# Multiple Antenna Broadcast Channels with Shape Feedback and Limited Feedback

Peilu Ding, David J. Love, and Michael D. Zoltowski
School of Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana 47907
(Ph) 765-496-6797 (Fax) 765-494-0880
Email:{pding, djlove, mikedz}@ecn.purdue.edu

**Abstract**

In this paper, we consider two different models of partial channel state information (CSI) at the basestation for multiple antenna broadcast channels: i.) the shape feedback model where the normalized channel vector of each user is available at the basestation and ii.) the limited feedback model where each user quantizes its channel vector according to a rotated codebook which is optimal in the sense of mean square error and feeds back the codeword index. The paper is focused on characterizing the sum rate performance of both zero-forcing dirty paper coding (ZFDPC) systems and channel inversion (CI) systems under the given two partial basestion CSI models. Intuitively speaking, a system with shape feedback loses the sum rate gain of adaptive power allocation. However, shape feedback still provides enough channel knowledge for ZFDPC and CI to approach their own optimal throughput in the high SNR regime. As for limited feedback, we derive sum rate bounds for both signaling schemes and link their throughput performance to some basic properties of the quantization codebook. Interestingly, we find that limited feedback employing a fixed codebook leads to a sum rate ceiling for both schemes for asymptotically high SNR.

**Index Terms**

MIMO, Broadcast channel, Sum rate, Partial CSI, Shape feedback, Limited feedback.

Supported by the SBC Foundation and the NSF under grant CCF0513916.

## I. Introduction

Because of the need for high data rate multi-user systems, it is imperative to understand how to leverage multiple antenna technology to increase the data rate and user capacity of wireless networks. Broadcast channels model the situation where a basestation is sending information to a number of users (receivers) [1]. Simple examples of broadcast channels include the downlink in a cellular network and the broadcast scenario in a wireless local area network (LAN) where the access point is transmitting to multiple users.

In recent years, broadcast channels with multiple antennas installed at the basestation have received significant research interest because of their spectral efficiency improvement and potential for commercial application in wireless systems [2]. It was shown in [3]–[6] that the multiple antennas at the basestation provide a sum rate capacity increase that grows linearly with the minimum of the number of transmit antennas and users. The resulting sum rate advantage can be achieved through dirty paper coding [7] which eliminates the cross user interference at the transmitter side assuming perfect channel state information (CSI) is available at the basestation. Besides these information theoretic results, there has also been some work recently in the area of practical signaling for the multiple antenna broadcast channel. For example, [8] studied zero-forcing beamforming methods for the downlink of multi-user multiple-input multiple-output (MIMO) channels. Peel *et al*. proposed the regularized channel inversion (CI) scheme and combined this technique with vector perturbation [9], [10]. Another interesting research topic for the multiple antenna broadcast channel is multi-user scheduling (selection). Recent progress on this topic showed that by judiciously selecting the active user set, the optimal throughput scaling can be achieved by even suboptimal signaling schemes, such as CI, when the number of users is very large [11]–[13].

For multiple antenna broadcast channels, when the basestation does not have any channel knowledge, the sum rate loss compared to the case that the basestation has perfect CSI is substantial. In fact, the optimal transmission scheme for the basestation without any CSI is to transmit to a randomly selected user during each time slot. In this case, the sum rate does not even grow either as the number of antennas increases or as the number of users increases (no multi-user diversity gain). While the perfect CSI assumption for the basestation can be argued in the case of a time division duplexing (TDD) system, the assumption is highly unrealistic for frequency division duplexing (FDD). Some recent multi-user scheduling algorithms only require partial CSI at the basestation to maintain full sum rate growth when the number of users is large [11], [14], [15]. However, it is still not clear what kind of partial channel knowledge is essential for multiple antenna broadcast systems to obtain the sum rate advantage.

The benefits of designing a point-to-point multiple antenna signal using some form of partial CSI has received much interest over the past few years. Several different models for partial channel knowledge have been proposed and analyzed, including the statistical partial CSI model [16]–[19], the channel subspace model [20], and the limited feedback model [21]–[30]. However, for multiple antenna broadcast channels, the partial CSI problem is not as well addressed as in the single user MIMO case. Recent progress on this topic can be found in [14], [15], [31]–[34].

In this paper, we consider two different kinds of partial CSI at the basestation for multiple antenna broadcast channels and focus on the setting where the basestation has multiple antennas and each mobile has a single antenna due to size and battery constraints. The first model considered is shape feedback. In the shape feedback model, the basestation is able to obtain the normalized channel vector, i.e., the *shape*[1], of each user. Even though it is simple, the shape feedback model is especially helpful for us to understand the throughput sensitivity on the channel gain knowledge for multiple antenna broadcast channels. In many practical systems, the basestation can have unreliable or even unobtainable channel gain values but reliable channel shape information of each user. For example, when analog feedback [35]–[37] is used in an FDD system, the magnitude value at the basestation is usually outdated much faster than the shape information. This is because the shape vector mainly captures the directional knowledge of paths, which usually varies much more slowly than the amplitude of the channel, especially in the outdoor scenario [38]. Also the channel gain in an analog feedback system might be unusable because the system gains in the feedback channel are often not well calibrated. Furthermore, if blind channel estimation is combined with channel reciprocity at the basestation of a TDD system, the channel gain

[1]Any nonzero vector $\boldsymbol{a} \in \mathbb{R}^n$ can always be decomposed into a gain $\|\boldsymbol{a}\|$ which is a scalar and a shape $\boldsymbol{a}/\|\boldsymbol{a}\|$ which is a vector on the unit sphere of $\mathbb{R}^n$.

knowledge is usually unobtainable due to the amplitude ambiguity nature of the blind methods [39].

The general limited feedback model is also considered. In this model, CSI is conveyed from each user to the basestation over a feedback channel. The basestation and the users have access to a CSI codebook which is designed offline. Each user sends the binary index of the best codevector from the codebook through a zero-error, zero-delay feedback channel to the basestation. We note that this finite rate partial CSI model was independently studied in [40] for the channel inversion scheme. Here a limited feedback framework designed for multiple antenna broadcast systems is first proposed. The key differences between our scheme and the limited feedback for single user MIMO channels are the following:

1) For multiple antenna broadcast channels, each receiver only knows its own channel instead of the full (i.e. all users') CSI and the users cannot cooperate. Each user is unable to obtain the optimal precoding or beamforming structures which are computed from the full CSI. Therefore, in our scheme, vector quantization is applied to the channel vector itself instead of to the beamforming vector or precoding matrices, which is usually the case for single user MIMO systems where the receiver has full CSI [41];
2) The codebook of each user should be different from others. Otherwise, there is a chance that two or more users quantize their channel vectors to the same codevector which will cause a rank loss in the quantized channel matrix composed by those codevectors. To avoid this situation, we let every user rotate a general codebook by a random unitary matrix that is also known at the basestation so that the CSI matrix at the basestation is full rank with probability one. Also under Rayleigh fading, the randomly rotated codebooks used by different users are all equivalent in the sense of average quantization error.

