

Research Article

A Design Framework for Scalar Feedback in MIMO Broadcast Channels

Ruben de Francisco and Dirk T. M. Slock

Eurecom Institute, BP 193, 06904 Sophia-Antipolis Cedex, France

Correspondence should be addressed to Ruben de Francisco, ruben.defrancisco@ieee.org

Received 15 June 2007; Revised 6 October 2007; Accepted 13 November 2007

Recommended by Markus Rupp

Joint linear beamforming and scheduling are performed in a system where limited feedback is present at the transmitter side. The feedback conveyed by each user to the base station consists of channel direction information (CDI) based on a predetermined codebook and a scalar metric with channel quality information (CQI) used to perform user scheduling. In this paper, we present a design framework for scalar feedback in MIMO broadcast channels with limited feedback. An approximation on the sum rate is provided for the proposed family of metrics, which is validated through simulations. For a given number of active users and average SNR conditions, the base station is able to update certain transmission parameters in order to maximize the sum-rate function. On the other hand, the proposed sum-rate function provides a means of simple comparison between transmission schemes and scalar feedback techniques. Particularly, the sum rate of SDMA and time division multiple access (TDMA) is compared in the following extreme regimes: large number of users, high SNR, and low SNR. Simulations are provided to illustrate the performance of various scalar feedback techniques based on the proposed design framework.

Copyright © 2008 R. de Francisco and D. T. M. Slock. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

Multiple-input multiple-output (MIMO) systems can significantly increase the spectral efficiency by exploiting the spatial degrees of freedom created by multiple antennas. In point-to-point MIMO systems, the capacity increases linearly with the minimum of the number of transmit/receive antennas, irrespective of the availability of channel state information (CSI) [1, 2]. In the MIMO broadcast channel, it has recently been proven [3] that the sum capacity is achieved by dirty paper coding (DPC) [4]. However, the applicability of DPC is limited due to its computational complexity and the need for full channel state information at the transmitter (CSIT). Downlink techniques based on space division multiple access (SDMA) have been proposed [5], achieving the same asymptotic sum rate as that of DPC.

The capacity gain of multiuser MIMO systems is highly dependent on the available CSIT. While having full CSI at the receiver can be assumed, this assumption is not reasonable at the transmitter side. Several limited feedback approaches have been considered in point-to-point systems [6–8], where

each user sends to the transmitter the index of a quantized version of its channel vector from a codebook. An extension for MIMO broadcast channels is made in [9], in which each mobile feeds back a finite number of bits regarding its channel realization at the beginning of each block based on a codebook.

Besides channel direction information (CDI), we consider limited feedback scenarios in which each user conveys channel quality information (CQI) to the base station for the purpose of user scheduling. In [10], an SDMA extension of opportunistic beamforming [11] using partial CSIT in the form of individual signal-to-interference-plus-noise ratio (SINR) is proposed, achieving optimum capacity scaling for large number of users. A simple scheme for joint scheduling and beamforming with limited feedback is proposed in [12, 13]. The receivers compute and feed back a scalar metric that can be interpreted as an upper bound on the SINR. Note that a scheme with similar metric is also reported in [14]. Assuming certain orthogonality constraints between beamforming vectors, a lower bound on the instantaneous or average SINR can be computed as scalar feedback, as shown in

[15, 16], respectively. The total amount of feedback overhead in the system can be reduced by appropriately setting minimum desired SINR thresholds while controlling each user's quality of service (QoS). A performance comparison of several scalar metrics for scheduling is provided in [17] for systems with zero-forcing beamforming (ZFBF) transmission.

In this paper, we present a design framework for scalar feedback in MIMO broadcast channels, which generalizes previously proposed techniques. A family of metrics is presented based on individual SINRs, which are computed at the receivers and fed back to the base station as channel quality information. The framework here presented can be applied to any system in which codebooks are employed for channel direction quantization. Moreover, additional orthogonality constraints between beamforming vectors may be considered with the purpose of simplifying the task of user scheduling and controlling the amount of multiuser interference.

An approximation on the ergodic sum rate is provided for the proposed family of metrics. The resulting sum-rate function fits well the simulated sum rate as shown through simulations, even in cells with reduced number of active users. This function, as we show, can be a powerful design tool and at the same time it greatly simplifies system analysis. On the one hand, we can envisage a cellular system in which, given certain average SNR conditions and number of active users, the base station sets the different parameters so as to maximize the sum-rate function. On the other hand, as shown in the analysis, the sum-rate function provides a means of simple comparison between different transmission schemes and scalar feedback techniques in extreme regimes, without the need of extreme value theory. Particularly, we compare the sum rate of SDMA and TDMA approaches in scenarios with large number of users, high SNR, and low SNR regimes. Simulations are provided to illustrate the performance of different scalar feedback techniques based on the proposed design framework.

The paper is organized as follows. Section 2 introduces the system model. Linear beamforming with limited feedback is introduced in Section 3, presenting system assumptions on codebook design, beamforming design, and user scheduling. Design guidelines for scalar feedback are given in Section 4 and the corresponding sum-rate function is provided in Section 5. Section 6 shows a comparison of SDMA and TDMA in different extreme regimes, namely, large number of users, high SNR, and low SNR. Section 7 shows numerical results and conclusions are drawn in Section 8.

2. SYSTEM MODEL

We consider a multiple antenna broadcast channel consisting of M antennas at the transmitter and $K \geq M$ single-antenna receivers. The received signal y_k of the k th user is mathematically described as

$$y_k = \mathbf{h}_k^H \mathbf{x} + n_k, \quad k = 1, \dots, K, \quad (1)$$

where $\mathbf{x} \in \mathbb{C}^{M \times 1}$ is the transmitted signal, $\mathbf{h}_k \in \mathbb{C}^{M \times 1}$ is an i.i.d. Rayleigh flat fading channel vector, and n_k is additive white Gaussian noise at receiver k . We assume that each

of the receivers has perfect and instantaneous knowledge of its own channel \mathbf{h}_k , and that n_k is independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian with zero mean and variance $\sigma^2 = 1$. The transmitted signal is subject to an average transmit power constraint P , that is, $\mathbb{E}\{\|\mathbf{x}\|^2\} = P$. Note that, since unit-variance noise is assumed, P takes on the meaning of average SNR. Let \mathcal{S} denote the set of users selected for transmission at a given time slot, with cardinality $|\mathcal{S}| = M_o$, $1 \leq M_o \leq M$. Let \mathbf{v}_k be the unit-norm beamforming vector for user k . Assuming equal power allocation to the M_o scheduled users, the received signal at the k th mobile is given by

$$y_k = \sqrt{\frac{P}{M_o}} \sum_{i \in \mathcal{S}} \mathbf{h}_k^H \mathbf{v}_i s_i + n_k, \quad k = 1, \dots, K. \quad (2)$$

Hence, the SINR of user k is

$$\text{SINR}_k = \frac{|\mathbf{h}_k^H \mathbf{v}_k|^2}{\sum_{i \in \mathcal{S}, i \neq k} |\mathbf{h}_k^H \mathbf{v}_i|^2 + M_o/P}. \quad (3)$$

We focus on the ergodic sum rate (SR) which, assuming Gaussian inputs, is equal to

$$\text{SR} = \mathbb{E} \left\{ \sum_{k \in \mathcal{S}} \log[1 + \text{SINR}_k] \right\}. \quad (4)$$

Notation: We use bold upper and lower case letters for matrices and column vectors, respectively. $(\cdot)^H$ stands for Hermitian transpose. $\mathbb{E}(\cdot)$ denotes the expectation operator. The notation $\|\mathbf{x}\|$ refers to the Euclidean norm of the vector \mathbf{x} , and $\angle(\mathbf{x}, \mathbf{y})$ refers to the angle between vectors \mathbf{x} and \mathbf{y} .

3. LINEAR BEAMFORMING WITH LIMITED FEEDBACK

Joint linear beamforming and scheduling are performed in a system where limited feedback is present at the transmitter side. The feedback conveyed by each user to the base station consists of channel direction information based on a predetermined codebook and a scalar metric with channel quality information used to perform user scheduling.

In such systems, the design of appropriate scalar metrics in scenarios with realistic number of users and average SNR values remains a challenge. These metrics must contain information of the users' channel gains as well as channel quantization errors, as discussed in [18]. If the users have additional knowledge of the beamforming technique used at the transmitter side, an estimate on the multiuser interference at the receiver can be computed. This information can be encapsulated together with the channel gain, quantization error, and average noise power into a scalar metric ξ , which consists of an estimate on the SINR. In our work, we consider such scalar feedback strategies, as discussed in detail in next section. User selection is carried out based on these metrics and the users' spatial properties, obtained from channel quantizations.

