposed which solves (7).

Lemma 1. *Algorithm 1 solves the optimization problem* (7).

The proof can be found in Appendix B. The function Ψ is defined in (5). The convergence rate of Algorithm 1 is illustrated in Figure 4 where $\epsilon = 10^{-3}$ is used. In Figure 4, it can be observed that for larger chunk sizes the convergence rate is faster because the objective function is lower and there are less optimization variables. This fast convergence is a mandatory prerequisite to embed Algorithm 1 in the iterative water-filling algorithm for weighted sum rate optimization in the next section.

3.2. Multiuser chunk processing: weighted sum capacity

In the multiuser setting, we study the uplink scenario with SIC at the base and solve the optimization problem

$$\max_{\mathbf{Q}_{1,1}, \ldots, \mathbf{Q}_{K,M}} \sum_{k=1}^{K} w_k \sum_{m=1}^{M} R_{k,m}$$

$$\text{s.t. } \mathbf{Q}_{k,m} \succeq \mathbf{0}, 1 \le m \le M, \quad \sum_{m=1}^{M} \mathrm{tr}\,(\mathbf{Q}_{k,m}) \le P_k, 1 \le k \le K,$$
$$(8)$$

where $R_{k,m}$ is the mutual information by user k in chunk m. The individual achievable rates depend on the SIC order. Reference [29, Proposition 2] shows that the optimal decoding order π satisfies $w_{\pi_1} \ge w_{\pi_2} \ge \cdots \ge w_{\pi_K} \ge 0$. By inserting the optimal decoding order into (8) and collecting two succeeding terms in the sum, we obtain the programming problem in (3).

The optimization problem (3) is a convex-optimization problem because the objective function is the positive-weighted sum of functions which are jointly concave in the set of transmit covariance matrices $\{\mathbf{Q}_{1,1}, \ldots, \mathbf{Q}_{K,M}\}$ and the constraint set is convex. Furthermore, the number of optimization variables is too large, for example, for $N = 2048, B = 2, n_T = 4, K = 20$ there are 20480 covariance matrices of size 4×4 involved, to directly apply a convex optimization method, for example, an interior point method. Instead, the structure of the optimization problem is taken into account and the problem is decomposed into single-user problems with colored noise. The fundamental difference to
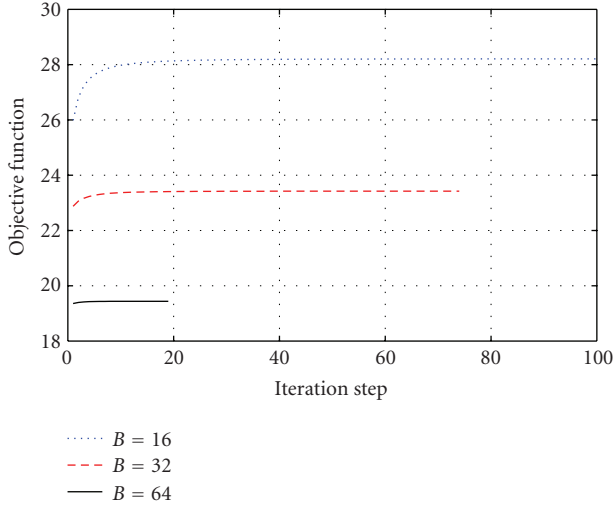
FIGURE 4: Convergence rate of single-user MIMO OFDM chunk optimization with $n_T = 2, n_R = 2, N = 1024$ and different chunk sizes for ideal iid Rayleigh fading channel model.

standard iterative water filing [16] is the single-user step and the additional spectral power allocation.

Algorithm 2 first initializes the covariance matrices to identity matrices. Next, for all users the single-user multicarrier chunk optimization from Algorithm 1 is performed. Since the objective function is increasing in each step and there is a unique global optimum, the algorithm converges to the optimum. The formal proof is similar to the proof in [29, Proposition 7] and is therefore omitted. The general conditions for convergence and the convergence speed of the alternating optimization approach are given in [36, Theorems 2 and 3].

4. SYSTEM DESIGN AND OPTIMAL CHUNK SIZE

In this section, we use the developed algorithm to show how practical limitations, namely, quantization and finite modulation and coding schemes (MCS) can be incorporated. Furthermore, the performance measure is introduced which takes the feedback overhead into account. Later simulations will all be based on this net throughput.

The control unit decides on transmit strategies, that is, linear precoding matrices $\mathbf{Q}_{1,1}, \ldots, \mathbf{Q}_{K,N}$, modulation, and coding for each user at each carrier. Feedback from base to mobile is required. A full rank $\mathbf{Q}_k$ has n_T^2 complex entries, however it can be reduced to $n_T + (n_T - 1) \cdot n_T = n_T^2$ real entries since the matrix is Hermitian. Thus, the worst case feedback ($\eta = n_T$) from base to mobiles are $K \cdot N \cdot n_T^2$ real values. By applying different chunk sizes, the feedback overhead and signal processing complexity can be decreased, reducing thereby the performance of the system.

In this section, a measure for the effective overall transmission rate is derived. Furthermore, several practical aspects as quantization of the linear precoding matrices and MCS are discussed.

4.1. Net throughput

Following the feedback computation above, the amount of feedback as a function of the number of transmit antennas n_T, the number of users K, the quantization q, the coherence time T, the number of carriers N, the chunk size B, and feedback channel data rate R_d is defined by

$$\alpha = \frac{N \cdot K \cdot \zeta}{B \cdot R_d \cdot T} \qquad (9)$$

with ζ as the number of feedback bits for one transmit covariance matrix. As an example, assume a scalar quantization and an 8-bit quantization per real value. This leads to $\zeta = n_T^2 q$ and $K \cdot N \cdot n_T^2 \cdot 8$ bits feedback. Consider for example $K = 10, N = 1024, n_T = 2$. Then 320 Kbits per coherence time (or per frame) are necessary. Further on, the signal processing at transmitter needs N multiplications of transmit data block with $n_T \times n_T$ matrices. Assume a feedback rate $R_d = 320$ bits per channel use and $T = 500$ channel uses. The resulting feedback amount in (9) is given by $\alpha = 0.002N/B$.

In our approach, the control overhead reduces the transmission rate R to the effective transmission rate R_e,

$$R_{e,1} = R(1 - \alpha) = R\left(1 - \frac{N \cdot K \cdot \zeta}{B \cdot R_d \cdot T}\right). \qquad (10)$$

This approach considers only the uplink and the feedback reduces the transmission rate directly.

As in other communications systems, there are complex tradeoffs between design parameters and performance in multiuser MIMO OFDM MAC. In (10), there are two tradeoffs. The first is with respect to the chunk size B. The larger B, the worse is the performance but the smaller is also the feedback overhead. The second tradeoff is with respect to the quantization level q. The larger q, the better is the performance because the linear precoding matrices are represented better, but the higher is also the feedback overhead.

4.2. Quantized linear precoding

In [37], methods and performance results of quantized feedback approaches for multiple antenna channels are described and compared. A concrete vector quantization scheme based on Grassmannian subspace packing is proposed in [38] for single-user beamforming without power allocation. In the multiuser setting, it often happens (see multiuser illustrations in Figure 9) that only a small number of streams with different powers are allocated. Therefore, the Grassmannian subspace packing can be extended with a rough quantization of the power allocation to arrive at a full transmit covariance matrix. The channel optimized covariance matrix quantization is beyond the scope of this paper.