For these two partial CSI models, two widely accepted transmission schemes are considered: i.) the asymptotically optimal zero-forcing dirty paper coding (ZFDPC) scheme and ii.) the channel inversion (CI) method that is suboptimal but more practical. By characterizing the sum rate performance of the two partial CSI models for the given signaling schemes, we find that these two kinds of partial CSI result in quite different throughput performance. It is shown that for the high SNR regime, ZFDPC with shape feedback is asymptotically optimal in the sum rate sense. When the CI scheme is used with shape feedback, we give a simple but effective power allocation strategy that provides sum rate performance better than perfect CSI channel inversion with equal power allocation. Simulation results further show that it is very close to the maximum achievable sum rate for CI which requires perfect CSI and adaptive power allocation.

For the limited feedback case, the mismatch between the quantized channel vectors available at the basestation and the exact channel vectors results in additional cross-user interference for both ZFDPC and channel inversion. We derive bounds for the sum rate performance of limited feedback under these two transmission schemes. These bounds link the throughput performance with some basic properties of the codebook and provide important insights into the impact of the use of limited feedback. From these bounds, we find that both signaling schemes experience sum rate ceilings for a fixed size codebook as the SNR increases.

Throughout this paper, we use $(\cdot)^T$ to denote the transpose, $(\cdot)^H$ the conjugate transpose, $\mathbb{E}\{\cdot\}$ the expectation and $(\cdot)^*$ the complex conjugate. $\mathrm{tr}\,(\boldsymbol{A})$ means the trace of matrix $\boldsymbol{A}$, and $\mathrm{diag}\,(\boldsymbol{a})$ denotes the diagonal matrix whose diagonal line is composed by the elements of vector $\boldsymbol{a}$. $(\boldsymbol{A})_i^T$ represents the $i$th row vector of $\boldsymbol{A}$ .

## II. SYSTEM OVERVIEW

### A. Channel Model

Consider a broadcast channel consisting of an $N$-antenna basestation and $K$ single-antenna users. Assuming that the channel is flat-fading, the discrete-time complex baseband signal received by user $i$ at a given time slot is

$$y_i = \boldsymbol{h}_i^T \boldsymbol{x} + v_i$$

where $\boldsymbol{h}_i^T = [h_{i,1}, \ldots, h_{i,N}]$ is the channel fading vector between the basestation and the $i$th user, $\boldsymbol{x} = [x_1, \ldots, x_N]^T$ is the transmitted signal, and $v_i$ is the zero mean complex white noise with variance one. We assume independent and identically distributed (i.i.d.) Rayleigh fading, $\boldsymbol{h}_i \sim \mathcal{C}N(0, \boldsymbol{I}_N)$ where

$\boldsymbol{I}_N$ is an $N \times N$ identity matrix and $\mathbb{E}\{\boldsymbol{h}_i \boldsymbol{h}_j^H\} = \boldsymbol{0}$ if $i \neq j$. By stacking the received signals of all $K$ users into $\boldsymbol{y} = [y_1, y_2, \ldots, y_K]^T$, we have

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{v} \tag{1}$$

where $\boldsymbol{H} = [\boldsymbol{h}_1, \ldots, \boldsymbol{h}_K]^T$ and $\boldsymbol{v} = [v_1, \ldots, v_K]^T$. The model (1) looks the same as a single user MIMO channel, but the key difference is that the receive antennas cannot cooperate with each other in the broadcast scenario. User $i$ only sees its own received signal $y_i$. As a consequence, user $i$ can only obtain the state information of its own channel which is $\boldsymbol{h}_i$ in this case.

Generally the number of the users $K$ is greater than the number of the transmit antennas $N$, that is[2] $K > N$. As we have mentioned, how to select the appropriate users for transmission under some sort of CSI knowledge is an important subject and deserves its own treatment. We leave this topic for future work and focus on the system setting where $K = N$. This can be understood as the scenario where $N$ users are chosen randomly according to a uniform distribution if $K > N$. The channel matrix $\boldsymbol{H}$ is assumed to be full rank in the following part of the paper (this happens with probability one given i.i.d. Rayleigh fading). We also omit the user ordering issue which does not affect the performance of channel inversion and the asymptotic sum rate for ZFDPC [3].

The system sum rate is defined as

$$R = \sum_{i=1}^{K} R_i$$

where $R_i$ is the transmission rate between the basestation and the $i$th user. It is measured in bits per channel use (or bits/s/Hz) in this paper. We focus on the ergodic sum rate for time-varying channels, which means the sum rate is averaged over all the channel states according to the channel distribution. This ergodic sum rate can be approached by a fixed rate scheme which codes across a long time period during which the channel experiences the ergodic states according the distribution. In this case, the channel fading is fast enough so that the channel state changes ergodically during the transmission of one codeword.

[2]When $K < N$, there is a loss in the transmission degree of freedom [9].

The following power constraint is applied to the transmitted signal

$$\mathbb{E}\left\{\|\boldsymbol{x}\|^2\right\} = P \tag{2}$$

where $\|\cdot\|$ is the Euclidean norm and $P$ is the maximum total transmit power over one time slot. Since the noise power is normalized to one, $P$ also represents the SNR.

### *B. Zero-Forcing Dirty Paper Coding*

The sum rate capacity of multiple antenna broadcast channels is achieved through dirty paper coding at the basestation. The optimal dirty paper coding capacity involves a joint optimization over a set of covariance matrices under a chosen power constraint, which is too complex for implementation. In [3], a relatively simple ZFDPC scheme is proposed and is shown to provide an optimal throughput for asymptotically large SNR.

In the ZFDPC scheme, the basestation collects the perfect channel knowledge $\boldsymbol{H}$ and then decomposes it into

$$\boldsymbol{H} = \boldsymbol{G}\boldsymbol{Q}$$

where $\boldsymbol{G}$ is a $K \times N$ lower triangular matrix and $\boldsymbol{Q}$ is an $N \times N$ unitary matrix under the assumption that $\boldsymbol{H}$ is full rank. Applying $\boldsymbol{Q}^H$ to the original source signal $\boldsymbol{s} = [s_1, \ldots, s_N]^T$ as a precoding matrix gives a transmitted signal

$$\boldsymbol{x} = \boldsymbol{Q}^H \boldsymbol{s}$$

and the input-output relationship for the $i$th user

$$y_i = g_{ii} s_i + \sum_{j<i} g_{ij} s_j + v_i, \tag{3}$$

where $g_{ij}$ is the $(i, j)$ element in $\boldsymbol{G}$. By treating $\sum_{j<i} g_{ij} s_j$ as the known interference and judiciously generating $s_i$ according to dirty paper coding, these $N$ cross interfering subchannels have the same capacity as $N$ parallel Gaussian channels with fading gains $g_{ii}$, $i = 1, \ldots, N$. The resulting sum rate is

$$R^{dpc} = \sum_{i=1}^{N} \mathbb{E}\left\{\log_2\left(1 + |g_{ii}|^2 P_i\right)\right\} \tag{4}$$

where $P_i$ is the power allocated to user $i$ and satisfies $\sum_{i=1}^{N} P_i = P$.