As simple transmission technique we consider transmit matched filtering (TxMF) which consists of using as normalized beamforming vectors the quantized channel directions

```

MS
Compute & Feedback  $\xi_k$ 
    quantization index  $i \in \{1, \dots, L\}$ 
BS
Initialize Set  $\mathcal{S} = \emptyset$ 
Loop For  $i : 1, \dots, M_o$  repeat
    Set  $\xi_{\max}^i = 0$ 
    Loop For  $k : 1, \dots, K, k \notin \mathcal{S}$  repeat
        If  $\xi_k > \xi_{\max}^i$  and  $|\mathbf{v}_k^H \mathbf{v}_j| \leq \epsilon \ \forall j \in \mathcal{S}$ 
             $\xi_k \rightarrow \xi_{\max}^i$  and  $k_i = k$ 
    Select  $k_i \rightarrow \mathcal{S}$ 

```

ALGORITHM 1: Outline of scheduling algorithm.

of users scheduled for transmission. The normalized channel vector of user k to be quantized is $\mathbf{h}_k = \mathbf{h}_k / \|\mathbf{h}_k\|$, which corresponds to the channel direction. A B -bit quantization codebook \mathcal{V}_k is considered, containing $L = 2^B$ unit norm vectors in \mathbb{C}^M , which is assumed to be known to both the receiver and the transmitter. Similar to [7, 8], we assume that each receiver quantizes its channel to the vector that maximizes the inner product

$$\mathbf{v}_k = \arg \max_{\mathbf{v} \in \mathcal{V}_k} |\mathbf{h}_k^H \mathbf{v}|^2 = \arg \max_{\mathbf{v} \in \mathcal{V}_k} \cos^2(\angle(\mathbf{h}_k, \mathbf{v})). \quad (5)$$

Each user sends the corresponding quantization index back to the transmitter through an error-free and zero-delay feedback channel using B bits. Note that this model is equivalent to the finite rate feedback model proposed by [7, 9].

The optimal vector quantizer is difficult to find and the solution to this problem is not yet known. As codebook design goes beyond the scope of the paper, we adopt the geometrical framework presented in [8]. The resulting quantization error is defined as $\sin^2 \theta_k = \sin^2(\angle(\mathbf{h}_k, \mathbf{v}_k)) = 1 - |\mathbf{h}_k^H \mathbf{v}_k|^2$ [8, 19], where \mathbf{v}_k is the quantized channel direction of user k . Using this framework, the cumulative distribution function (cdf) of the quantization error is given by [8, 19],

$$F_{\sin^2 \theta_k}(x) = \begin{cases} \delta^{1-M} x^{M-1}, & 0 \leq x \leq \delta, \\ 1, & x > \delta, \end{cases} \quad (6)$$

where $\delta = 2^{-B/(M-1)}$.

Let the orthogonality factor ϵ denote the maximum degree of nonorthogonality between two unit-norm vectors. The columns of the normalized beamforming matrix $\mathbf{V}(\mathcal{S})$ are constrained to be ϵ -orthogonal and thus

$$|\mathbf{v}_i^H \mathbf{v}_j| \leq \epsilon \quad \forall i, j \in \mathcal{S}, i \neq j. \quad (7)$$

An outline of the proposed scheduling algorithm is shown in Algorithm 1. In case M_o users with ϵ -orthogonality cannot be found, the algorithm stops and distributes the power equally among the scheduled users, setting $M_o = |\mathcal{S}|$. Note that this greedy algorithm is equivalent to the one proposed in [5, 20, 21]. The first user is selected from the set $\mathcal{Q}^0 = \{1, \dots, K\}$ as the one having the highest channel quality, that

is, $k_1 = \arg \max_{k \in \mathcal{Q}^0} \xi_k$. For $i = 1, \dots, M_o - 1$, the $(i + 1)$ th user is selected as $k_{i+1} = \arg \max_{k \in \mathcal{Q}^i} \xi_k$ among the user set $\mathcal{Q}^i = \{1 \leq k \leq K : |\mathbf{v}_k^H \mathbf{v}_j| \leq \epsilon, 1 \leq j \leq i\}$.

The number of active beams for transmission M_o and orthogonality factor ϵ is system parameters fixed by the base station (BS) that can be adapted in order to maximize the system sum rate.

4. SCALAR FEEDBACK DESIGN

In this section, we present design guidelines for scalar metrics based on signal-to-interference-plus-noise ratios, which are computed at the receivers and fed back to the base station as channel quality information. Complemented with channel quantizations as CDI, user scheduling at the base station of a MIMO broadcast channel is performed. The design framework for scalar feedback here presented can be applied to any system in which codebooks are employed for channel quantization, known both to the base station and mobile users.

These metrics must contain information of different nature in order to exploit the multiuser diversity of the MIMO broadcast channel. Moreover, additional information on the orthogonality constraints between beamforming vectors can be taken into account, thus providing a QoS estimate at the receiver side. The total amount of feedback overhead can be reduced by appropriately setting minimum desired SINR thresholds. Hence, in a practical system each user may send feedback to the base station only if a minimal QoS can be guaranteed.

Besides signal and noise power, the following information may be encapsulated by each user in such scalar metrics:

- (i) channel power gain: $\|\mathbf{h}_k\|^2$,
- (ii) quantization error: $\sin^2 \theta_k$,
- (iii) orthogonality factor: ϵ ,
- (iv) number of active beams: M_o .

As shown in [18], channel power gain and quantization error information are necessary in order to exploit the available multiuser diversity. The quantization error is a function of the number of codebook bits, as shown in the previous section. By increasing the codebook size, the multiplexing gain of the system can be increased (better resolution) and at the same time the multiuser diversity gets increased, due to lower quantization error. The orthogonality factor ϵ can be used to bound the amount of expected multiuser interference, which in turn can be used to compute a lower bound on the SINR. In our work, we assume that the number of active beams (nonzero power) is a parameter appropriately set by the base station to maximize the system sum rate.

Multiuser interference

For user k and index set \mathcal{S} , the multiuser interference can be expressed as $I_k(\mathcal{S}) = \sum_{i \in \mathcal{S}, i \neq k} (P/M_o) |\mathbf{h}_k^H \mathbf{v}_i|^2 = (P/M_o) \|\mathbf{h}_k\|^2 I_k(\mathcal{S})$, where $I_k(\mathcal{S})$ denotes the interference over the normalized channel \mathbf{h}_k . Let $\mathbf{U}_k \in \mathbb{C}^{M \times (M-1)}$ be an orthonormal basis spanning the null space of \mathbf{v}_k^H and define the matrix $\Psi_k = \sum_{i \in \mathcal{S}, i \neq k} \mathbf{v}_i \mathbf{v}_i^H$ and the operator $\lambda_{\max}\{\cdot\}$,

which returns the largest eigenvalue. Define I_{UBk} as the upper bound on I_k and $\theta_k = \angle(\mathbf{h}_k, \mathbf{v}_k)$. As proven in [18] for systems with arbitrary orthogonality between beamforming vectors, the multiuser interference of user k can be bounded as follows:

$$I_{UBk} = \alpha_k \cos^2 \theta_k + \beta_k \sin^2 \theta_k + 2\gamma_k \sin \theta_k \cos \theta_k, \quad (8)$$

where

$$\begin{aligned} \alpha_k &= \mathbf{v}_k^H \Psi_k \mathbf{v}_k, \\ \beta_k &= \lambda_{\max} \{ \mathbf{U}_k^H \Psi_k \mathbf{U}_k \}, \\ \gamma_k &= \| \mathbf{U}_k^H \Psi_k \mathbf{v}_k \|. \end{aligned} \quad (9)$$

Family of metrics

In the proposed design framework, any scalar feedback metric can be described as follows:

$$\xi = \frac{\| \mathbf{h}_k \|^2 \cos^2 \theta_k}{\| \mathbf{h}_k \|^2 (\alpha \cos^2 \theta_k + \beta \sin^2 \theta_k + 2\gamma \sin \theta_k \cos \theta_k) + M_o/P}. \quad (10)$$

The numerator in the expression above reflects the effective received power in a system with channel quantization. On the other hand, the denominator accounts for the noise power and provides a measure of the interference experienced by the user, for instance, an upper or lower bound, by exploiting the structure of the beamforming matrix. By choosing different values for the parameters α , β , γ , and M_o , the meaning of the proposed metric is modified, yielding different SINR measures. In next section, a sum-rate function is derived based on this metric structure, for arbitrary values of these parameters. When setting these parameters as in (9), the metric ξ becomes a lower bound for the SINR described in (3). Note that, even though ϵ -orthogonality beamformers are imposed at the transmitter, we may choose not to include this information in the scalar feedback metric. In addition, even though M_o is in principle a parameter that may be modified by the base station, a simplified case with $M_o = M$ may be considered for feedback design.