In the effective rate definition (10), q is the quantization level of every real number that is needed to parameterize the channel covariance matrix. In the worst case, $n_T^2 q$ bits are needed, that is, one transmit covariance matrix $\mathbf{Q}$ is described by $n_T^2 q$ bits. Since this number is large even for small number of antennas and quantization levels we restrict our attention to the random vector quantization approach

---

Result: Solve optimization problem (3)
Input: Channel realizations $\mathbf{H}_{1,1,1}, \ldots, \mathbf{H}_{K,M,B}$ and power constraints $P_k > 0$ for
$\qquad 1 \le k \le K$
initialization: for all $1 \le m \le M$ and $1 \le k \le K : \mathbf{Q}_{k,m}^0 = \mathbf{0}, \mathbf{Q}_{k,m}^1 = (P_k/M \cdot n_T)\mathbf{I}$, and set $\ell = 1$;
While $\| \sum_{m=1}^M \sum_{k=1}^K \sum_{b=1}^B (w_k - w_{k+1}) \log (\det (\mathbf{I} + \rho \sum_{j=1}^k \mathbf{H}_{j,m,b} \mathbf{Q}_{j,m}^\ell \mathbf{H}_{j,m,b}^H)/ \det (\mathbf{I} + \rho \sum_{j=1}^k \mathbf{H}_{j,m,b} \mathbf{Q}_{j,m}^{\ell-1} \mathbf{H}_{j,m,b}^H)) \| > \epsilon$ **do**
$\quad \ell = \ell + 1$;
$\quad$ For $1 \le k \le K, 1 \le m \le M$ set $\mathbf{Q}_{k,m}^\ell = \mathbf{Q}_{k,m}^{\ell-1}$;
$\quad$ **for** $k = 1, \ldots, K$ **do**
$\qquad \{\mathbf{Q}_{k,1}^\ell, \ldots, \mathbf{Q}_{k,B}^\ell\} = \arg \max_{\mathbf{Q}_1, \ldots, \mathbf{Q}_M} \sum_{j=k}^K (w_j - w_{j+1})$
$\qquad \sum_{m=1}^M \sum_{b=1}^B \log \det (\mathbf{I} + \rho \sum_{l=1, j \ne k}^j \mathbf{H}_{l,m,b} \mathbf{Q}_{j,m}^\ell \mathbf{H}_{l,m,b}^H + \rho \mathbf{H}_{k,m,b} \mathbf{Q}_m \mathbf{H}_{k,m,b}^H)$
$\qquad$ s.t. $\mathbf{Q}_m \succeq \mathbf{0}, 1 \le m \le M$ and $\sum_{m=1}^M \operatorname{tr}(\mathbf{Q}_m) \le P_k$ by Algorithm 1;
$\quad$ **end**
end
Output: Optimal set of transmit covariance matrices $\mathbf{Q}_{1,1}, \ldots, \mathbf{Q}_{K,M}$

ALGORITHM 2: Multiple user optimal MIMO OFDM chunk processing.

[7, 39]. We use $(n_T - 1)q$ bits for the power quantization and the remaining bits for beamformer quantization.

Consider, for example, the case in which the mobiles have two transmit antennas and $q = 8$, we generate 16 777 216 random vectors for beamforming quantization. The two eigenvalues of the covariance matrix corresponding to the power allocation are uniformly quantized according to 16 levels between 0 and the maximum transmit power.

4.3. Modulation and coding schemes

In the ideal simulations, the mobiles use independent Gaussian code books. However, in practice finite modulation and coding schemes are employed. These limitations influence the resource allocation and limit the performance of a single stream. In order to show the impact of finite modulation and coding schemes (MCS), we present also results with respect to the MCS shown in Figure 5. At high SNR, the maximum achievable rate is bounded by 4.5 bit/s (64-QAM with code rate 3/4). The MCS used in Figure 5 are defined in [40].

The conversion from the rates achievable with Gaussian code books to finite MCS works via the SINR values of the individual data streams. The receiver applies the optimum combining (OC) method [41]. Hence, the SINR for data stream s of user k in chunk b on carrier θ is given by(We omit the indices b and θ for convenience.)

$$\mathrm{SINR}_{k,s} = \widetilde{\mathbf{h}}_{k,s}^H \left[\sum_{t \ne s} \widetilde{\mathbf{h}}_{k,t} \widetilde{\mathbf{h}}_{k,t}^H + \mathbf{Z}_k \right]^{-1} \widetilde{\mathbf{h}}_{k,s} \qquad (11)$$

with effective channel after precoding $\widetilde{\mathbf{h}}_{k,s} = \mathbf{H}_k \mathbf{Q}_{k,s}^{1/2}$, where $\mathbf{Q}_{k,s}^{1/2} = \mathbf{v}_{k,s} p_{k,s}^{1/2}$ is the beamforming vector $\mathbf{v}_{k,s}$ and power allocation $p_{k,s}$ of user k and stream s and with noise plus multiple access interference after SIC (For sum rate optimization the SIC order is arbitrary. For weighted-sum rate op-
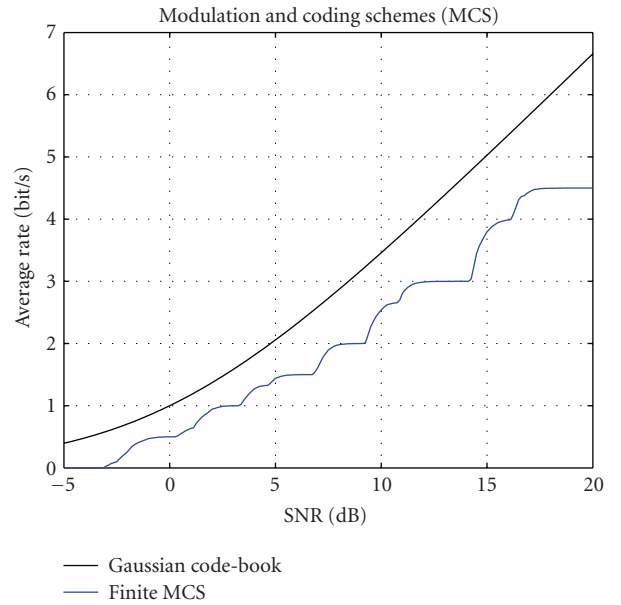

FIGURE 5: Average rate versus SNR for Gaussian code-book and for finite modulation and coding schemes.

timization, we assume that the users are ordered according to $w_1 \ge w_2 \ge \cdots \ge w_K \ge 0$.),

$$\mathbf{Z}_k = \sum_{l=k+1}^K \mathbf{H}_l \mathbf{Q}_l \mathbf{H}_l^H + \sigma_n^2 \mathbf{I}. \qquad (12)$$

Note that the linear precoding matrices as well as the optimal decoding order hold only for Gaussian code books. However, the optimization of the weighted sum rate under finite MCS constraints is a combinatorial nonlinear problem with high computational complexity. Therefore, we optimize first under the Gaussian signalling assumption and map then the SINR values to finite MCS achievable rates. This approach is suboptimal.

As it can be seen in Figure 5, the difference between the rates achievable with finite MCS and the Gaussian codebook is characterized by the following behavior. First the MCS curve is shifted to the right and second, that at high SNR the rate achievable with finite MCS is bounded by 4.5 bit/s/Hz. The first difference can be resolved by the SINR-gap concept [42, 43]. For high SNR, the second difference leads to a problem because increasing the SINR from a certain point does not increase the achievable rate of finite MCS. On the one hand, this problem occurs seldom because the SINR is limited by multiple access interference. On the other hand, it occurs in sparse resource allocation scenarios where only a single user is scheduled on one chunk, this may lead to a performance loss. One remedy is to increase the finite MCS for higher SINR. Another remedy could be to include this restriction into the original optimization problem without destroying the convenient structure. This is left as an open research problem.