The maximum sum rate is achieved through waterfilling power allocation with a sum rate of

$$R^{dpc*} = \sum_{i=1}^{N} \mathbb{E}\left[\log_2\left(\mu|g_{ii}|^2\right)\right]_+$$

where $\mu$ is the solution of the waterfilling equation $\sum_{i=1}^{N}[\mu - |g_{ii}|^{-2}]_+ = P$. The sum rate under equal power allocation is also of interest to us

$$R^{dpc-eq} = \sum_{i=1}^{N} \mathbb{E}\left\{\log_2\left(1 + |g_{ii}|^2 P/N\right)\right\}.$$

### *C. Channel Inversion Scheme*

Channel inversion, which is also called zero-forcing beamforming, is suboptimal but easy to implement compared to dirty paper coding based schemes. It decouples the channel into orthogonal subchannels with linear precoding. Since we assume $K = N$, the channel matrix $\boldsymbol{H}$ has full rank with probability one under i.i.d. Rayleigh fading. The precoding matrix is the direct inverse of the channel[3] which gives

$$\boldsymbol{x} = \boldsymbol{H}^{-1}\boldsymbol{s}.$$

The effective input-output relation is just a set of $N$ additive white Gaussian noise subchannels without fading

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{v} = \boldsymbol{s} + \boldsymbol{v}.$$

Just as receiver zero forcing leads to noise enhancement, the CI scheme usually causes signal power reduction. Let the average power of the $i$th substream be $\mathbb{E}\{|s_i|^2\} = P_i$, the total power constraint (2) now has the form

$$\sum_{i=1}^{N} P_i \left[(\boldsymbol{H}\boldsymbol{H}^H)^{-1}\right]_{i,i} = P$$

where $[\cdot]_{i,i}$ denotes the $(i,i)$ element in the matrix. The sum rate of CI is the summation of the data rate of each substream

$$R^{ci} = \sum_{i=1}^{N} \mathbb{E}\left\{\log_2\left(1 + P_i\right)\right\}. \quad (5)$$

[3]When $N < K$, the precoding matrix should be the pseudoinverse of the channel, $\boldsymbol{x} = \boldsymbol{H}^H(\boldsymbol{H}\boldsymbol{H}^H)^{-1}\boldsymbol{s}$.

Again, the optimal sum rate is achieved when the power is allocated according to the waterfilling solution yielding

$$R^{ci*} = \sum_{i=1}^{N} \mathbb{E}\left[\log_2\left(\mu / \left[(\boldsymbol{H}\boldsymbol{H}^H)^{-1}\right]_{i,i}\right)\right]_+$$

where $\mu$ is the solution of $\sum_{i=1}^{N}\left[\mu - 1/\left[(\boldsymbol{H}\boldsymbol{H}^H)^{-1}\right]_{i,i}\right]_+ = P$ [3]. When the power is equally allocated, we have

$$P_i = \frac{P}{\sum_{i=1}^{N}\left[(\boldsymbol{H}\boldsymbol{H}^H)^{-1}\right]_{i,i}} = \frac{P}{\|\boldsymbol{H}^{-1}\|_F^2}, \text{ for } i = 1, \ldots, N \tag{6}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. Replacing the $P_i$ in (5) with (6), we get the resulting sum rate

$$R^{ci-eq} = N\mathbb{E}\left\{\log_2\left(1 + P\|\boldsymbol{H}^{-1}\|_F^{-2}\right)\right\}. \tag{7}$$

Both schemes usually assume perfect channel knowledge at the basestation which can be impractical when $K$ and $N$ become large. In the following sections, we will consider partial channel knowledge and its effect on the sum rate performance.

## III. Sum Rate of Shape Feedback

For the multiple antenna broadcast setting described in Section II, every user experiences a multiple-input single-output (MISO) channel $\boldsymbol{h}_i$. If the system has only one user, that is $K = 1$, the transmitter only needs to know the normalized channel vector $\hat{\boldsymbol{h}}_1 = \boldsymbol{h}_1/\|\boldsymbol{h}_1\|$ to achieve a sum rate arbitrarily close to the channel capacity. In other words, the channel throughput is insensitive to the loss of $\|\boldsymbol{h}_1\|$ knowledge at the transmitter.

To study the sum rate sensitivity of the channel magnitude knowledge for the multiuser case, we consider shape feedback. The normalized channel vectors for each user are available at the basestation by reciprocity or user feedback. In this case, the CSI at the basestation is

$$\hat{\boldsymbol{H}} = [\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2, \ldots, \hat{\boldsymbol{h}}_N]^T$$

where $\hat{\boldsymbol{h}}_i = \boldsymbol{h}_i/\|\boldsymbol{h}_i\|$, $i = 1, \ldots, N$, are all unit norm vectors. $\hat{\boldsymbol{H}}$ is linked with the perfect channel knowledge according to

$$\boldsymbol{H} = \boldsymbol{\Lambda}\hat{\boldsymbol{H}}$$

where $\mathbf{\Lambda} = \mathrm{diag}\left([\alpha_1, \ldots, \alpha_K]\right)$ and $\alpha_i = \|\boldsymbol{h}_i\|$ are the amplitudes of the channel vectors. Compared to the perfect channel vectors, channel shape vectors are uniformly distributed on the complex unit sphere in $N$-dimensions under i.i.d. Rayleigh fading [42].

### *A. Shape Feedback with ZFDPC*

The first observation about shape feedback is that the basestation is still able to get the exact precoding matrix for ZFDPC as in the full CSI case. Because the QR-type decomposition of $\hat{\boldsymbol{H}}$ has the form

$$\hat{\boldsymbol{H}} = \hat{\boldsymbol{G}}\boldsymbol{Q} = (\mathbf{\Lambda}^{-1}\boldsymbol{G})\boldsymbol{Q}$$

where $\boldsymbol{G}\boldsymbol{Q}$ is the QR-type decomposition of the perfect channel matrix $\boldsymbol{H}$. After precoding, the transmission relation for the $i$th user is

$$y_i = \alpha_i(\hat{g}_{ii}s_i + \sum_{j<i} \hat{g}_{ij}s_j) + v_i \tag{8}$$

where the signal $\hat{g}_{ii}s_i$ and the known interference $\sum_{j<i} \hat{g}_{ij}s_j$ experience the same multiplicative gain $\alpha_i = \|\boldsymbol{h}_i\|$.

The second observation is that without the channel magnitude knowledge $\alpha_i = \|\boldsymbol{h}_i\|$, the basestation is unable to get the exact inflation factor

$$\beta_i = \alpha_i^2|\hat{g}_{ii}|^2 P_i/(\alpha_i^2|\hat{g}_{ii}|^2 P_i + 1)$$

which is required to optimally implement dirty paper coding [7]. Fortunately, the ZFDPC scheme can still be applied approximately optimally in the high SNR regime because $\beta_i \approx 1$ in this region and the basestation can just fix the inflation factor to be $1$.

The third effect of shape feedback is that there is no meaning to adaptive power allocation for ZFDPC. When the channels of different users are symmetric in distribution, equal power allocation ($P_i = P/N$ for $i = 1, \ldots, N$) is the optimal strategy.

Therefore, in the high SNR regime, shape feedback can be used with ZFDPC under equal power loading and gives the following ergodic sum rate

$$R^{dpc}_{\mathrm{sha},P\gg 0} = \sum_{i=1}^{N} \mathbb{E}\left\{\log_2\left(1 + \frac{|g_{ii}|^2 P}{N}\right)\right\} = R^{dpc-eq},\ \text{for } P \gg 0.$$

Compared to $R^{dpc*}$, $R^{dpc}_{\text{sha},P\gg 0}$ loses the adaptive power loading gain. But it is easy to show that $R^{dpc}_{\text{sha},P\gg 0}$ is still optimal for asymptotically high SNR.