In the remainder of this section we present several scalar metrics complying with this structure.

Metric 1. Let \mathbf{u}_{jk} be the j th column vector of the matrix \mathbf{U}_k . The vector \mathbf{u}_{jk} is isotropically distributed over an $M - 1$ dimensional hyperplane orthogonal to \mathbf{v}_k , under the assumption that \mathbf{v}_k is isotropically distributed over the unit norm hypersphere. Given a fixed unit-norm vector \mathbf{v}_i in \mathbb{C}^M , the random variable $|\mathbf{v}_i^H \mathbf{u}_{jk}|^2$ follows a beta distribution with parameters $(1, M - 2)$ [22]. The mean value of this random variable is $1/(M - 1)$, and thus we have that $\mathbb{E}[\sum_{i=1, i \neq k}^{M_o} |\mathbf{v}_i^H \mathbf{u}_{jk}|^2] = (M_o - 1)/(M - 1)$. Using this result in (9) and the fact that nonorthogonality between pairs of

beamforming vectors is upper bounded by ϵ , we propose in [18] the following values for this metric:

$$\begin{aligned} \alpha &= \frac{(M_o - 1)^2}{M - 1} \epsilon^2, & \beta &= \frac{(M_o - 1)}{M - 1} [1 + (M_o - 2)\epsilon], \\ \gamma &= \frac{(M_o - 1)^2}{M - 1} \epsilon, & 1 &\leq M_o \leq M. \end{aligned} \quad (11)$$

Note that averaging the inverse of the resulting metric yields an upper bound on the average of the inverse SINR. Hence, the average value of this metric tends to be a lower bound on the average SINR.

Metric 2. As a particular case, we consider $\epsilon = 0$ in the metric computation and assume a fixed number of active beams

$$\begin{aligned} \alpha &= 0, & \beta &= 1, \\ \gamma &= 0, & M_o &= M. \end{aligned} \quad (12)$$

This metric can be interpreted as an upper bound on the SINR when exactly $M_o = M$ beams are used for transmission and equal power allocation is performed. Note that this metric was proposed in parallel in [12–14].

Metric 3. Another option consists of computing a lower bound on the instantaneous SINR [15]. As opposed to Metric 1, no averaging over the distribution of $|\mathbf{v}_k^H \mathbf{u}_{jk}|$ is performed and thus this lower bound is less tight in average. The metric parameters are given by

$$\begin{aligned} \alpha &= (M_o - 1)\epsilon^2, \\ \beta &= \begin{cases} 0, & \text{if } M_o = 1, \\ 1 + (M_o - 2)\epsilon, & \text{otherwise,} \end{cases} \\ \gamma &= (M_o - 1)\epsilon, \quad 1 \leq M_o \leq M. \end{aligned} \quad (13)$$

Taking into account ϵ in the SINR computation may mask the contribution of the channel power gains in the SINR expression, hence reducing the benefits of multiuser diversity. However, this approach offers the advantage of avoiding outage events in the communication link.

Metric 4. A straightforward improvement of Metric 2 can be done by setting a variable number of active beams $1 \leq M_o \leq M$, keeping the same values for α , β , and γ .

Note that, for a given scenario and feedback metric, there is an optimal pair of system parameters ϵ and M_o that maximizes the sum rate. Increasing the value of ϵ relaxes the ϵ -orthogonality constraint and thus more users are taken into account for scheduling, increasing the multiuser diversity benefit. However, as ϵ increases, so does the multiuser interference. On the other hand, increasing the number of active beams M_o exploits the spatial multiplexing gain, at the expense of increasing the interference. Hence, for a given average SNR and number of active users K in the cell, the base station must appropriately set ϵ and M_o in order to balance the multiuser diversity and multiplexing gains and to maximize the system sum rate. In practice, this may be carried

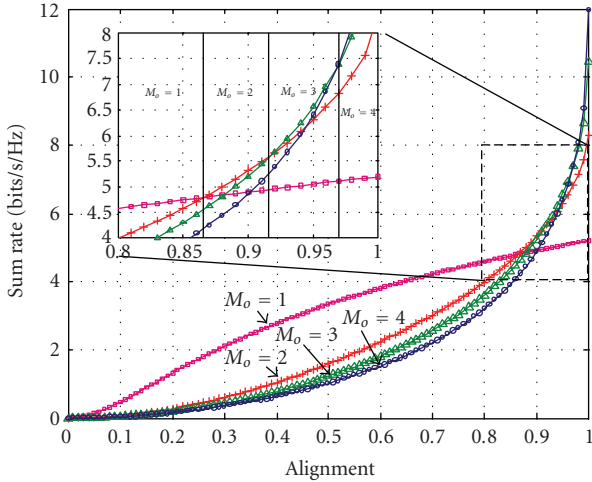


FIGURE 1: Approximated lower bound on the sum rate using Metric 1 versus the alignment $\cos \theta_k$ for $M = 4$ antennas, variable number of active beams M_o , orthogonality factor $\epsilon = 0.1$ and SNR = 10 dB.

out by storing lookup tables at the base station, so that ϵ and M_o can be quickly adapted whenever the average SNR or the number of active users changes. If the system parameters need to be updated, the base station broadcasts the new values to the users, which are used to compute the feedback metrics.

In Figure 1, an approximated lower bound on the system sum rate is plotted as a function of the alignment $\cos \theta_k$, computed as $SR \approx M_o \log(1 + \xi_k^I)$, where ξ_k^I denotes the feedback Metric 1 of user k . This approximation assumes that the M_o scheduled users have the same ξ_k^I value and thus the same estimated lower bound on the achievable rate. The system under consideration is assumed to have $M = 4$ antennas, $\epsilon = 0.1$, and average SNR = 10 dB. The sum rate is evaluated for different number of active beams to observe the impact of appropriately choosing M_o . Note that the case of $M_o = 1$ corresponds to TDMA, whereas $M_o > 1$ corresponds to SDMA. The system with $M_o = 1$ exhibits better performance for low and intermediate values of $\cos \theta_k$, that is, TDMA provides higher rates than SDMA in most cases. Only for large values of $\cos \theta_k$, $M_o > 1$ provides higher rates, which in practice occurs for large number of quantization bits B or large number of users K . Since the amount of bits B is generally low due to bandwidth limitations, SDMA will be chosen over TDMA when $M_o > 1$ users with small quantization errors can be found, with higher probability as the number of users in the cell increases. As the parameter ϵ increases, the crossing points of the curves in Figure 1 shift to the right and thus the range for which TDMA performs better also increases. This is due to the fact that the bound in ξ_k^I becomes looser for increasing ϵ values. As shown in this example, for $\epsilon > 0$ there exist M possible modes of transmission, that is, $M_o = 1, \dots, M$. However, for the case of $\epsilon = 0$ and varying M_o as considered in Metric 4, it can be proven that the modes of transmission exhibiting higher rates are reduced to 2, namely, $M_o = 1, M$.

5. SUM-RATE FUNCTION

In this section, we derive a function to approximate the ergodic sum rate that a system with linear beamforming and limited feedback can provide, given knowledge of each user’s SINR metric. A general and simple solution is derived based on the generic metric representation of ξ , given in (10). Note that the different metrics described in the previous section follow as particular cases of ξ by setting accordingly the values of α, β, γ , and M_o . The sum-rate function we provide is a tool that enables simple analysis and comparison of SDMA and TDMA approaches. Moreover, as shown in the simulations, it approximates well the system number even when the number of users in the cell is small. In our analysis, we are interested in the actual sum rate that can be achieved. Hence, the metric takes on the meaning of either an upper or lower SINR bound as needed in order to compare SDMA and TDMA in the extreme regimes under study.

First, an approximation on the cdf of ξ is derived, using mathematical tools from [23].

Proposition 1. *In the low-resolution regime (small B), the cdf of ξ can be approximated as follows:*

$$F_\xi(s) \approx 1 - \frac{e^{-M_o s/P(1-\alpha s)}}{\delta^{M-1} (1+m)^{M-1}}, \tag{14}$$

where $m = (2\gamma s[\gamma s + \sqrt{\gamma^2 s^2 + (1-\alpha s)\beta s}] + (1-\alpha s)\beta s)/(1-\alpha s)^2$.

Proof. See Appendix A. □

Note that the above cdf is a generalization for arbitrary ϵ and M_o of the cdf derived in [13]. Also, the result provided in [10] follows as a particular case by selecting $\epsilon = 0, M_o = M$, and $B = 0$.