5. ILLUSTRATIONS

In this section, we illustrate the theoretical results as well as the practical implications. First, the rate region is completely computed for an ideal channel model without quantization and MCS constraints but with chunk constraint. These results show the performance gain of the proposed algorithm compared to existing algorithms. Next, the IEEE 802.11n channel model is used to illustrate a particular chunk size optimization (again without quantization and MCS). Finally, the WIM2 channel model is used to illustrate all the practical limitations.

5.1. Rate region for ideal Rayleigh channels

In Figure 6 the achievable rate region of a realization of an identically and independently distributed (iid) Rayleigh fading channel with $L = 6$ taps, equal power delay profile, $N = 32$ carriers, and two users is shown for different chunk sizes. The region is computed with Algorithm 2 for 33 different weights $\mathbf{w} = [\omega, 2 - \omega]$ with ω ranging from 0.01 to 1.99 in steps of 0.06. We assume nonquantized precoding matrices. The feedback overhead is not considered in the rates R_1 and R_2 shown in Figure 6.

In Figure 6 it can be observed that even for a chunk size of $B = 2$ the region shrinks compared to perfect feedback with $B = 1$ although the coherence bandwidth is larger than two carriers. The performance degradation between $B = 1$ and $B = 32$ at the sum rate point is about 50%.

We compare the achievable rate region with the suboptimal scheme which takes the average channel matrix within each chunk for optimization. This scheme is optimal for small SNR [44] only. The advantage of the proposed Algorithms 1 and 2 can be clearly observed especially for larger chunk sizes.

5.2. Sum rate in IEEE 802.11n uplink channels

In Table 1 the chunk size and the corresponding feedback overhead in percent, the number of OFDM symbols used for
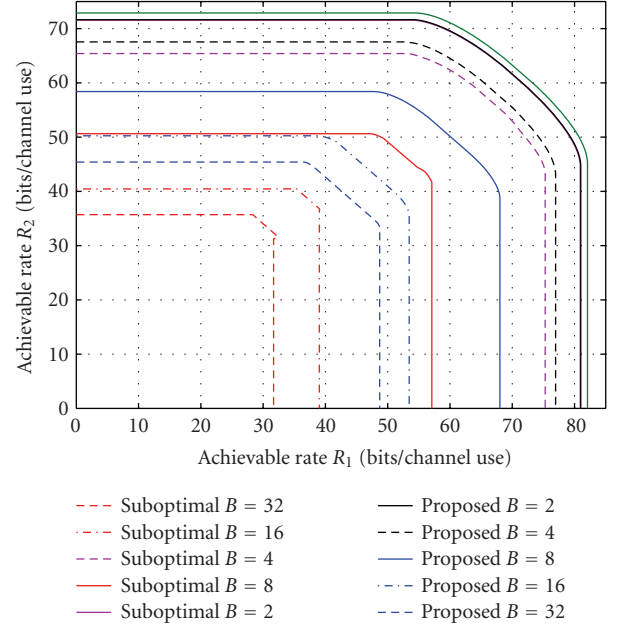

FIGURE 6: Two user rate region for different chunk sizes in ideal frequency-selective iid Rayleigh fading channel.

TABLE 1: Feedback overhead, number of OFDM symbols for feedback, and sum rate for different chunk sizes for 20 users in IEEE 802.11n channel model.

Chunk size B	Feedback overhead in %	# of OFDM symbols	Sum rate (Mbit/symb)
32	45.31	10	929
64	22.66	5	905
128	11.33	2-3	865
256	5.66	1-2	828
512	2.83	1	791
1024	1.42	<1	768

feedback, and the sum rate R are shown for a multiuser scenario with $K = 20$ users, $n_T = n_R = 2$ antennas at 15 dB SNR based on the IEEE 802.11n channel model. The precoding matrices are fedback without quantization.

From Table 1, we observe that the feedback overhead can be reduced significantly with only a small penalty in the achievable sum rate. Note that if only a maximum of 4 OFDM symbols is allowed for feedback signaling (which is equivalent to 18% overhead), the chunk size has to be larger than 128.

The results in Table 1 show that the sum rate decreases only slowly by increasing the chunk size. This behavior depends on the SNR, the channel model, and the number of users. For asymptotically high SNR, equal power allocation is optimal and therefore, the transmit strategies do not depend on the carrier. The performance loss increases with the frequency selectivity of the channel. In IEE802.11n model D and E, 18 taps are created by 3 and 4 clusters, respectively. The more users are available (the channels of the users are
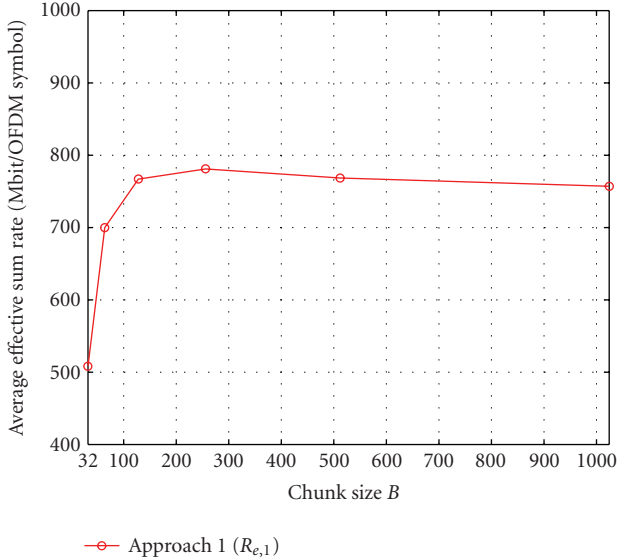
FIGURE 7: Effective average transmission rate R_e over chunk size B for 20 users in IEEE 802.11n channel model.

FIGURE 8: Effective average transmission rate R_e over chunk size B for 5 users in WIM2 local area channel model A1.

generated independently by the IEEE802.11n model D and E) the easier the algorithm can allocate chunks to users who do not fluctuate too much.

In Figure 7, the average sum rates for different chunk sizes are depicted with $n_T = 2$ and $n_R = 2$, $K = 20$ users and an SNR of 15 dB. From the figure, it can be observed that the maximum average efficient sum rate is achieved for $R_{e,1}$ for $B = 256$.

5.3. Sum rate in WINNER local area scenario

In Figure 8, the average effective sum rate of a five-user local area scenario are shown. The system parameters are according to the definition in [40] for the local area (LA) scenario, that is, eight cross-polarized base station antennas and two dual cross polarized antennas. 1840 out of 2048 carriers and a signal bandwidth of 81.25 MHz out of a system bandwidth of 100 MHz are used. The feedback load was set to 10^6. No quantization of the linear precoding matrices is assumed. The chunk sizes are varied between $16 \leq B \leq 1840$. Three different SNR, defined as individual power constraint divided by noise power, are studied from -5 dB to 15 dB.