For a Rayleigh fading channel (i.e. the elements of $\boldsymbol{H}$ are i.i.d. distributed as $\mathcal{C}N(0,1)$), each $|g_{ii}|^2$ is independently distributed as $\chi^2_{2(N-i+1)}$ which denotes a central Chi-squared distributed random variable with $2(N-i+1)$ degrees of freedom [3], [43]. Thus the closed-form expression of $R^{dpc}_{sha,P>>0}$ for Rayleigh fading can be easily derived by using the equality in [44] as

$$R^{dpc}_{\text{sha},P\gg 0} \quad = \quad e^{N/P}\log_2 e \sum_{i=0}^{N-1}(N-i)\mathcal{E}_{i+1}\left(\frac{N}{P}\right), \ \text{for } P \gg 0$$

where $\mathcal{E}_n(\cdot)$ is the exponential integral function of order $n$ [45], [46].

*B. Shape Feedback with Channel Inversion*

When shape feedback is combined with channel inversion, we use the inversion of $\hat{\boldsymbol{H}}$ as the precoding matrix. The received signal is

$$\boldsymbol{y} = \boldsymbol{H}\hat{\boldsymbol{H}}^{-1}\boldsymbol{s} + \boldsymbol{v} = \boldsymbol{\Lambda}\boldsymbol{s} + \boldsymbol{v}. \tag{9}$$

Compared to channel inversion with perfect channel knowledge, which results in $N$ additive white Gaussian noise channels, (9) represents $N$ subchannels with different multiplicative channel gains $\alpha_i$. The power constraint changes into

$$\mathbb{E}\left\{\|\boldsymbol{x}\|^2\right\} = \sum_{i=1}^{N} P_i \left[(\hat{\boldsymbol{H}}\hat{\boldsymbol{H}}^H)^{-1}\right]_{i,i} = P \tag{10}$$

It is of interest to see that for the channel inversion scheme, even without channel magnitude knowledge, the basestation can still improve the throughput by adaptively allocating the transmit power. Consider the power allocation scheme

$$P_i = \frac{P}{N[(\hat{\boldsymbol{H}}\hat{\boldsymbol{H}}^H)^{-1}]_{i,i}} \tag{11}$$

which results in the following sum rate

$$\begin{aligned} R^{ci}_{\text{sha}} &= \sum_{i=1}^{N}\mathbb{E}\left\{\log_2\left(1+\frac{\alpha_i^2 P}{N[(\hat{\boldsymbol{H}}\hat{\boldsymbol{H}}^H)^{-1}]_{i,i}}\right)\right\} \\ &= \sum_{i=1}^{N}\mathbb{E}\left\{\log_2\left(1+\frac{P}{N[(\boldsymbol{H}\boldsymbol{H}^H)^{-1}]_{i,i}}\right)\right\}. \end{aligned}$$

where the second equality comes from $(\hat{\boldsymbol{H}}\hat{\boldsymbol{H}}^H)^{-1} = \boldsymbol{\Lambda}^H(\boldsymbol{H}\boldsymbol{H}^H)^{-1}\boldsymbol{\Lambda}$.

Compared to the sum rate performance of perfect CSI with channel inversion, we have the following result.

*Lemma 1:* The sum rate performance of shape feedback with channel inversion and the power allocation defined in (11) is better than the sum rate of perfect CSI channel inversion with equal power allocation, that is $R^{ci}_{\text{sha}} \geq R^{ci-eq}$.

*Proof:* Let $\gamma_i = [(\boldsymbol{H}\boldsymbol{H}^H)^{-1}]_{i,i}$. We have

$$\begin{aligned} R^{ci}_{\text{sha}} &= N\sum_{i=1}^{N}\frac{1}{N}\mathbb{E}\left\{\log_2\left(1+\frac{P}{N\gamma_i}\right)\right\} \\ &\geq N\mathbb{E}\left\{\log_2\left(1+\frac{P}{\sum_{i=1}\gamma_i}\right)\right\} \\ &= R^{ci-eq} \end{aligned}$$

where the inequality comes from Jensen's inequality and the convexity of $\log_2(1+\frac{1}{x})$. ■

In the simulation results presented in Section V, we can see that the performance of $R^{ci}_{\text{sha}}$ is in fact very close to $R^{ci*}$ which is the ergodic maximum achievable sum rate of channel inversion scheme.

*Remarks:* Under shape feedback, the basestation does not know the the supportive rates of the current channel state. The ergodic sum rate can be approached by a fixed rate coding scheme whose codeword duration is long enough such that it experiences the channel states ergodically.

## IV. Sum Rate under Limited Feedback

The shape feedback model discussed in the previous section does not provide much flexibility on the CSI overhead rate for system implementation. In this section, we consider a different kind of partial CSI model, the limited feedback model, for the basestations of multiple antenna broadcast systems. The limited feedback model has been successfully used in single user MIMO systems. Its feedback overhead rate can be adjusted by changing the size of the quantization codebook. We first propose a limited feedback scheme designed for multiple antenna broadcast channels. We derive an upper bound for its sum rate performance under both ZFDPC and CI. The sum rate ceiling effect of limited feedback is also found.

### *A. Limited Feedback for Multiple Antenna Broadcast Channels*

Limited feedback for single user multiple antenna systems has been studied under different settings. In most of the cases, a codebook that is known at both the receiver and the transmitter is used to quantize the channel information necessary to construct the adaptive transmitted signal. The channel information to be quantized will not necessarily be the channel vector/matrix itself. For example, when limited feedback beamforming and optimal receive combining are used, the vector codebook is constructed to quantize the optimal beamforming vector which is the singular vector of $\boldsymbol{H}$ corresponding to the largest singular value [21], [22].

For a multiple antenna broadcast system, the key feature that distinguishes it from a single user MIMO system is that the receive antennas of different users cannot cooperate. At the receiver side the total CSI $\boldsymbol{H}$ is separated into $\{\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_K\}$ and distributed among the $K$ users. We assume no user cooperation, so user $i$ only knows $\boldsymbol{h}_i$ and is not able to obtain the information about the optimal transmission scheme (for example, the precoding matrix $\boldsymbol{Q}^H$ in the ZFDPC scheme) which is based on the full knowledge of $\boldsymbol{H}$. Therefore in our limited feedback scheme, the codebook is constructed for each user to directly quantize the channel vector itself.

Consider a codebook $\mathcal{W}$ that contains $L$ codevectors

$$\mathcal{W} = \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_L\}.$$

We use *minimum distance selection* and *mean square error* as the encoding function and distortion measure, respectively. Therefore, user $i$ encodes its channel vector $\boldsymbol{h}_i$ into

$$\mathcal{Q}_{\mathcal{W}}(\boldsymbol{h}_i) = \boldsymbol{w}_{l_i}$$

where $l_i = \mathrm{argmin}_{1 \le j \le L} \|\boldsymbol{h}_i - \boldsymbol{w}_j\|$. Every user sends its index $l_i$ back to the basestation, so that the channel knowledge at the basestation is

$$\boldsymbol{H}_w = [\boldsymbol{w}_{l_1}, \boldsymbol{w}_{l_2}, \ldots, \boldsymbol{w}_{l_K}]^T.$$

The average distortion introduced by quantization according to codebook $\mathcal{W}$ is defined as

$$D_{\mathcal{W}} = E_{\boldsymbol{h}} \left\{ \| \mathcal{Q}_{\mathcal{W}}(\boldsymbol{h}) - \boldsymbol{h} \|^2 \right\}. \tag{12}$$

A locally optimal codebook in the sense of (12) for a given size $L$ can be constructed by the generalized Lloyd algorithm [47].