Let the ordered variate $s_{i:K}$ denote the i th largest among K i.i.d. random variables. From known results of order statistics [24], we have that the cdf of $s_1 = \max_{1 \leq i \leq K} s_{i:K}$ is $F_{s_1} = (F_\xi(s))^K$. According to the proposed user selection algorithm, the SINR of the first-selected user is the maximum SINR over K i.i.d. random variables. However, at the i th selection step (i th beam) the search space gets reduced since the ϵ -orthogonality condition needs to be satisfied. Hence, the i th user is selected over K_i i.i.d. random variables yielding a cdf for the maximum SINR given by $F_{s_i} = (F_\xi(s))^{K_i}$. Since ξ is upper bounded by $1/\alpha$, its mean value is given by

$$\mathbb{E}(s_i) = \int_0^{1/\alpha} 1 - (F_\xi(s))^{K_i} ds. \tag{15}$$

An approximation of K_i can be calculated through the probability that a random vector in $\mathbb{C}^{M \times 1}$ is ϵ -orthogonal to a set with $i-1$ vectors in $\mathbb{C}^{M \times 1}$, which is equal to $I_{\epsilon^2}(i-1, M-i+1)$ [5], $I_x(a, b)$ being the regularized incomplete beta function. By using the law of large numbers [21], we can find the following approximation:

$$K_i \approx KI_{\epsilon^2}(i-1, M-i+1). \tag{16}$$

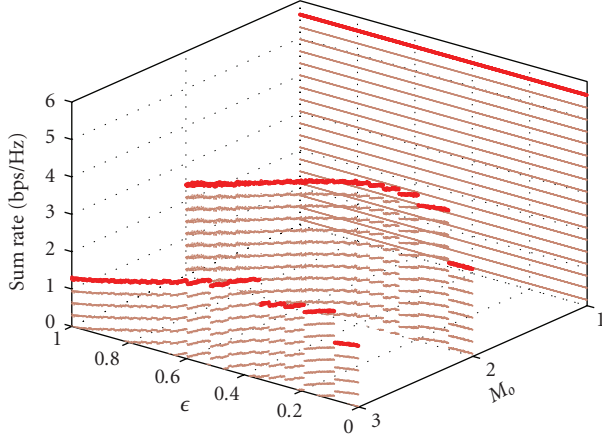


FIGURE 2: Sum-rate function using Metric 1 versus orthogonality factor ϵ and number of active beams M_o , for $K = 35$ users, SNR = 10 dB, and $B = 1$ bit.

The average sum rate in a system with M_o active beams can be bounded as follows by using Jensen's inequality:

$$\text{SR} = \sum_{i \in \mathcal{S}} \mathbb{E}[\log_2(1 + s_i)] \leq \sum_{i \in \mathcal{S}} \log_2[1 + \mathbb{E}(s_i)]. \quad (17)$$

Using (17) and solving the integral in (15) for the cdf of ξ described in (14), we obtain the following theorem after some approximations.

Theorem 1. Given ϵ -orthogonal transmission in a system with M_o active beams, the sum rate is approximated as follows:

$$R_{M_o} \approx \sum_{i=1}^{M_o} \log_2 \left[1 + \frac{1}{\alpha} \sum_{n=1}^{K_i} \mathcal{B}_n \mathcal{K}_{i,n} \mathcal{P}_n \right], \quad (18)$$

where

$$\begin{aligned} \mathcal{B}_n &= \frac{(-1)^{n-1}}{\delta^{n(M-1)}}, \\ \mathcal{K}_{i,n} &= \binom{K_i}{n}, \\ \mathcal{P}_n &= 1 + \frac{Cn}{\alpha} e^{Cn/\alpha} E_i \left(-\frac{Cn}{\alpha} \right), \end{aligned} \quad (19)$$

and $C = M_o/P + (M-1)\beta$. The exponential integral function is defined as $E_i(x) = -\int_{-x}^{\infty} (e^{-t}/t) dt$.

Proof. See Appendix B. \square

Note that the term \mathcal{B}_n reflects the influence of the codebook design, $\mathcal{K}_{i,n}$ together with the summation upper limit K_i inside the logarithm capture the amount of multiuser diversity exploited by the system and \mathcal{P}_n accounts for the dependency of the sum rate on the power.

Note that as a particular case of the equation above, a simpler expression can be derived for $M_o = 1$, given by

$$R_1 \approx \log_2 \left[1 + \sum_{n=1}^K \mathcal{B}_n \mathcal{K}_{1,n} \frac{P}{n} \right]. \quad (20)$$

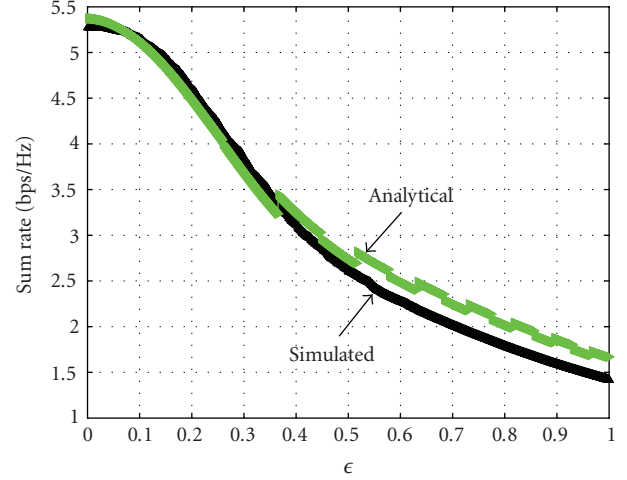


FIGURE 3: Comparison of analytical and simulated lower bounds on the sum rate using Metric 3, for $M = 2$ antennas, $K = 15$ users, SNR = 10 dB, and $B = 1$ bit.

Another case of interest is the case in which $\alpha = 0$. As α approaches zero, we have

$$\lim_{\alpha \rightarrow 0} \frac{1}{\alpha} \left[1 + \frac{Cn}{\alpha} e^{Cn/\alpha} E_i \left(-\frac{Cn}{\alpha} \right) \right] = \frac{1}{Cn}, \quad (21)$$

and thus the sum-rate function in this case becomes

$$\lim_{\alpha \rightarrow 0} R_{M_o} = \sum_{i=1}^{M_o} \log_2 \left[1 + \sum_{n=1}^{K_i} \mathcal{B}_n \mathcal{K}_{i,n} \frac{1}{Cn} \right]. \quad (22)$$

In Figure 2, the sum-rate function in (18) is plotted as a function of the number of active beams M_o and orthogonality factor ϵ , using the values for α , β , and γ as described in Metric 1. In this simulation, a system with $K = 35$ users has been considered, an average SNR = 10 dB and a simple codebook with $B = 1$ bit. Note that in this particular scenario, SDMA cannot guarantee better rates than TDMA regardless of the value of ϵ . In this context, the number of users is low, hence there is low probability of obtaining large values of $\cos \theta_k$. Thus, TDMA transmission is favored, which is consistent with the results obtained in the previous section.

In order to validate the obtained sum-rate function, we consider a simple scenario with $M = 2$ antennas and a system in which $M_o = 2$ if 2 ϵ -orthogonal users can be found in a given time slot and $M_o = 1$ otherwise. The probability of not finding 2 ϵ -orthogonal users is given by $p = [1 - \epsilon^2]^{K-1}$. Hence, the approximated rate in this simplified scenario is given by

$$R \approx pR_1 + (1-p)R_2, \quad (23)$$

where R_1 and R_2 (R_{M_o} with $M_o = 2$) are as described in (18) and (20), respectively. Figure 3 shows a comparison of analytical and simulated lower bounds on the sum rate in such a system, with $M = 2$ antennas, $K = 15$ users, and SNR = 10 dB. The values for α , β , and γ used are those of Metric 3, given in (14). Each user has a simple codebook designed as described in the previous section with $B = 1$ bit,

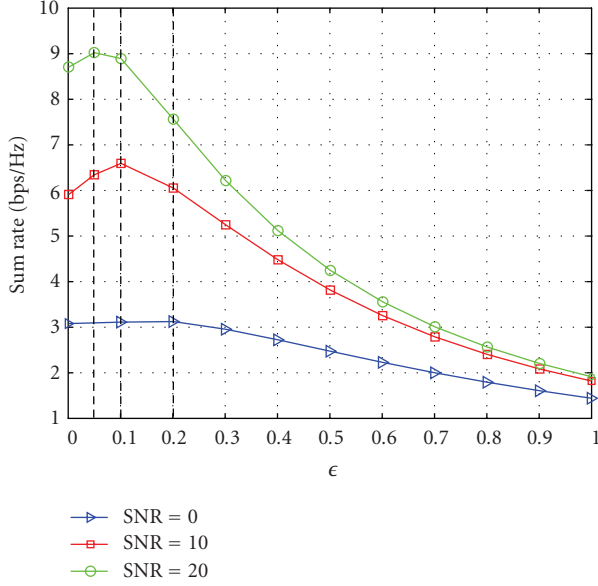


FIGURE 4: Simulated lower bound on the sum rate using Metric 3 as a function of the orthogonality factor ϵ for large K .

different from user to user. Note that the jitter in the analytical curve is due to the rounding effect of K_i .