There are several observations in Figure 8. At first, the degradation due to finite MCS fluctuates between 20% for high SNR, 40% for medium SNR, and 30% for small SNR. The main source of rate loss is the upper bound on the rate of the finite MCS (at 4.5 bit in Figure 5). At medium and low SNR, the absolute loss due to finite MCS is decreased, for medium SNR, the average sum rate even increases with increasing chunk size from $B = 920$ to $B = 1840$. The reason for this lies in the fact that with individual power constraints and only one large chunk, all users are scheduled simultaneously (In the uplink scenario with individual power constraints it can be easily shown that all users should transmit with maximum individual power to be Pareto optimal.) on
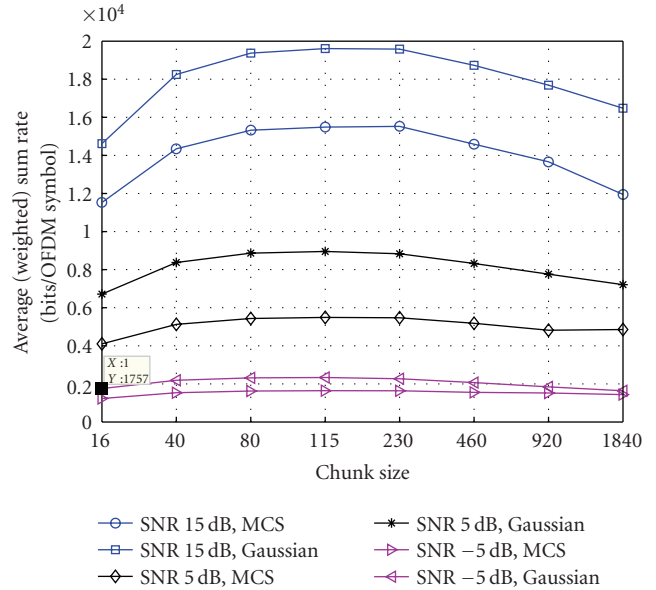
that chunk and the individual SINRs of the data streams are clearly interference limited. No data stream is saturated with respect to the maximum data rate of the MCS.

The power allocation in the left upper subfigure in Figure 9 shows that indeed too much power is allocated to two users on a single chunk and hence the SINR for those users is too high. However, the remaining three users distributed their power over the chunks. Therefore, there is the sum rate loss of about 35% for a chunk size of 115. For larger $B = 230$, there is much more multiple access interference (see Figure 9 right-hand side) and thus the loss due to finite MCS is smaller, about 20%.

Second, for all SNR values, there is an optimal chunk size at $B = 115$ which is larger than the coherence bandwidth of the channel (between 8 and 16 carriers). Another important observation is that the loss between the optimal chunk size and the minimum chunk size $B = 16$ is for all SNR around 25–27%.

5.4. Resource allocation in WINNER LA

Note that the solution of the optimization problem (8) contains implicitly the mapping of users to chunks because multiple transmit covariance matrices $\mathbf{Q}_{k,m}$ will be zero and thus user k will not be scheduled on chunk m.

Figure 9 shows a typical power allocation of the users over the chunks for one fixed channel realization of the WIM2 A1 channel model at SNR 5 dB. The channel model is for indoor small office or residential scenario with line-of-sight (LOS) with velocities between 0 and 5 km/h. Note that the sum powers of all users are identical. Two different chunk sizes are compared.

In Figure 9, it can be observed that there are two types of power allocations, namely, a peaky power allocation of user 3 and 4 and a flat power allocation for users 1, 2, and 5. These
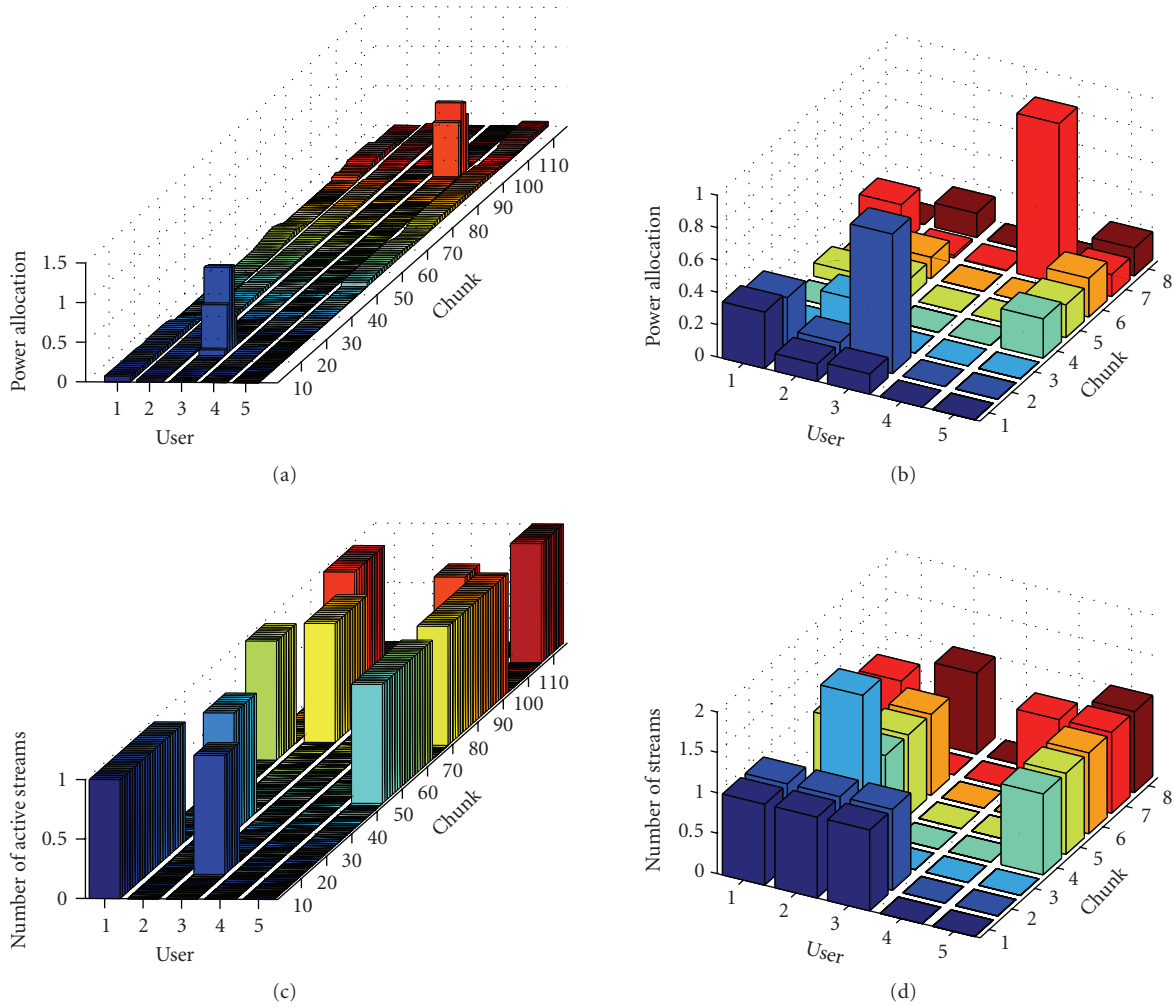
(a)

(b)

(c)

(d)

Figure 9: Power allocation and number of active streams of users over chunks for different chunk sizes: $B = 16$ and $B = 230$.
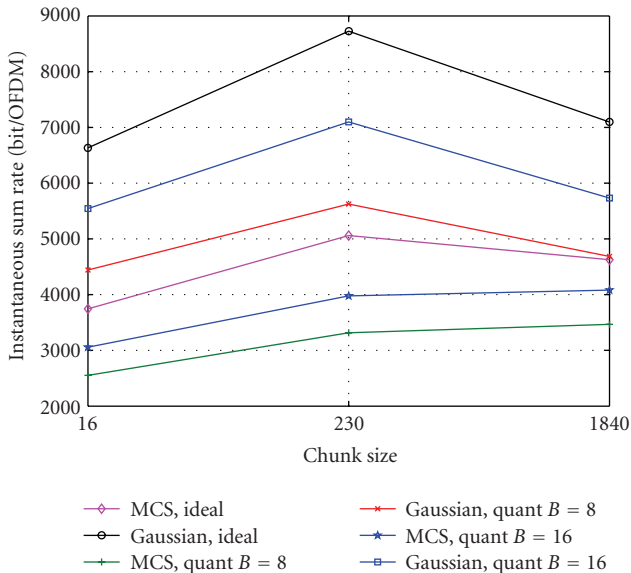

Figure 10: Impact of transmit covariance matrix quantization on the instantaneous sum rate for 5 users in WIM2 A1 channel.

peaky power allocations lead to the rate loss for finite MCS described above. If the chunk size is increased, more and more users are scheduled on the same chunk. For $B = 230$, three users are loaded on one chunk on average.