Notice that if a general codebook is used by all the users, the limited feedback may result in the ill-conditioning of $\boldsymbol{H}_w$. This could happen when $l_j = l_k$ for $j \neq k$, meaning that two or more users select the same codevector in the codebook. In that case, the channel knowledge at the basestation $\boldsymbol{H}_w$ is not full rank. For multiple antenna broadcast channels, the rank loss in the CSI matrix can be seen as reducing the number of transmit antennas which will cause a large sum rate degradation. To avoid such degradation, we propose to use different codebooks at each user. Let $\mathcal{W}^{(i)} = \{\boldsymbol{w}_1^{(i)}, \ldots, \boldsymbol{w}_L^{(i)}\}$ be the codebook used by user $i$. We want

$$\boldsymbol{w}_l^{(i)} \neq \boldsymbol{w}_m^{(j)} \text{ for } i \neq j; l, m = 1, \ldots, L. \tag{13}$$

We will also require that the codebooks provide the same average distortion, i.e.

$$D_{\mathcal{W}_i} = D_{\mathcal{W}_j} \text{ for } i, j = 1, \ldots, K. \tag{14}$$

To achieve (13) and (14) with probability one, a general codebook $\mathcal{W}$ is first generated. Then every user rotates the common codebook by a random unitary matrix $\boldsymbol{T}_i$, $\boldsymbol{T}_i^H \boldsymbol{T}_i = \boldsymbol{I}_N$. These rotation matrices can be randomly generated either at the basestation side or at the user side in a distributed manner. In both cases, the basestation should have full knowledge of all the rotation matrices while each user only needs to know its own rotation matrix.

Thus, the codebook used at user $i$ is

$$\mathcal{W}_i = \boldsymbol{T}_i \mathcal{W} = \{\boldsymbol{T}_i \boldsymbol{w}_1, \ldots, \boldsymbol{T}_i \boldsymbol{w}_L \}.$$

Under i.i.d. Rayleigh fading, the channel vector is i.i.d. Gaussian distributed which is invariant with respect to unitary rotation. This means these rotated codebooks have the same mean square quantization errors

$$D_{\mathcal{W}} = D_{\mathcal{W}_i}, \ i = 1, \ldots, K.$$

The channel knowledge at the basestation is then

$$\boldsymbol{H}_w = [\mathcal{Q}_{\mathcal{W}_1}(\boldsymbol{h}_1), \ldots, \mathcal{Q}_{\mathcal{W}_K}(\boldsymbol{h}_K)]^T$$

with $\mathrm{rank}(\boldsymbol{H}_w) = N$ with probability one. Furthermore, we will model the codebook's conditional behavior as

$$\mathbb{E}\left\{(\boldsymbol{h}_i - \mathcal{Q}_{\mathcal{W}_i}(\boldsymbol{h}_i))\,(\boldsymbol{h}_i - \mathcal{Q}_{\mathcal{W}_i}(\boldsymbol{h}_i))^H \mid \mathcal{Q}_{\mathcal{W}_i}(\boldsymbol{h}_i)\right\} = \frac{D}{N}\boldsymbol{I}_N \tag{15}$$

and

$$\mathbb{E}\left\{\boldsymbol{h}_i - \mathcal{Q}_{\mathcal{W}_i}(\boldsymbol{h}_i) \mid \mathcal{Q}_{\mathcal{W}_i}(\boldsymbol{h}_i)\right\} = \boldsymbol{0}. \tag{16}$$

*B. Limited Feedback with ZFDPC*

Now we are ready to analyze the sum rate performance of the limited feedback ZFDPC scheme. The basestation assumes $\boldsymbol{H}_w$ to be the perfect CSI and applies the QR-type decomposition $\boldsymbol{H}_w = \boldsymbol{G}_w\boldsymbol{Q}_w$ to get the precoding matrix $\boldsymbol{Q}_w^H$. The resulting input-output relation is

$$\begin{aligned}
\boldsymbol{y} &= \boldsymbol{H}\boldsymbol{Q}_w^H\boldsymbol{s} + \boldsymbol{v} \\
&= (\boldsymbol{H}_w + \boldsymbol{\Delta})\boldsymbol{Q}_w^H\boldsymbol{s} + \boldsymbol{v} \\
&= \boldsymbol{G}_w\boldsymbol{s} + \boldsymbol{\Delta}\boldsymbol{Q}_w^H\boldsymbol{s} + \boldsymbol{v}
\end{aligned}$$

where $\boldsymbol{\Delta} = \boldsymbol{H} - \boldsymbol{H}_w$ is the difference between the quantized channel matrix and the perfect channel matrix. Its $i$th row vector $(\boldsymbol{\Delta})_i^T$ is just the quantization error at user $i$

$$(\boldsymbol{\Delta})_i = \boldsymbol{h}_i - \mathcal{Q}_{\mathcal{W}_i}(\boldsymbol{h}_i).$$

For the $i$th user, we have

$$y_i = g_{ii}^w s_i + \sum_{j<i} g_{ij}^w s_j + (\boldsymbol{\Delta})_i^T \boldsymbol{Q}_w^H \boldsymbol{s} + v_i \tag{17}$$

where $g_{ij}^w$ is the $(i, j)$ element in $\boldsymbol{G}_w$. The gain $g_{ii}^w$ is revealed to user $i$. We will assume that $\boldsymbol{s}$ is generated using successive dirty paper coding with Gaussian codebooks such that $\mathbb{E}\left\{\boldsymbol{s}\boldsymbol{s}^H\right\} = \frac{P}{N}\boldsymbol{I}_N$. The basestation encoder is assumed to have perfect knowledge of the noise variance and the quantizer distortion. The receiver $i$ uses minimum distance decoding to recover the transmitted codeword assuming a multiplicative channel value of $g_{ii}^w$. A discussion of decoding for dirty paper coding can be found in [48].

Comparing (17) with (3), we see there is an additional term $(\boldsymbol{\Delta})_i^T \boldsymbol{Q}_w^H \boldsymbol{s}$ in the received signal. Since the receiver $i$ only has knowledge of $(\boldsymbol{\Delta})_i^T$ and the basestation only has knowledge of $\boldsymbol{Q}_w^H$, $(\boldsymbol{\Delta})_i^T \boldsymbol{Q}_w^H$ is unknown at both the basestation and the receiver. Therefore, $u_i = (\boldsymbol{\Delta})_i^T \boldsymbol{Q}_w^H \boldsymbol{s}$ appears as cross user interference to the $i$th receiver and is treated as noise during decoding. The interference term $u_i$ has the following property.