6. STUDY OF EXTREME REGIMES

In this section, we analyze several extreme regimes, namely, scenarios with large number of users, high SNR, and low SNR regime. The results intuitively clarify the cases in which SDMA is better than TDMA and the role of ϵ in the comparison of both techniques. Previous works in the literature focus on the study of the asymptotic scaling with P or K by using results from extreme value theory, as shown in [10, 13]. Here, we base our study on simpler mathematical tools. The ratios between the sum rates provided by SDMA and TDMA are computed in different limiting cases, by using the sum-rate functions derived in the previous section.

6.1. Large number of users

In this subsection, we provide asymptotical results showing that SDMA can provide higher rates than TDMA in near-orthogonal MIMO systems as the number of users increases, which is consistent with the work presented in [25]. First, note that the number of available users at the i th step can be bounded as $K_i \geq K\epsilon^{2(M-1)}$ as shown in [5]. For finite SNR, we can easily obtain from (18) and (20) the following result.

Theorem 2. *Given an arbitrary ϵ , SDMA outperforms TDMA asymptotically with the number of users*

$$\lim_{K \rightarrow \infty} \frac{R_{M_o}}{R_1} = M_o. \quad (24)$$

Proof. As shown in Figure 3, it can be seen from (18) that R_{M_o} , as function of ϵ , is lower bounded by $R_{M_o}|_{\epsilon=1}$. Thus, here we focus on a lower bound on the SINR, as described

by Metric 3, in order to provide a lower bound on the actual sum rate. The value $\epsilon = 1$ results in a pessimistic SINR lower bound in the metric given in (9). Setting $\epsilon = 1$, we obtain that in each selection step $K_i = K - i + 1$, $i = 1, \dots, M_o$, and thus

$$R_{M_o} \geq \sum_{i=1}^{M_o} \log_2 \left[1 + \frac{1}{\alpha} \sum_{n=1}^{K-i+1} \mathcal{B}_n \mathcal{K}_{1,n} \mathcal{P}_n \right], \quad (25)$$

where $\mathcal{P}_n = 1 + (C n / \alpha) e^{Cn/\alpha} E_i(-C n / \alpha)$, $C = C|_{\epsilon=1}$, and $\alpha = \alpha|_{\epsilon=1}$. Therefore, we get the following lower bound on the ratio between R_{M_o} and R_1 :

$$\begin{aligned} & \lim_{K \rightarrow \infty} \frac{R_{M_o}}{R_1} \\ & \geq \lim_{K \rightarrow \infty} \frac{R_{M_o}|_{\epsilon=1}}{R_1} \\ & \stackrel{(a)}{=} \lim_{K \rightarrow \infty} \frac{\sum_{i=1}^{M_o} \log_2 \left[\left(\frac{K-i+1}{(K-i+1)/2} \right) (1/\alpha) \mathcal{B}_{(K-i+1)/2} \mathcal{P}_{(K-i+1)/2} \right]}{\log_2 \left[\left(\frac{K}{K/2} \right) \mathcal{B}_{K/2} (P/K/2) \right]} \\ & \stackrel{(b)}{=} \lim_{K \rightarrow \infty} \frac{\sum_{i=1}^{M_o} \log_2 \left(\frac{K-i+1}{(K-i+1)/2} \right)}{\log_2 \left(\frac{K}{K/2} \right)} \stackrel{(c)}{=} M_o, \end{aligned} \quad (26)$$

where (a) follows from selecting the highest exponent terms of K in the numerator and denominator and (b) from applying the logarithm property $\log(xy) = \log(x) + \log(y)$, keeping the relevant terms for the computation of the limit; (c) follows by realizing that $\lim_{K \rightarrow \infty} (\log_2(\frac{K-a}{(K-a)/2}) / \log_2(\frac{K}{K/2})) = 1$ for any finite integer a .

Similar to the lower bound obtained on R_{M_o}/R_1 , it can be shown that $\lim_{K \rightarrow \infty} (R_{M_o}/R_1) \leq M_o$ by assuming an upper bound on the SINR as metric with $1 \leq M_o \leq M$, which corresponds to the case of using Metric 4. Setting $K_i = K - i + 1$, $i = 1, \dots, M_o$, and using the sum-rate function for the particular case of $\alpha = 0$, given in (22), yields the desired result. \square

6.2. High SNR regime

This scenario corresponds to the interference-limited region, in which the multiuser interference limits the system performance rather than the average SNR. The number of users K is considered to be finite in the analysis of this regime.

Theorem 3. *Given an arbitrary ϵ , TDMA outperforms SDMA in the high SNR regime*

$$\lim_{P \rightarrow \infty} \frac{R_{M_o}}{R_1} = 0. \quad (27)$$

Proof. The bounded behavior of SDMA as function of the power P is intuitively reflected in the proposed rate function. It suffices to realize that the power dependent part of R_{M_o} can be upper bounded as follows:

$$\mathcal{P}_n \leq 1. \quad (28)$$

In order to provide a proof for the theorem, we focus here on Metric 4, which yields an upper bound on the SDMA sum rate with variable number of active beams. Since in this case we have that $\alpha = 0$, the sum rate is described by (22). The power dependent part is bounded by the following constant:

$$\lim_{P \rightarrow \infty} \frac{1}{C} = \lim_{P \rightarrow \infty} \frac{P}{M_o + (M-1)\beta P} = \frac{1}{(M-1)\beta}. \quad (29)$$

Hence, when transmitting $M_o > 1$ active beams, the sum rate is bounded regardless of the transmitted power. Thus we have that

$$\lim_{P \rightarrow \infty} \frac{R_{M_o}}{R_1} \leq \lim_{P \rightarrow \infty} \frac{\sum_{i=1}^{M_o} \log_2 \left[1 + \sum_{n=1}^{K_i} \mathcal{B}_n \mathcal{K}_{i,n} (1/Cn) \right]}{\log_2 \left[1 + \sum_{n=1}^K \mathcal{B}_n \mathcal{K}_{1,n} (P/n) \right]} = 0, \quad (30)$$

where the inequality follows from the fact that an upper bound on the SDMA sum rate is used, based on Metric 4 with $\alpha = 0$. The equality comes from the fact that when taking the limit, the numerator is not a function of P as shown in (29). Since both R_{M_o} and R_1 are greater than or equal to zero, we obtain the desired result. \square

Note that the above result is consistent with the work in [9], in which the interference-limited behavior of MIMO broadcast channels is studied in a system where limited feedback is available in the form of channel direction information.

6.3. Low SNR regime

This scenario corresponds to the noise-limited region. In this regime, the choice of ϵ has an impact on the optimal choice of transmission technique, that is, SDMA or TDMA. In Figure 4 we show the evolution of the optimal value of ϵ for varying SNR in a cell with large number of users, $K = 1000$, $M = 2$ antennas and a codebook of $B = 1$ bit. The simulated system adapts the optimal number of active beams as a function of ϵ so that the lower bound on the sum rate computed on the basis of Metric 3. Fixing $\epsilon = 0$ implies that the system forces a TDMA solution since there is zero probability of finding two quantized random channels perfectly orthogonal, assuming different quantization codebooks for each user. A shift to the right in the position of the maximum implies that the number of ϵ -orthogonal users found at the second step (K_2) also increases, hence using 2 beams for transmission and thus exploiting the benefits of SDMA rather than TDMA. Therefore, Figure 4 shows that as the SNR decreases, a system based on near-orthogonal transmission tends to select SDMA over TDMA.

However, if the system parameter ϵ is set independently of the average SNR value (or equivalently the power P for normalized noise power), we obtain the following theorem for finite number of users.