For a chunk size of $B = 1840$ all users transmit simultaneously on the same chunk. One interesting question is whether the users perform single-stream beamforming or spatial multiplexing. For the channel realization from Figure 9, only one user performs spatial multiplexing whereas all other users perform single-stream beamforming. This observation corresponds to the results in [45, 46].

5.5. Impact of quantization in WINNER LA

In Figure 10, the impact of the quantization of the transmit covariance matrix is illustrated for one instantaneous channel realization. For every transmit covariance matrix 16 bits or 8 bits are allocated. The same setting as in Figure 8 is used. In Figure 10, it can be observed that the degradation due to finite quantization of the precoding matrices is about 20% for $q = 16$ and 35% for $q = 8$.

Another observation from Figure 10 is that for finite MCS with quantization, the achievable rate increases monotonically with the chunk size. The reason for this behavior is that the gain by adapting better to the channels for small chunk sizes is lost due to the rough quantization of the transmit covariance matrices. In this scenario, it seems to be better to use only one roughly quantized transmit covariance matrix over the maximum chunk size.

6. DISCUSSIONS AND CONCLUSIONS

The work solves the weighted-sum rate optimization of the MIMO OFDM MAC under individual power constraints and under chunk constraints. In order to reduce the feedback overhead by informing the mobiles of their linear precoding matrices, the optimal chunk size is numerically found. Finite MCS and quantization of the linear precoding matrices are also addressed. The performance is illustrated by numerical simulations according to different channel models.

6.1. Further applications

In the iterative water-filling algorithm that was proposed for the weighted-sum rate optimization of flat fading MIMO channels, the update rule (6) can be directly applied to solve the inner single-user update step, that is, to solve [29, Section VII.B], first step in 6.a.] Note that this step can be also solved by a modified MAXDET [30] software. However, depending on the chunk size and the number of transmit and receive antennas, the algorithm converges slowly or does not stop at all.

Another application of the update rule and also of Algorithm 1 is in the context of code-book design for MIMO systems with limited feedback. In [47], an iterative algorithm was proposed to find the set of covariance matrices for quantized feedback with statistical CSI at the transmitter. The algorithm is similar to Lloyd algorithm which finds a local optimum in channel optimized vector quantization [48, 49]. For fixed index mapping and set of channels $\mathcal{H}$, the optimal covariance matrices are found by solving

$$\arg\max_{\mathbf{Q} \geq \mathbf{0}} \frac{1}{|\mathcal{H}|} \sum_{\mathbf{H} \in \mathcal{H}} \log \det\left(\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^H\right) \tag{13}$$
$$\text{s.t. } \operatorname{tr}\mathbf{Q} \leq P.$$

The programming problem (13) can be directly solved with the update rule (6).

6.2. Open problems

In the work, the optimization under chunk constraints is performed first and then the additional finite MCS and quantization constraints are considered. The performance can be improved if the limitations by finite MCS and quantized precoding matrices are modeled as additional constraints in the programming problem (3). Thereby the constraints are to be formulated such that the properties of the constraints are conserved, especially convexity and continuity. The channel optimization quantization of the transmit covariance matrices in this multiuser setting is also unsolved.

In the work, it is assumed that all active users in the system have data to transmit, that is, all users have full buffers. However, the weighted sum rate optimization can be used to incorporate the queuing behavior of the system, that is, assume arrival processes at the mobiles. The interesting performance measure is then the system stability.

APPENDICES

A. PROOF OF THEOREM 1

The optimization problem is a convex optimization problem with a feasible start point. Therefore, the duality gap is zero and we derive the KKT optimality conditions for (2) in standard form (minimization of convex objective) [50]. The Lagrangian is

$$L(\mathbf{Q}, \mu, \boldsymbol{\Xi}) = -\Psi(\mathbf{Q}) + \mu\left(\operatorname{tr}(\mathbf{Q}) - P\right) - \operatorname{tr}(\boldsymbol{\Xi}\mathbf{Q}). \tag{A.1}$$

One of the KKT conditions is obtained by setting the first derivative of $L(\mathbf{Q})$ with respect to $\mathbf{Q}$ to zero, that is,

$$\frac{\partial L(\mathbf{Q}, \mu, \boldsymbol{\Xi})}{\partial \mathbf{Q}} = -\sum_{b=1}^{B} c_b \mathbf{H}_b^H \left[\mathbf{Z}_b + \mathbf{H}_b \mathbf{Q} \mathbf{H}_b^H\right]^{-1} \mathbf{H}_b + \mu\mathbf{I} - \boldsymbol{\Xi} = \mathbf{0}. \tag{A.2}$$

Since from the other optimality conditions follows $\mathbf{Q}\boldsymbol{\Xi} = \mathbf{0}$, we solve (A.2) for the first summand and multiply both sides with $\mathbf{Q}^{1/2}$ from the left and from the right

$$\mathbf{Q} = \frac{1}{\mu}\sum_{b=1}^{B} c_b \mathbf{Q}^{1/2} \mathbf{H}_b^H \left[\mathbf{Z}_b + \mathbf{H}_b \mathbf{Q} \mathbf{H}_b^H\right]^{-1} \mathbf{H}_b \mathbf{Q}^{1/2} \tag{A.3}$$

with μ chosen such that $\operatorname{tr}(\mathbf{Q}) = P$. Note that the RHS of (A.3) is positive semidefinite. The characterization in (A.3) leads directly to an iterative algorithm with the updating rule

$$\mathbf{Q}^{\ell+1} = \frac{1}{\mu}\mathbf{Q}^{\ell,1/2}\Psi(\mathbf{Q}^\ell)'\mathbf{Q}^{\ell,1/2}$$
$$= \frac{1}{\mu}\sum_{b=1}^{B} c_b \mathbf{Q}^{\ell,1/2} \mathbf{H}_b^H \left[\mathbf{Z}_b + \mathbf{H}_b \mathbf{Q}^\ell \mathbf{H}_b^H\right]^{-1} \mathbf{H}_b \mathbf{Q}^{\ell,1/2} \tag{A.4}$$

and the water level

$$\mu = \frac{P}{\operatorname{tr}\left(\sum_{b=1}^{B} c_b \mathbf{Q}^{\ell,1/2} \mathbf{H}_b^H \left[\mathbf{Z}_b + \mathbf{H}_b \mathbf{Q}^\ell \mathbf{H}_b^H\right]^{-1} \mathbf{H}_b \mathbf{Q}^{\ell,1/2}\right)}. \tag{A.5}$$