*Lemma 2:* For a Gaussian source signal $\boldsymbol{s}$ with equal power allocation $\mathbb{E}\{\boldsymbol{s}\boldsymbol{s}^H\} = \frac{P}{N}\boldsymbol{I}_N$, the cross user interference $u_i$ has variance

$$\mathbb{E}\left\{u_i u_i^H \mid \boldsymbol{H}_w\right\} = DP/N$$

where $D$ is the average distortion of the codebook defined in (12).

*Proof:* Due to the unitary property of $\boldsymbol{Q}_w^H$, $\boldsymbol{Q}_w^H \boldsymbol{s}$ is also i.i.d. Gaussian distributed with zero mean and variance $\frac{P}{N}\boldsymbol{I}_N$. The source signal $\boldsymbol{s}$ is independent of the quantization error $(\Delta)_i^T$. Therefore, the variance of the interference is

$$\begin{aligned} \mathbb{E}\{u_i u_i^H \mid \boldsymbol{H}_w\} &= \frac{P}{N}\mathbb{E}\left\{(\boldsymbol{\Delta})_i^T (\boldsymbol{\Delta})_i^* \mid \boldsymbol{H}_w\right\} \\ &= \frac{P}{N}\mathbb{E}\left\{\|\mathcal{Q}_{\mathcal{W}_i}(\boldsymbol{h}_i) - \boldsymbol{h}_i\|^2 \mid \boldsymbol{H}_w\right\} \\ &= \frac{DP}{N}. \end{aligned}$$

The third equality comes from

$$\mathbb{E}\left\{\|\mathcal{Q}_{\mathcal{W}_i}(\boldsymbol{h}_i) - \boldsymbol{h}_i\|^2 \mid \boldsymbol{H}_w\right\} = D, \tag{18}$$

which is a direct result of (15). Here we take average over the quantization error by assuming each codeword experiences the channel state ergodically. ■

Thus, under limited feedback, the ZFDPC precoding gives us a lower triangular channel $\boldsymbol{G}_w$ with conditional interference-plus-noise power equal to $1+DP/N$ for each subchannel. By applying successive dirty paper coding to the lower triangular channel $\boldsymbol{G}_w$, the supremum of all achievable sum rates using limited feedback and Gaussian ZFDPC encoding, $R_{\lim}^{dpc}$, is bounded as

$$R_{\lim}^{dpc} \le \sum_{i=1}^{N} \mathbb{E}\left\{\log_2\left(1 + \frac{|g_{ii}^w|^2 P/N}{1+DP/N}\right)\right\} \tag{19}$$

where equal power allocation is assumed and the expectation is over $\boldsymbol{H}_w$. This follows from the generalized mutual information work in [49]–[51] using i) the assumption of Gaussian codebooks, ii) (15) and (16), iii) Lemma 2, and iv) the fact that $D$ and $\mathbb{E}\{|v_i|^2\} = 1$ are known at the encoder. Note that the bound in (19) is a necessary, rather than sufficient, condition for achievability.

Comparing $R_{\lim}^{dpc}$ with $R^{dpc-eq}$, we see that not only is the $\mathrm{SNR} = P/N$ replaced by the $\mathrm{SINR} = P/(N+DP)$ but also the effective channel gain is now $g_{ii}^w$ instead of $g_{ii}$. Since it is hard to quantify $g_{ii}^w$, we derive the following upper bound.

*Theorem 1:* $R_{\lim}^{dpc}$ in (19) is upper bounded by

$$R_{\lim}^{dpc} \le N \log_2\left(1 + \frac{P(N-D)}{N+PD}\right). \tag{20}$$

*Proof:* Using Jensen's inequality on (19), we have

$$R_{\lim}^{dpc} \le N\mathbb{E}\left\{\log_2\left(1 + \frac{P}{N+DP}\frac{1}{N}\sum_{i=1}^{N}|g_{ii}^w|^2\right)\right\}. \tag{21}$$

For a QR-type decomposition, we have

$$|g_{ii}^w|^2 \le \sum_{j=1}^{i}|g_{ij}^w|^2 = \|(\boldsymbol{H}_w)_i^T\|^2 = \|\boldsymbol{w}_{l_i}\|^2 \tag{22}$$

where the equality $\sum_{j=1}^{i}|g_{ij}^{w}|^2 = \|(\boldsymbol{H}_w)_i^T\|^2$ comes from the orthogonality of the columns of the $\boldsymbol{Q}$ in the decomposition. By substituting (22) into (21), we have

$$\begin{aligned} R_{\lim}^{dpc} &\leq N\mathbb{E}\left\{\log_2\left(1+\frac{P}{N+DP}\frac{1}{N}\sum_{i=1}^{N}\|\boldsymbol{w}_{l_i}\|^2\right)\right\} \\ &\leq N\log_2\left(1+\frac{P}{N+DP}\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\{\|\boldsymbol{w}_{l_i}\|^2\}\right) \end{aligned} \quad (23)$$

where the second inequality is from Jensen's inequality. Since the channel vectors are i.i.d. distributed we have

$$\mathbb{E}\{\|\boldsymbol{w}_{l_i}\|^2\} = \mathbb{E}\{\|\boldsymbol{w}_{l_j}\|^2\} \text{ for } i,j=1,\ldots,K.$$

Let $\bar{\eta} = \mathbb{E}\{\|\boldsymbol{w}_{l_i}\|^2\}$. It can be seen that $\mathbb{E}\{\|\boldsymbol{w}_{l_i}\|^2\} = N - D$ by taking expectation on both sides of (18) over $\mathcal{Q}_{\mathcal{W}_i}(\boldsymbol{h}_i) = \boldsymbol{w}_{l_i}$. After replacing this into (23), we get the bound. ■

For the case $L \to \infty$, i.e., the perfect feedback case, we have $D \to 0$ and $\bar{\eta} \to \mathbb{E}\left\{\|\boldsymbol{h}_i\|^2\right\} = N$. The bound changes to

$$R^{dpc} \leq N\log_2(1+P)$$

which can be seen as the result of applying the Jensen's inequality to the throughput of $N$ non-interfering MISO channels with i.i.d. Rayleigh fading and transmit power $P/N$.

This bound circumvents the difficult problem of analytical evaluation of $g_{ii}^{w}$ and provides important insights into the effect of limited feedback on the sum rate performance of multiple antenna broadcast system. First, it implies that a ceiling effect occurs on the sum rate under limited feedback for asymptotically high SNR.