Theorem 4. *Given an arbitrary ϵ , set independently of SNR, TDMA provides the same or better performance than SDMA in the low SNR regime:*

$$\lim_{P \rightarrow 0} \frac{R_{M_o}}{R_1} \leq 1. \quad (31)$$

Proof. In order to proof the theorem, we first proof the following asymptotic relation between SDMA and TDMA in 2 extreme cases:

$$0 \leq \lim_{P \rightarrow 0} \frac{R_{M_o}}{R_1} \leq \frac{1}{M_o} \quad \text{if } \epsilon = 0, \quad (32)$$

$$0 \leq \lim_{P \rightarrow 0} \frac{R_{M_o}}{R_1} \leq 1 \quad \text{if } \epsilon = 1. \quad (33)$$

First, we note that the relation $\lim_{P \rightarrow 0} (R_{M_o}/R_1) \geq 0$ follows from the fact that both R_{M_o} and R_1 are greater than zero for positive P . In order to proof the upper bound on $\lim_{P \rightarrow 0} (R_{M_o}/R_1)$ for $\epsilon = 0, 1$, we consider an upper bound on the sum rate, provided by using Metric 4. Since in this case $\alpha = 0$, we use the sum-rate function given in (22). We obtain the following result:

$$\begin{aligned} & \lim_{P \rightarrow 0} \frac{R_{M_o}}{R_1} \\ & \leq \lim_{P \rightarrow 0} \frac{\sum_{i=1}^{M_o} \log_2 \left[1 + \sum_{n=1}^{K_i} \mathcal{B}_n \mathcal{K}_{i,n} (1/Cn) \right]}{\log_2 \left[1 + \sum_{n=1}^K \mathcal{B}_n \mathcal{K}_{1,n} (P/n) \right]} \\ & \stackrel{(a)}{=} \lim_{P \rightarrow 0} \left(\frac{\sum_{i=1}^{M_o} \sum_{n=1}^{K_i} (\mathcal{B}_n \mathcal{K}_{i,n}/n) (1/C)' }{1 + \sum_{n=1}^{K_i} \mathcal{B}_n \mathcal{K}_{i,n} (1/Cn)} \right) \left(\frac{1 + \sum_{n=1}^K \mathcal{B}_n \mathcal{K}_{1,n} (P/n)}{\sum_{n=1}^K (\mathcal{B}_n \mathcal{K}_{1,n}/n)} \right) \\ & \stackrel{(b)}{=} \frac{1}{M_o} \frac{\sum_{i=1}^{M_o} \sum_{n=1}^{K_i} (\mathcal{B}_n \mathcal{K}_{i,n}/n)}{\sum_{n=1}^K (\mathcal{B}_n \mathcal{K}_{1,n}/n)}, \end{aligned} \quad (34)$$

where (a) follows from applying L'Hôpital's rule, with $(1/C)' = \partial(1/C)/\partial P = M_o/[M_o + (M-1)\beta P]^2$, and (b) follows from $\lim_{P \rightarrow 0} (1/C)' = 1/M_o$. For the case $\epsilon = 0$, we have that $K_1 = K$, and $K_i = 0$ for $i \geq 2$. Hence, it can be seen from (34) that the ratio becomes $1/M_o$, thus yielding (32). For the case $\epsilon = 1$, we get $K_i = K - i + 1$, $i = 1, \dots, M_o$. For simplicity, we provide a looser upper bound by considering $K_i = K - i + 1$, $i = 1, \dots, M_o$, which yields the result described in (33). Since intermediate values of ϵ independent of the SNR will yield values for (34) in the range $(1/M_o, 1)$, we obtain the desired result. \square

7. NUMERICAL RESULTS

Figure 5 shows a performance comparison in terms of sum rate versus orthogonality factor ϵ for various levels of channel state information at the transmitter (CSIT). The simulated system has $M = 2$ antennas and a simple codebook of $B = 1$ bits. The number of active users is $K = 10$ and the average SNR = 20 dB. The upper curve corresponds to the sum rate obtained with transmit matched filtering, with perfect CSIT and exhaustive search. Hence, its average rate is not a function of the orthogonality factor. The lower curve corresponds to the sum rate that the system can guarantee when the CSIT consists of quantized channel directions and Metric 3 as scalar feedback (equivalent to Metric 1 for $M = 2$). Thus, this curve corresponds to a lower bound on the actual sum rate that the system can achieve. Finally, the third curve corresponds to the sum rate of a system with

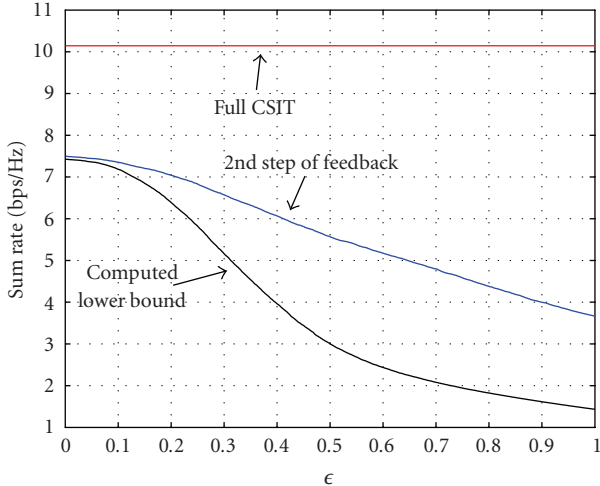


FIGURE 5: Comparison of simulated lower bound on the sum rate using Metric 3, and actual sum rates obtained with second step of feedback and full CSIT. $M = 2$ antennas, $K = 10$ users, SNR = 20 dB, and $B = 1$ bit.

second step of full CSIT feedback, which means that given a set of users selected for transmission by using Metric 3, the BS requests full channel information from those users to perform transmit matched filtering. We can see that the bound becomes looser as ϵ increases, since the bound on the SINR becomes more pessimistic. In the simulated system with $K = 10$ users, the maximum average sum rate occurs when the system sets orthogonality $\epsilon = 0$. This means that the system forces that at each time slot only one beam will be active, since there is zero probability of finding two quantized random channels perfectly orthogonal, assuming different quantization codebooks for each user. Thus, in the simulated scenario with reduced number of users, TDMA (one active beam per time slot) is the optimal transmission technique while in systems with large number of users SDMA is optimal as shown in previous section.

In the remainder of this section, we compare the actual sum rate achieved by systems based on different scalar feedback: Metrics 1, 2, 3, and 4, for $M = 3$ antennas and $B = 9$ bits. For comparison, the performances of random beamforming (RBF) [10] and TxMF with perfect CSIT and exhaustive-search user selection are provided. The systems using Metrics 1, 2, and 4 are assumed to appropriately set M_o and ϵ both for transmission and metric computation, maximizing the sum rate for each K and SNR pair. On the other hand, the scheme with Metric 2 uses optimal ϵ values in each scenario.

Figure 6 shows a performance comparison in terms of sum rate versus number of users for SNR = 10 dB, in a cell with realistic number of active users. The scheme based on Metric 1 provides slightly better performance than the other schemes. The scheme based on Metric 3 exhibits worse scaling with the number of users, thus exploiting less effectively the multiuser diversity. Note that all schemes exhibit slightly worse scaling than RBF and the perfect CSIT solution. This is due to the fact that a simple transmission technique has been

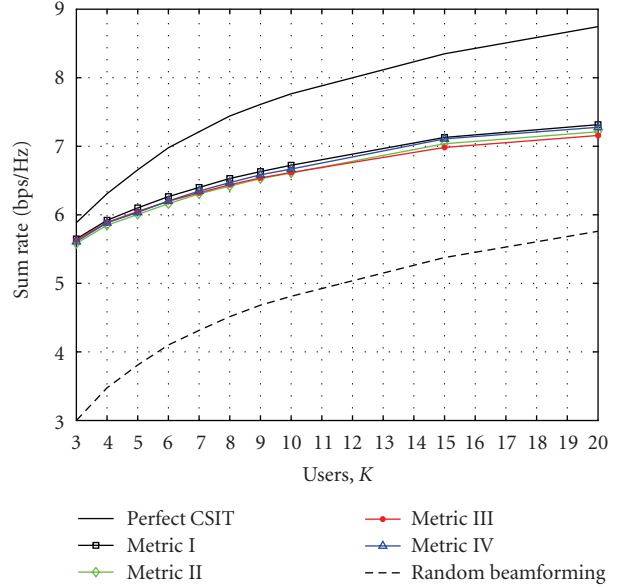


FIGURE 6: Sum rate achieved by different feedback approaches as a function of the number of users, for $B = 9$ bits, $M = 3$ transmit antennas, and SNR = 10 dB.

used, TxMF, since beamforming design is beyond the scope of this paper. In order to restore the optimal scaling with K , zero-forcing beamforming (ZFBF) can be performed at the transmitter based on the available channel quantizations, as discussed in [13].