Due to the convexity of the programming problem, in the optimum the update rule will give the same covariance matrix and this is used as a stopping criteria. The representation in (6) follows with positive semidefinite $\mathbf{X}$ and positive definite $\mathbf{Z}$ from

$$\mathbf{X}^{1/2}[\mathbf{Z} + \mathbf{X}]^{-1}\mathbf{X}^{1/2} = \mathbf{I} - \mathbf{Z}^{-1/2}[\mathbf{Z} + \mathbf{X}]^{-1}\mathbf{Z}^{-1/2}. \tag{A.6}$$

The second part proves that the update rule converges to the global optimum by showing that the objective function is increased in every step of (6). Note that this part

of the proof works for any objective function $\Psi(\mathbf{A})$. Consider the ℓth step from $\mathbf{A} = \mathbf{Q}^\ell$ to the new point $\mathbf{Q}^{\ell+1} = \mathbf{A}^{1/2}\Psi(\mathbf{A})'\mathbf{A}^{1/2}/\operatorname{tr}(\mathbf{A}\Psi(\mathbf{A})')$. The directional derivative of the objective function at the point $\mathbf{A}$ in the update direction

$$\Delta = \frac{\mathbf{A}^{1/2}\Psi(\mathbf{A})'\mathbf{A}^{1/2}}{\operatorname{tr}(\mathbf{A}\Psi(\mathbf{A})')} - \mathbf{A} \qquad (A.7)$$

is given by

$$\frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}\Psi(\mathbf{A}+\epsilon\Delta) = \operatorname{tr}(\Delta\Psi'(\mathbf{A}))$$
$$= \operatorname{tr}\left(\frac{\mathbf{A}^{1/2}\Psi'(\mathbf{A})\mathbf{A}^{1/2}\Psi'(\mathbf{A})}{\operatorname{tr}(\mathbf{A}\Psi'(\mathbf{A}))} - \mathbf{A}\Psi'(\mathbf{A})\right). \qquad (A.8)$$

Let us define $\mathbf{B} = \Psi'(\mathbf{A})$, then it holds

$$\operatorname{tr}\left(\frac{\mathbf{A}^{1/2}\mathbf{B}\mathbf{A}^{1/2}\mathbf{B}}{\operatorname{tr}(\mathbf{A}\mathbf{B})} - \mathbf{A}\mathbf{B}\right)$$
$$= \frac{1}{\operatorname{tr}(\mathbf{A}\mathbf{B})}\left(\operatorname{tr}(\mathbf{A}^{1/2}\mathbf{B}\mathbf{A}^{1/2}\mathbf{B}) - \operatorname{tr}(\mathbf{A}\mathbf{B})^2\right). \qquad (A.9)$$

By assumption, the matrix $\mathbf{A} = \mathbf{Q}^\ell$ is trace constrained and we assume without loss of generality that $\operatorname{tr}(\mathbf{A}) = 1$. The matrix $\mathbf{B}$ in (A.9) is positive semidefinite. Next, we find the matrix $\mathbf{B}^*$ which minimizes the expression in brackets in (A.9), that is, we define

$$g(\mathbf{B}) = \operatorname{tr}(\mathbf{A}^{1/2}\mathbf{B}\mathbf{A}^{1/2}\mathbf{B}) - \operatorname{tr}(\mathbf{A}\mathbf{B})^2 \qquad (A.10)$$

and solve the minimization problem

$$\mathbf{B}^* = \arg\min_{\mathbf{B}\succeq\mathbf{0}} g(\mathbf{B}). \qquad (A.11)$$

The Lagrangian for (A.11) is $L(\mathbf{B},\mathbf{\Gamma}) = \operatorname{tr}(\mathbf{A}^{1/2}\mathbf{B}\mathbf{A}^{1/2}\mathbf{B}) - \operatorname{tr}(\mathbf{A}\mathbf{B})^2 + \operatorname{tr}(\mathbf{B}\mathbf{\Gamma})$. Note that by construction the Lagrangian multiplier $\mathbf{\Gamma}$ and $\mathbf{B}$ commute. The necessary optimality condition is [51, Equation (15)]

$$2\mathbf{A}^{1/2}\mathbf{B}\mathbf{A}^{1/2} - 2\operatorname{tr}(\mathbf{A}\mathbf{B})\mathbf{A} = \mathbf{\Gamma}. \qquad (A.12)$$

Multiply (A.12) from left and right with $\mathbf{B}^{1/2}$ to obtain

$$\mathbf{B}^{1/2}\mathbf{A}^{1/2}\big[\mathbf{B} - \operatorname{tr}(\mathbf{A}\mathbf{B})\mathbf{I}\big]\mathbf{A}^{1/2}\mathbf{B}^{1/2} = \mathbf{0}. \qquad (A.13)$$

From (A.13) it follows that the extreme points of $\mathbf{B}$ are either $\mathbf{B}^* = c\mathbf{I}$ for $c > 0$ or $\mathbf{A}^{1/2}\mathbf{B}^{*,1/2} = \mathbf{0}$. Note that g in (A.10) is zero for both values of $\mathbf{B}^*$, that is, $g(\mathbf{B}^*) = 0$.

Finally, we show that the extreme points $\mathbf{B}^*$ are indeed minima and therefore $g(\mathbf{B}) \geq 0$. Denote the eigenvalue decompositions of $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}_A\mathbf{U}^H$ and $\mathbf{B} = \mathbf{V}\mathbf{\Lambda}_B\mathbf{V}^H$, respectively. Define a new positive semidefinite matrix $\mathbf{C} = \mathbf{U}^H\mathbf{B}\mathbf{U}$ to write $g(\mathbf{C})$ as

$$g(\mathbf{C}) = \operatorname{tr}(\mathbf{\Lambda}_A^{1/2}\mathbf{C}\mathbf{\Lambda}_A^{1/2}\mathbf{C}) - \operatorname{tr}(\mathbf{\Lambda}_A\mathbf{C})^2. \qquad (A.14)$$

Note that the function g in (A.14) is invariant against the trace of $\mathbf{C}$ and we will assume without loss of generality that the $\operatorname{tr}(\mathbf{C}) = \delta$. Parameterize the unitary matrix by a skew

Hermitian $\mathbf{S} = -\mathbf{S}^H$ as $\exp(-\epsilon\mathbf{S})$. First, we show that the local minima are obtained for diagonal $\mathbf{C}$ by contradiction, that is, for the parameterized function $g(\epsilon)$,

$$\frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}g(\epsilon) = \frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}\operatorname{tr}\big(\mathbf{\Lambda}_A^{1/2}\exp(\epsilon\mathbf{S})\mathbf{B}\exp(-\epsilon\mathbf{S})$$
$$\times \mathbf{\Lambda}_A^{1/2}\exp(\epsilon\mathbf{S})\mathbf{B}\exp(-\epsilon\mathbf{S})\big)$$
$$- \operatorname{tr}\big(\mathbf{\Lambda}_A\exp(\epsilon\mathbf{S})\mathbf{B}\exp(-\epsilon\mathbf{S})\big)^2$$
$$= 2\operatorname{tr}(\mathbf{\Lambda}_A\mathbf{B})\big[\operatorname{tr}(\mathbf{\Lambda}_A\mathbf{S}\mathbf{B}) - \operatorname{tr}(\mathbf{\Lambda}_A\mathbf{B}\mathbf{S})\big]$$
$$= 2\operatorname{tr}(\mathbf{\Lambda}_A\mathbf{B})\operatorname{tr}\big([\mathbf{\Lambda}_A\mathbf{B} - \mathbf{B}\mathbf{\Lambda}_A]\mathbf{S}\big) \geq 0, \qquad (A.15)$$