*Corollary 1:* For a given general codebook $\mathcal{W}$, there exists a sum rate ceiling of $R_{\lim}^{dpc}$ as the SNR increases asymptotically high,

$$\lim_{P\to\infty} R_{\lim}^{dpc} = N\log_2\left(\frac{N}{D}\right). \quad (24)$$

*Proof:* Let the SNR increase in the upper bound in Theorem 1; we have

$$\begin{aligned} \lim_{P\to\infty} R_{\lim}^{dpc} &\leq \lim_{P\to\infty} N\log_2\left(1+\frac{P(N-D)}{N+DP}\right) \\ &\leq N\log_2\left(\frac{N}{D}\right). \end{aligned}$$

Since $N/D$ is fixed for a given codebook, the asymptotic bound does not increase as the SNR becomes large. Therefore, it is a sum rate ceiling for ZFDPC scheme under limited feedback. ■

Intuitively speaking, the ceiling effect is because the power of the cross user interference caused by the mismatch between $\boldsymbol{H}$ and $\boldsymbol{H}_w$ is related to the signal power. For a system with a fixed codebook, as the signal power $P$ increases, the power of the interference-plus-noise also increases linearly. To avoid the ceiling effect, we should at least let the interference power keep constant as $P$ increases. This also enables us to roughly compute the feedback rate required for an applicable limited feedback system. For the interference power $\frac{DP}{N}$ to be at least constant as $P$ increases, we should have $D$ of order $O(N/P)$. From the rate-distortion theorem [47], we know that the number of bits $b$ necessary for each user to represent its $N \times 1$ channel vector $\boldsymbol{h}_i$ with average distortion $D$ is

$$b = N \log_2(N/D). \tag{25}$$

By replacing $D$ with $O(N/P)$ in the rate-distortion function (25), we have

$$b = O(N \log_2 P)$$

which is the approximate number of bits necessary for the system to avoid the sum rate ceiling. We see that $b$ has to increase logarithmically with $P$ and should scale linearly as the number of transmit antenna $N$ grows. For example, when $P = 10\,\mathrm{dB}$ and $N = K = 4$, we have $b \approx 13$ bits for each user.

We can see that the sum rate ceiling is directly linked with the *mean square error* $D$ of the codebook used. This shows the validity of using *minimum distance selection* and *mean square error* distortion in quantization codebook design for multiple antenna broadcast channels. Intuitively speaking, the sum rate ceiling is caused by the cross user interference term whose variance increases as $\mathbb{E}\{\|\mathcal{Q}_{\mathcal{W}}(\boldsymbol{h}_i) - \boldsymbol{h}_i\|^2\}$ increases. To minimize the cross user interference, we want to select the vector which is closest to $\boldsymbol{h}_i$ in the codebook. For codebooks that are optimal in the sense of average squared error, the average distortion $D$ decreases as the size of the codebook $L$ increases. Thus it means the ceiling becomes higher as $L$ increases.

*C. Limited Feedback with Channel Inversion*

When limited feedback is combined with the channel inversion scheme, the basestation uses $\boldsymbol{H}_w^{-1}$ as the beamforming matrix. The received signal is

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{H}\boldsymbol{H}_w^{-1}\boldsymbol{s} + \boldsymbol{v} \\ &= \boldsymbol{s} + \boldsymbol{\Delta}\boldsymbol{H}_w^{-1}\boldsymbol{s} + \boldsymbol{v}. \end{aligned}$$

Due to the power constraint, we have

$$\mathbb{E}\left\{\|\boldsymbol{H}_w^{-1}\boldsymbol{s}\|^2 \mid \boldsymbol{H}_w\right\} = \sum_{i=1}^{K} P_i[(\boldsymbol{H}_w\boldsymbol{H}_w^H)^{-1}]_{i,i} = P.$$

We will assume that $\boldsymbol{s}$ is generated using Gaussian codebooks. For analytical simplicity, we assume equal power allocation, that is,

$$\mathbb{E}\left\{\boldsymbol{s}\boldsymbol{s}^H \mid \boldsymbol{H}_w\right\} = \frac{P\boldsymbol{I}_N}{\sum_{i=1}^{K}[(\boldsymbol{H}_w\boldsymbol{H}_w^H)^{-1}]_{i,i}} = \|\boldsymbol{H}_w^{-1}\|_F^{-2}P\boldsymbol{I}_N. \tag{26}$$

As before, we assume the receiver uses minimum distance decoding (assuming a multiplicative channel gain of one) to try to recover the transmitted codeword. Thus, the receiver performs decoding as if the signal was transmitted over an additive white Gaussian noise channel.

Again, there is the cross user interference $\zeta_i = (\boldsymbol{\Delta})_i^T\boldsymbol{H}_w^{-1}\boldsymbol{s}$ in the received signal of the $i$th user.

*Lemma 3:* For a Gaussian source signal $\boldsymbol{s}$ with equal power allocation $\mathbb{E}\left\{\boldsymbol{s}\boldsymbol{s}^H \mid \boldsymbol{H}_w\right\} = \frac{P}{\|\boldsymbol{H}_w^{-1}\|_F^2}\boldsymbol{I}_N$, the cross user interference $\zeta_i$ has variance

$$\mathbb{E}\left\{\zeta_i\zeta_i^H \mid \boldsymbol{H}_w\right\} = DP/N$$

where $D$ is the average distortion of the codebook defined in (12).

*Proof:* We have

$$\begin{aligned} \mathbb{E}\left\{\zeta_i\zeta_i^H \mid \boldsymbol{H}_w\right\} &= \mathbb{E}\left\{(\boldsymbol{\Delta})_i^T\boldsymbol{H}_w^{-1}\boldsymbol{s}\boldsymbol{s}^H\boldsymbol{H}_w^{-H}(\boldsymbol{\Delta})_i^* \mid \boldsymbol{H}_w\right\} \\ &= \mathbb{E}\left\{\frac{P}{\|\boldsymbol{H}_w^{-1}\|_F^2}(\boldsymbol{\Delta})_i^T\boldsymbol{H}_w^{-1}\boldsymbol{H}_w^{-H}(\boldsymbol{\Delta})_i^* \mid \boldsymbol{H}_w\right\} \\ &= \mathbb{E}\left\{\frac{P}{\|\boldsymbol{H}_w^{-1}\|_F^2}\operatorname{tr}\left(\boldsymbol{H}_w^{-H}(\boldsymbol{\Delta})_i^*(\boldsymbol{\Delta})_i^T\boldsymbol{H}_w^{-1}\right) \mid \boldsymbol{H}_w\right\}. \end{aligned} \tag{27}$$

According to (15), we have $\mathrm{tr}\left(\boldsymbol{H}_w^{-H}\mathbb{E}\left\{(\boldsymbol{\Delta})_i^*(\boldsymbol{\Delta})_i^T \mid \boldsymbol{H}_w\right\}\boldsymbol{H}_w^{-1}\right) = \|\boldsymbol{H}_w^{-1}\|_F^2 \frac{D}{N}$. Replacing it into (27), we get

$$\mathbb{E}\left\{\zeta_i\zeta_i^H \mid \boldsymbol{H}_w\right\} = DP/N.$$

■

Since the conditional interference-plus-noise power of each user is $1 + DP/N$, we can obtain the following ergodic sum rate bound. The supremum of all achievable sum rates using a limited feedback codebook of distortion $D$ and Gaussian codebooks, denoted $R_{\mathrm{lim}}^{ci}$, is bounded as

$$R_{\mathrm{lim}}^{ci} \leq N\mathbb{E}\left\{\log_2\left(1 + \frac{P\|\boldsymbol{H}_w^{-1}\|_F^{-2}}{1+PD/N}\right)\right\}. \tag{28}$$

Again, this result uses the generalized mutual information work in [49]–[51]. Also as before, the bound in (28) is a necessary condition for achievability.

This bound (28) can be further bounded as follows.