Figure 7 depicts the performances of different schemes in the low-mid SNR region, in a setting with $K = 10$ users. As the average SNR in the system increases, the sum rate of schemes using Metrics 1 and 3 for feedback converges to the same value. They exhibit linear increase in the high SNR region as expected, which corresponds to a TDMA solution. The scheme that uses Metric 4 for scheduling also benefits from a variable number of active beams, although providing worse performance than the systems using Metrics 1 and 3. Since in the simulated system the number of codebook bits B is not increased proportionally to the average SNR, as discussed in [9], the scheme using Metric 2 ($M_o = M$) exhibits an interference-limited behavior, flattening out at high SNR.

8. CONCLUSIONS

A design framework for scalar feedback in MIMO broadcast channels with limited feedback has been presented. In order to perform user scheduling, these metrics may contain information such as channel power gain, quantization error, orthogonality factor between beamforming vectors, and/or number of active beams. An approximation on the sum rate has been provided for the proposed family of metrics, which has been validated through simulations. As it has been shown, the proposed sum-rate function is a powerful design tool and enables simple analysis. A sum-rate comparison between SDMA and TDMA has been provided in several extreme regimes. Particularly, SDMA outperforms TDMA as the number of users becomes large. TDMA provides better

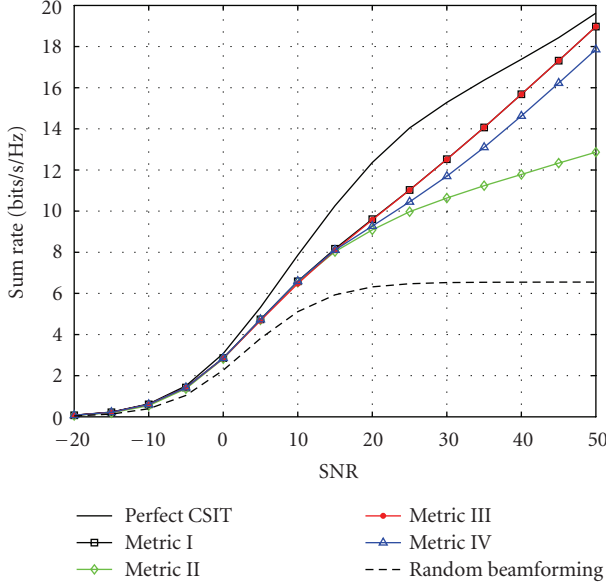


FIGURE 7: Sum rate achieved by different feedback approaches versus average SNR, for $B = 9$ bits, $M = 3$ transmit antennas, and $K = 10$ users.

rates than SDMA in the high SNR regime (interference-limited region). Moreover, the importance of optimizing the orthogonality factor ϵ in the low SNR regime has been highlighted. Several metrics have been presented based on the proposed design framework, illustrating their performances through numerical simulations. The system sum rate can be drastically improved by considering a variable number of active beams adapted to each scenario. In addition, scalar metrics based on SINR lower bounds can provide benefits from a point of view of QoS and feedback reduction.

APPENDICES

A. PROOF OF PROPOSITION 1

Define the following changes of variables:

$$\begin{aligned} \psi &:= \sin^2 \theta_k, & x &:= \frac{1}{\delta} \phi(1 - \psi), \\ \phi &:= \|\mathbf{h}_k\|^2, & y &:= \frac{1}{\delta} \phi \psi. \end{aligned} \quad (\text{A.1})$$

Then, the metric in (10) can be expressed as

$$\xi = \frac{x}{\alpha x + \beta y + 2\gamma\sqrt{xy} + \lambda}, \quad (\text{A.2})$$

where $\lambda = \delta M_o/P$. Note that $\xi \leq 1/\alpha$, with equality for $P \rightarrow \infty$. The Jacobian of the transformation $x = f(\phi, \psi)$, $y = g(\phi, \psi)$ described in (A.1) is given by

$$J(\phi, \psi) = \begin{vmatrix} \frac{\partial x}{\partial \phi} & \frac{\partial x}{\partial \psi} \\ \frac{\partial y}{\partial \phi} & \frac{\partial y}{\partial \psi} \end{vmatrix} = \frac{\phi}{\delta^2}. \quad (\text{A.3})$$

Expressing ϕ and ψ as a function of x and y , we have $\phi = \delta(x + y)$ and $\psi = y/(x + y)$. Substituting in the Jacobian, we get $J(x, y) = (x + y)/\delta$. Since ϕ and ψ are independent random variables for i.i.d. channels, the joint probability density function (pdf) of x and y is obtained from $f_{xy}(x, y) = (1/J(x, y))f_\phi[\delta(x + y)]f_\psi[y/(x + y)]$. The pdf of ϕ is

$$f_\phi(\phi) = \frac{\phi^{M-1}}{\Gamma(M)} e^{-\phi}, \quad (\text{A.4})$$

where $\Gamma(M) = (M-1)!$ is the complete gamma function. The pdf f_ψ is obtained from the cdf of ψ given in (6). Hence, we get the joint density

$$f_{xy}(x, y) = \frac{\delta}{\Gamma(M-1)} e^{-\delta(x+y)} y^{M-2}. \quad (\text{A.5})$$

The cdf of the proposed SINR metric is found by solving the integral

$$F_\xi(s) = \iint_{x,y \in D_s} f_{xy}(x, y) dx dy. \quad (\text{A.6})$$

The bounded region D_s in the xy -plane represents the region where the inequality $x/(\alpha x + \beta y + 2\gamma\sqrt{xy} + \lambda) \leq s$ holds. Isolating x on the left side of the inequality, D_s can be equivalently described as $x \leq g(y)$, with $g(y)$ given by

$$\begin{aligned} g(y) &= \frac{(2\gamma^2 s^2 + \beta s(1 - \alpha s))y + 2\gamma s \sqrt{(y^2 s^2 + \beta s(1 - \alpha s))y^2 + \lambda s(1 - \alpha s)}}{(1 - \alpha s)^2} \\ &\quad + \varphi(s), \end{aligned} \quad (\text{A.7})$$

where $\varphi(s) = \lambda s/(1 - \alpha s)$. Since using $g(y)$ in the integration limits yields difficult integrals, we use the following linear approximation:

$$g(y) \approx m(s)y + \varphi(s), \quad (\text{A.8})$$

where the slope $m(s)$ corresponds to the oblique asymptote of $g(y)$:

$$m(s) = \lim_{y \rightarrow \infty} \frac{\partial g(y)}{\partial y} = \frac{2\gamma s (\gamma s + \sqrt{y^2 s^2 + \beta s(1 - \alpha s)}) + \beta s(1 - \alpha s)}{(1 - \alpha s)^2}. \quad (\text{A.9})$$

Note that, since $0 \leq s \leq 1/\alpha$, then $m(s) \geq 0$ for all s . In addition, since the domain of ψ is $D_\psi = [0, \delta]$, we also obtain the inequalities $y/(x + y) \geq 0$, $y/(x + y) \leq \delta$, and thus $x \geq ((1 - \delta)/\delta)y$. Hence, $F_\xi(s)$ is obtained by integrating $f_{xy}(x, y)$ over the first quadrant of the xy -plane, in the region defined by $x \leq g(y)$ and $x \geq ((1 - \delta)/\delta)y$. Depending on the slopes of these linear boundaries, the integral in (A.6) is carried out over different regions

$$F_\xi(s) \approx \begin{cases} \int_0^\infty \int_{((1-\delta)/\delta)y}^{m(s)y + \varphi(s)} f_{xy}(x, y) dx dy, & m \geq \frac{1 - \delta}{\delta}, \\ \int_0^{y_c} \int_{((1-\delta)/\delta)y}^{m(s)y + \varphi(s)} f_{xy}(x, y) dx dy & 0 \leq m < \frac{1 - \delta}{\delta}. \end{cases} \quad (\text{A.10})$$

The upper integration limit y_c along the y axis in the region $0 \leq m < (1 - \delta)/\delta$ corresponds to the value of y in which the linear boundaries intersect

$$y_c = \frac{\lambda s(1 - \alpha s)\delta}{(1 - \alpha s)^2(1 - \delta) - \beta s(1 - \alpha s)\delta - 2\gamma s(\gamma s + \sqrt{\beta s(1 - \alpha s) + \gamma^2 s^2})\delta}. \quad (\text{A.11})$$