where the last inequality follows from the choice $\mathbf{S} = \mathbf{B}\mathbf{\Lambda}_A - \mathbf{\Lambda}_A\mathbf{B}$. Equality holds when $\mathbf{B}$ and $\mathbf{\Lambda}_A$ commute. Therefore, the extreme points are obtained for diagonal $\mathbf{A}$ and diagonal $\mathbf{B}$. Define $\boldsymbol{\beta} = [\beta_1,\ldots,\beta_n] = \operatorname{diag}(\mathbf{\Lambda}_B)$, $\boldsymbol{\alpha} = [\alpha_1,\ldots,\alpha_n] = \operatorname{diag}(\mathbf{\Lambda}_A)$, and

$$g(\boldsymbol{\beta}) = \sum_{k=1}^n \alpha_k\beta_k^2 - \left(\sum_{k=1}^n \alpha_k\beta_k\right)^2. \qquad (A.16)$$

The last step is to show that $g(\boldsymbol{\beta})$ is Schur-convex with respect to $\boldsymbol{\beta}$. This implies that the minimum is obtained for $\mathbf{B}^* = c\mathbf{I}$. We verify Schur's condition [52, Lemma 2.1], that is, we sort $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_n$ and

$$\frac{\partial f(\boldsymbol{\beta})}{\partial\beta_1} - \frac{\partial f(\boldsymbol{\beta})}{\partial\beta_2} = 2\alpha_1\beta_1 - 2\alpha_2\beta_2 - 2\gamma\alpha_1\beta_1 + 2\gamma\alpha_2\beta_2$$
$$\geq 2\alpha_1\beta_1 - 2\alpha_1\beta_1 - 2\alpha_2\beta_2 + 2\alpha_2\beta_2 = 0 \qquad (A.17)$$

with $\gamma = \sum_{k=1}^n \alpha_k\beta_k$. The inequality follows because $\sum_{k=1}^n \alpha_k = 1$ and we can choose $\sum_{k=1}^n \beta_k = \delta \ll 1$ arbitrary small.

Hence, the first derivative at the point $\mathbf{Q}^\ell$ in direction of the update in (6) is positive

$$\frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}\Psi(\mathbf{A}+\epsilon\Delta) \geq 0. \qquad (A.18)$$

The remaining step is to show that the first directional derivative at the updated point $\mathbf{Q}^{\ell+1}$ in the direction of $\mathbf{Q}^\ell$ is nonpositive. This can be shown by the same procedure as above starting from (A.8). By concavity of the objective function follows that there cannot be the maximum between $\mathbf{Q}^\ell$ and $\mathbf{Q}^{\ell+1}$. Therefore, every fix-point algorithm defined in the way as (6) works for concave objective functions.

B. PROOF OF LEMMA 1

The proof follows along the same line as the proof of Theorem 1. The optimality conditions are given by

$$\mathbf{Q}_m = \frac{1}{\mu}\mathbf{Q}_m^{1/2}\Psi'(\mathbf{Q}_m)\mathbf{Q}_m^{1/2} \quad \text{for } 1 \leq m \leq M \qquad (B.1)$$

with μ as the Lagrangian multiplier for the sum power constraint. Therefore, there is only one fix point which corresponds to the global maximizer of (7). The monotonicity of

the update rule is shown by the same arguments as in the proof of Theorem 1 with the super covariance matrix

$$\check{\mathbf{Q}} = \begin{pmatrix} \mathbf{Q}_1 & 0 & \dots & 0 \\ 0 & \mathbf{Q}_2 & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & \dots & \mathbf{Q}_M \end{pmatrix} \qquad (B.2)$$

of size $M \times M$. The update rule in **Algorithm 1** can be simultaneously expressed for all transmit covariance matrices of all chunks super channel matrix

$$\check{\mathbf{H}}_{\mathbf{b}} = \mathrm{diag}[\mathbf{H}_{1,b}, \dots, \mathbf{H}_{M,b}] \qquad (B.3)$$

and super noise plus interference covariance matrix

$$\check{\mathbf{Z}}_{\mathbf{b}} = \mathrm{diag}[\mathbf{Z}_{1,b}, \dots, \mathbf{Z}_{M,b}] \qquad (B.4)$$

by

$$\check{\mathbf{Q}}^{\ell+1} = \frac{P \sum_{b=1}^{B} c_b \left(\mathbf{I} - \check{\mathbf{Z}}_b^{1/2} \left[\check{\mathbf{Z}}_b + \check{\mathbf{H}}_b \check{\mathbf{Q}}^{\ell} \check{\mathbf{H}}_b^H \right]^{-1} \check{\mathbf{Z}}_b^{1/2} \right)}{\mathrm{tr} \left(\sum_{b=1}^{B} c_b \left(\mathbf{I} - \check{\mathbf{Z}}_b^{1/2} \left[\check{\mathbf{Z}}_b + \check{\mathbf{H}}_b \check{\mathbf{Q}}^{\ell} \check{\mathbf{H}}_b^H \right]^{-1} \check{\mathbf{Z}}_b^{1/2} \right) \right)}. \qquad (B.5)$$

The update rule in (B.5) assures that the resulting $\check{\mathbf{Q}}^{\ell+1}$ is also block-diagonal and positive semidefinite and fulfills the trace constraint. Therefore, the same arguments as in the proof of Theorem 1 in Appendix A apply.

ACKNOWLEDGMENTS

REFERENCES

[1] A. J. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE Jourunal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–702, 2003.

[2] S. A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Transactions on Wireless Communications*, vol. 3, no. 4, pp. 1165–1175, 2004.

[3] W. Rhee and J. M. Cioffi, "On the capacity of multiuser wireless systems with multiple antennas," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2580–2595, 2003.

[4] A. Pascual-Iserte, A. I. Pérez-Neira, and M. A. Lagunas, "An approach to optimum joint beamforming design in a MIMO-OFDM multiuser system," *EURASIP Journal on Wireless Communications and Networking*, vol. 2004, no. 2, pp. 210–221, 2004.

[5] Y. J. Zhang and K. B. Letaief, "An efficient resource-allocation scheme for spatial multiuser access in MIMO/OFDM systems," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 107–116, 2005.

[6] Y.-C. Liang, R. Zhang, and J. M. Cioffi, "Transmit optimization for MIMO-OFDM with delay-constrained and no-delayconstrained traffics," *IEEE Transactions on Signal Processing*, vol. 54, no. 8, pp. 3190–3199, 2006.

[7] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 5045–5060, 2006.

[8] M. Bengtsson, "From single link MIMO to multi-user MIMO," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '04)*, vol. 4, Montreal, Quebec, Canada, May 2004.

[9] D. Gesbert, M. Kountouris, R. Heath, C.-B. Chae, and T. Salzer, "From single user to multiuser communications: shifting the MIMO paradigm," *IEEE Signal Processing Magazine*, vol. 24, no. 5, p. 36, 2007.

[10] M. Kobayashi and G. Caire, "An iterative water-filling algorithm for maximum weighted sum-rate of gaussian MIMO-BC," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1640–1646, 2006.

[11] M. Kobayashi and G. Caire, "Iterative waterfilling for weighted rate sum maximization in MIMO-OFDM broadcast channels," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '07)*, vol. 3, pp. 5–8, 2007.