Expressing the regions of the domain of $F_\xi(s)$ as function of s_c , defined as the crossing point between $m(s)$ and $(1 - \delta)/\delta$, and substituting (A.5) into (A.10), the cdf of ξ is found from the following integrals:

$$F_\xi(s) \approx \begin{cases} \frac{\delta}{\Gamma(M-1)} \int_0^\infty e^{-\delta y} y^{M-2} \int_{((1-\delta)/\delta)y}^{my+\varphi} e^{-\delta x} dx dy, & s_c \leq s < 1/\alpha, \\ \frac{\delta}{\Gamma(M-1)} \int_0^{y_c} e^{-\delta y} y^{M-2} \int_{((1-\delta)/\delta)y}^{my+\varphi} e^{-\delta x} dx dy, & 0 \leq s < s_c, \end{cases} \quad (\text{A.12})$$

where s_c is given by

$$s_c = \frac{\alpha(1 - \delta)^2 + \beta(1 - \delta)\delta - 2\sqrt{\gamma^2(1 - \delta)^3\delta}}{\alpha^2(1 - \delta)^2 + 2\alpha\beta(1 - \delta)\delta + \delta(\beta^2\delta - 4\gamma^2(1 - \delta))}. \quad (\text{A.13})$$

Solving the integrals in (A.12), the resulting cdf becomes

$$F_\xi(x) = \begin{cases} 1 - \frac{e^{-M_o s/P(1-\alpha s)}}{\delta^{M-1}(1+m)^{M-1}}, & s_c \leq s < 1/\alpha, \\ 1 - \frac{e^{-M_o s/P(1-\alpha s)}}{\delta^{M-1}(1+m)^{M-1}} + \Phi(s), & 0 \leq s < s_c, \end{cases} \quad (\text{A.14})$$

where $\Phi(s) = 1/\Gamma(M-1)[(e^{-M_o s/P(1-\alpha s)}/\delta^{M-1}(1+m)^{M-1})\Gamma(M-1, \delta(s+1)y_c) - \Gamma(M-1, y_c)]$ and $\Gamma(a, x) = \int_x^\infty t^{a-1} e^{-t} dt$ is the (upper) incomplete gamma function.

Note that this is a generalization of previous results in the literature. In the particular case of $B = 0$, then $\delta = 1$ and thus s_c becomes 0, yielding the cdf derived in [10] for random beamforming. If the metric refers to an upper bound on the SINR, with $\epsilon = 0$, then $s_c = (1 - \delta)/\delta$. If in addition $M_o = M$ is considered as in Metric 2, the cdf of (A.14) becomes the one provided in [13].

In order to obtain a tractable expression for $F_\xi(s)$, we assume that s_c is small so that $F_\xi(s)$ can be approximated as described in (14). Note that a small s_c value corresponds to a low value of B and thus the obtained cdf approximates better the low resolution regime.

B. PROOF OF THEOREM 1

Given M_o beams active for transmission, using (17) we approximate the rate as

$$\text{SR} \approx \sum_{i=1}^{M_o} \log_2 [1 + \mathbb{E}(s_i)]. \quad (\text{B.15})$$

From (15), $\mathbb{E}(s_i)$ is computed as follows:

$$\mathbb{E}(s_i) = \int_0^{1/\alpha} 1 - \left[1 - \frac{e^{-M_o s/P(1-\alpha s)}}{\delta^{M-1}(1+m)^{M-1}} \right]^{K_i} ds. \quad (\text{B.16})$$

Expanding the binomial in the integral, we get

$$\mathbb{E}(s_i) = \sum_{i=1}^{M_o} \frac{(-1)^{n-1}}{\delta^{n(M-1)}} \binom{K_i}{n} \int_0^{1/\alpha} \left[\frac{e^{-M_o s/P(1-\alpha s)}}{\delta^{M-1}(1+m)^{M-1}} \right]^n ds. \quad (\text{B.17})$$

A closed-form solution for the integral in the above equation cannot be found, and thus we use the Bernoulli inequality to obtain an approximation

$$\int_0^{1/\alpha} \left[\frac{e^{-M_o s/P(1-\alpha s)}}{\delta^{M-1}(1+m)^{M-1}} \right]^n ds \geq \int_0^{1/\alpha} e^{[-M_o s/P(1-\alpha s) + (M-1)m]n} ds. \quad (\text{B.18})$$

Note that the integral above is also difficult to solve, since m is a nonlinear function of s , as shown in Theorem 1. In order to provide good sum rates, ϵ will take in general small values. Under this assumption, the following approximation can be made:

$$m \approx \frac{\beta s}{1 - \alpha s}. \quad (\text{B.19})$$

Let $C = M_o/P + (M-1)\beta$, then the integral in (B.17) is approximated by the following integral:

$$\int_0^{1/\alpha} e^{-Cns/(1-\alpha s)} ds = \frac{1}{\alpha} \left[1 + \frac{Cn}{\alpha} e^{Cn/\alpha} E_i\left(-\frac{Cn}{\alpha}\right) \right], \quad (\text{B.20})$$

where $E_i(x)$ is the exponential integral function, defined as $E_i(x) = -\int_{-x}^\infty (e^{-t}/t) dt$. By substituting the approximated value of the integral found above into (B.17), and using the definitions of \mathcal{B}_n , $\mathcal{K}_{i,n}$, and \mathcal{P}_n given in Theorem 2, we obtain the desired approximation for the sum rate.

REFERENCES

- [1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.
- [2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.
- [3] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian MIMO broadcast channel," in *Proceedings of the International Symposium on Information Theory (ISIT '04)*, p. 174, Chicago, Ill, USA, June-July 2004.

- [4] M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, 1983.
- [5] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 528–541, 2006.
- [6] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1423–1436, 1998.
- [7] 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.
- [8] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple-antenna systems," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2562–2579, 2003.
- [9] N. Jindal, "MIMO broadcast channels with finite rate feedback," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '05)*, vol. 3, pp. 1520–1524, St. Louis, Mo, USA, November–December 2005.
- [10] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channels with partial side information," *IEEE Transactions on Information Theory*, vol. 51, no. 2, pp. 506–522, 2005.
- [11] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1277–1294, 2002.
- [12] N. Jindal, "Finite rate feedback MIMO broadcast channels," in *Proceedings of the Workshop on Information Theory and Its Applications (ITA '06)*, San Diego, Calif, USA, February 2006, invited paper.
- [13] T. Yoo, N. Jindal, and A. Goldsmith, "Finite-rate feedback MIMO broadcast channels with a large number of users," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '06)*, pp. 1214–1218, Seattle, Wash, USA, July 2006.
- [14] M. Kountouris, R. de Francisco, D. Gesbert, D. T. M. Slock, and T. Sälzer, "Efficient metric for scheduling in MIMO broadcast channels with limited feedback," FT060303—France Telecom R&D internal report, March 2006.
- [15] R. de Francisco, D. T. M. Slock, and Y.-C. Liang, "Balance of multiuser diversity and multiplexing gain in near-orthogonal MIMO systems with limited feedback," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '07)*, pp. 1269–1274, Kowloon, China, March 2007.
- [16] R. de Francisco and D. T. M. Slock, "On the design of scalar feedback techniques for MIMO broadcast scheduling," in *Proceedings of the 8th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC '07)*, Helsinki, Finland, June 2007.
- [17] M. Kountouris, R. de Francisco, D. Gesbert, D. T. M. Slock, and T. Sälzer, "Efficient metrics for scheduling in MIMO broadcast channels with limited feedback," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '07)*, vol. 3, pp. 109–112, Honolulu, Hawaii, USA, April 2007.
- [18] M. Kountouris, R. de Francisco, D. Gesbert, D. T. M. Slock, and T. Sälzer, "Low complexity scheduling and beamforming for multiuser MIMO systems," in *Proceedings of the 7th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC '06)*, pp. 1–5, Cannes, France, July 2006.
- [19] S. Zhou, Z. Wang, and G. B. Giannakis, "Quantifying the power loss when transmit beamforming relies on finite-rate feedback," *IEEE Transactions on Wireless Communications*, vol. 4, no. 4, pp. 1948–1957, 2005.
- [20] G. Dimić and N. D. Sidiropoulos, "On downlink beamforming with greedy user selection: performance analysis and a simple new algorithm," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3857–3868, 2005.
- [21] T. Yoo and A. Goldsmith, "Sum-rate optimal multi-antenna downlink beamforming strategy based on clique search," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '05)*, vol. 3, pp. 1510–1514, St. Louis, Mo, USA, November–December 2005.
- [22] J. C. Roh and B. D. Rao, "Transmit beamforming in multiple-antenna systems with finite rate feedback: a VQ-based approach," *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 1101–1112, 2006.
- [23] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, Boston, Mass, USA, 4th edition, 2002.
- [24] H. A. David and H. N. Nagaraja, *Order Statistics*, John Wiley & Sons, New York, NY, USA, 3rd edition, 2003.
- [25] M. Sharif and B. Hassibi, "A comparison of time-sharing, DPC, and beamforming for MIMO broadcast channels with many users," *IEEE Transactions on Communications*, vol. 55, no. 1, pp. 11–15, 2007.