[12] F. Boccardi, F. Tosato, and G. Caire, "Precoding schemes for the MIMO-GBC," in *Proceedings of the International Zurich Seminar on Digital Communications*, vol. 2006, pp. 10–13, 2006.

[13] A. Bayesteh and A. K. Khandani, "How much feedback is required in MIMO broadcast channels?" in *Proceedings of the IEEE International Symposium on Information Theory, (ISIT '06)*, pp. 1310–1314, Seattle, Wash, USA, July 2006.

[14] W. Dai, B. Rider, and Y. Lui, "Multi-access MIMO systems with finite rate channel state feedback," in *Proceedings of the Allerton Conference*, Monticello, Ind, USA, october 2005.

[15] T. Ratnarajah, "Limits of multi-user MIMO systems using scheduling and rate feedback," *Signal Processing*, vol. 87, no. 9, pp. 2165–2176, 2007.

[16] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for gaussian vector multiple-access channels," *IEEE Transactions on Information Theory*, vol. 50, no. 1, pp. 145–152, 2004.

[17] G. Wunder and T. Michel, "Optimal resource allocation for parallel gaussian broadcast channels: minimum rate constraints and sum power minimization," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4817–4822, 2007.

[18] A. Soysal and S. Ulukus, "Transmit directions and optimality of beamforming in MIMO-MAC with partial CSI at the transmitters," in *Proceedings of the Conference on Information Sciences and Systems (CISS '05)*, Baltimore, Md, USA, March 2005.

[19] E. A. Jorswieck, A. Sezgin, H. Boche, and E. Costa, "Multiuser MIMO MAC with statistical CSI and MMSE receiver: feedback strategies and transmitter optimization," in *Proceedings of the International Wireless Communications and Mobile Computing Conference (IWCMC '06)*, vol. 2006, pp. 455–460, 2006.

[20] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of gaussian MIMO broadcast channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2658–2668, 2003.

[21] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector gaussian broadcast channel and uplink-downlink duality," *IEEE Transactions on Information Theory*, vol. 49, no. 8, pp. 1912–1921, 2003.

[22] E. A. Jorswieck, "Lack of duality between SISO gaussian MAC and BC with statistical CSIT," *Electronics Letters*, vol. 42, no. 25, pp. 1466–1468, 2006.

[23] M. Sternad, T. Svensson, and G. Klang, "The WINNER B3G system MAC concept," in *Proceedings of the IEEE Vehicular Technology Conference*, pp. 3037–3041, 2006.

[24] S. Olonbayar and H. Rohling, "Multiuser diversity and subcarrier allocation in OFDM-FDMA systems," in *Proceedings of the OFDM Workshop*, San Francisco, Calif, USA,, February 2005.

[25] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1211–1226, 2002.

[26] IST-4-027756 WINNER II, "D1.1.1 WINNER II interim channel models," 2006.

[27] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, Cambridge, UK, 2005.

[28] K. Seong, R. Narasimhan, and J. M. Cioffi, "Queue proportional scheduling via geometric programming in fading broadcast channels," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1593–1602, 2006.

[29] H. Boche and M. Wiczanowski, "Optimization-theoretic analysis of stability-optimal transmission policy for multiple-antenna multiple-access channel," *IEEE Transactions on Signal Processing*, vol. 55, no. 6 I, pp. 2688–2702, 2007.

[30] S. P. Wu, L. Vandenberghe, and S. Boyd, "MAXDET: software for determinant maximization problems," http://www.stanford.edu/~boyd/papers/maxdet.html, 1996.

[31] Y. Labit, D. Peaucelle, and D. Henrion, "SeDuMi interface 1.02: a tool for solving LMI problems with SeDuMi," in *Proceedings of the CACSD Conference*, London, UK, September 2002.

[32] E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[33] E. A. Jorswieck, A. Sezgin, Y. Goa, and D. Kim, "On the tradeoff between feedback signaling and performance for tile based MIMO-OFDMA systems," in *Proceedings of the 6th International Symposium on Wireless Personal Multimedia Communications (WPMC '06)*, Abano Terme, Italy, September 2006.

[34] L. W. Hanlen and A. J. Grant, "Optimal transmit covariance for MIMO channels with statistical transmitter side information," in *IEEE International Symposium on Information Theory (ISIT 2005)*, pp. 1818–1822, Adelaide, South Australia, September 2005.

[35] C. T. Kelly, *Iterative Methods for Linear and Nonlinear Equations*, SIAM, Philadelphia, Pa, USA, 1995.

[36] J. C. Bezdek and R. J. Hathaway, "Some notes on alternating optimization," in *Proceedings of the Advances in Soft Computing—AFSS 2002: 2002 AFSS International Conference on Fuzzy Systems*, vol. 2275 of *Lecture Notes in Computer Science*, pp. 187–195, Calcutta, India, February 2002.

[37] D. J. Love, R. W. Heath Jr., W. Santipach, and M. L. Honig, "What is the value of limited feedback for MIMO channels?" *IEEE Communications Magazine*, vol. 42, no. 10, pp. 54–59, 2004.

[38] D. J. Love, R. W. Heath Jr., and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2735–2747, 2003.

[39] W. Santipach and M. Honig, "Capacity of a multi-antenna fading channel with quantized precoding matrix," to appear in *IEEE Transactions on Information Theory* http://arxiv.org/abs/0704.0217.

[40] IST-4-027756 WINNER II, "D6.13.7 test scenarios and calibration cases issue 2," December 2006.

[41] J. H. Winters, J. Salz, and R. D. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Transactions on Communications*, vol. 42, no. 2, pp. 1740–1751, 1994.

[42] J. M. Cioffi, *Digital Communication*, Lecture Notes, Stanford University, Stanford, Calif, USA, 2004.

[43] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communication*, Kluwer, Norwell, Mass, USA, 3rd edition, 2004.

[44] E. A. Jorswieck, W. B. Chamekh, and M. Weckerle, "Optimal chunk processing for multi-user MIMO OFDM wireless systems," in *Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '06)*, Helsinki, Finland, September 2006.

[45] W. Rhee, W. Yu, and J. Cioffi, "The optimality of beamforming in uplink multiuser wireless systems," *IEEE Transactions on Wireless Communications*, vol. 3, no. 1, pp. 86–96, 2004.

[46] W. Yu and W. Rhee, "Degrees of freedom in wireless multiuser spatial multiplex systems with multiple antennas," *IEEE Transactions on Communications*, vol. 54, no. 10, pp. 1747–1753, 2006.

[47] T. T. Kim, M. Bengtsson, E. G. Larsson, and M. Skoglund, "Combining long-term and low rate short-term channel state information over correlated MIMO channels," to appear in *IEEE Transactions on Wireless Communications*.

[48] G. Jöngren and M. Skoglund, "Quantized feedback information in orthogonal space-time block coding," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2473–2482, 2004.

[49] N. Farvardin, "A study of vector quantization for noisy channels," *IEEE Transactions on Information Theory*, vol. 36, no. 4, pp. 199–809, 1990.

[50] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004, http://www.stanford.edu/~boyd/cvxbook/.

[51] T. Minka, "Old and new matrix algebra useful for statistics," http://research.microsoft.com/~minka/papers/matrix/, 2000.

[52] A. Jorswieck and H. Boche, "Majorization and matrix monotone functions in wireless communications," in *Foundations and Trends in Communications and Information Theory*, S. Verdú, Ed., vol. 3, no. 6, pp. 553–701, Now, Hanover, Mass, USA, 2007.