*Theorem 2:* $R_{\mathrm{lim}}^{ci}$ in (28) is upper bounded by

$$R_{\mathrm{lim}}^{ci} \leq N\log_2\left(1 + \frac{P(N-D)}{N+PD}\right). \tag{29}$$

*Proof:* Since $\boldsymbol{H}_w$ is nonsingular with probability 1, we can express $\|\boldsymbol{H}_w^{-1}\|_F^{-2}$ in terms of the singular values of $\boldsymbol{H}_w$, that is,

$$\begin{aligned}\|\boldsymbol{H}_w^{-1}\|_F^{-2} &= \left(\|\boldsymbol{H}_w^{-1}\|_F^2\right)^{-1} \\ &= \left(\sum_{i=1}^{N}\frac{1}{|\lambda_i^w|^2}\right)^{-1}\end{aligned} \tag{30}$$

where $\lambda_i^w$ is the $i$th singular value of $\boldsymbol{H}_w$. Because $|\lambda_1^w|^2,\ldots,|\lambda_N^w|^2$ are all positive, their harmonic mean is less than their arithmetic mean, which gives

$$N\left(\sum_{i=1}^{N}\frac{1}{|\lambda_i^w|^2}\right)^{-1} \leq \frac{1}{N}\sum_{i=1}^{N}|\lambda_i^w|^2. \tag{31}$$

By combining (30) with (31) and replacing $\sum_{i=1}^{N}|\lambda_i^w|^2$ with $\|\boldsymbol{H}_w\|_F^2$, we get

$$\begin{aligned}\|\boldsymbol{H}_w^{-1}\|_F^{-2} &\leq \frac{1}{N^2}\|\boldsymbol{H}_w\|_F^2 \\ &= \frac{1}{N^2}\sum_{i=1}^{N}\|\boldsymbol{w}_{l_i}\|^2.\end{aligned} \tag{32}$$

Therefore, the $R_{\text{lim}}^{ci}$ in (28) is bounded by

$$R_{\text{lim}}^{ci} \leq N\mathbb{E}\left\{\log_2\left(1+\frac{P}{N^2(1+PD/N)}\sum_{i=1}^{N}\|\boldsymbol{w}_{l_i}\|^2\right)\right\}. \tag{33}$$

From Jensen's inequality and the fact that $\mathbb{E}\{\sum_{i=1}^{N}\|\boldsymbol{w}_{l_i}\|^2\} = N\bar{\eta} = N(N-D)$, we get

$$R_{\text{lim}}^{ci} \leq N\log_2\left(1+\frac{P(N-D)}{N+PD}\right)$$

which is the same bound as in Theorem 1. ■

Due to the same reason, CI also experiences a sum rate ceiling for limited feedback under a fixed codebook as in the ZFDPC case. To maintain the sum rate growth, we need approximately $O(N\log_2 P)$ feedback bits per user which makes the cross user interference keep constant as the signal power increases. We should remark that the bound is looser for $R_{\text{lim}}^{ci}$ than for $R_{\text{lim}}^{dpc}$ because $\|\boldsymbol{H}_w^{-1}\|_F^{-2}$ has a much higher probability of being near zero than $\frac{1}{N^2}\sum_{i=1}^{N}\|\boldsymbol{w}_{l_i}\|^2$. This is also the reason why CI has a lower throughput than ZFDPC under limited feedback. Even though the power of interference-plus-noise is the same, the effective channel gain of CI is more likely to be in a deep fade than the ZFDPC case.

## V. Simulation Results

In this section, we give some numerical results on the ergodic sum rate performance of these two partial CSI models discussed in the previous sections. Throughout the simulations, the channel is assumed to be independent Rayleigh fading, that is, each entry of $\boldsymbol{H}$ is independently $\mathcal{C}N(0,1)$ distributed. Since the noise power is normalized, the plotted SNR in the figures is $\text{SNR} = 10\log_{10} P$. For a given setting, the ergodic sum rate is obtained by averaging over at least 20000 channel realizations.

The ergodic ratio of $R_{\text{sha},P>>0}^{dpc}$ to $R^{dpc*}$ is plotted with respect to SNR in Fig. 1. The number of transmit antennas is $N = K = 8$. We can see that the $R_{\text{sha},P>>0}^{dpc}$ is almost the same as $R^{dpc*}$ when $\text{SNR} \geq 15\,\text{dB}$, which indicates the asymptotic optimality of ZFDPC with shape feedback in the sense of sum rate for high SNR.

Fig. 2 plots the growth of $R_{\text{sha},P>>0}^{dpc}$ with the number of transmit antennas for $\text{SNR} = 20\,\text{dB}$. It can be observed that the sum rate of shape feedback has the same growth rate as that of perfect transmitter

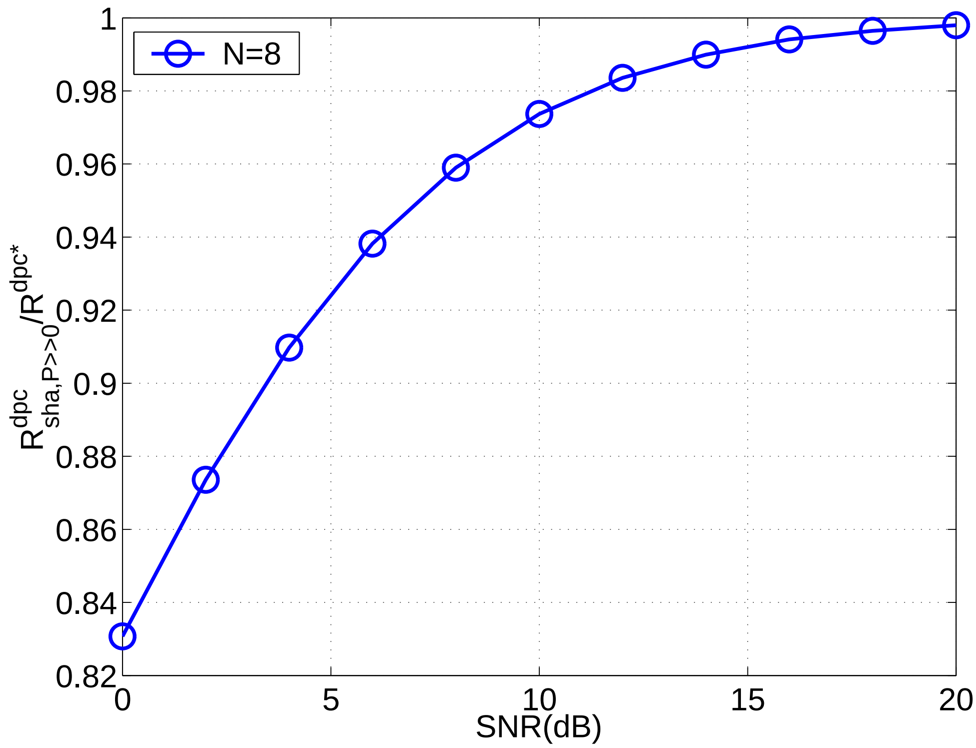


Fig. 1. The ergodic ratio of $R^{dpc}_{\mathrm{sha},P>>0}$ to $R^{dpc*}$ with respect to SNR for $N = 8$ and $K = 8$.

channel knowledge. It shows that shape feedback captures most of the throughput advantage provided by the multiple antennas through the ZFDPC encoding.

The sum rate performance of shape feedback with channel inversion is shown in Fig. 3 and Fig. 4. In both figures, the ergodic performance of $R^{ci}_{\mathrm{sha}}$ is superior to $R^{ci-eq}$ and very close to $R^{ci*}$. From Fig. 3, we see that the $R^{ci}_{\mathrm{sha}}$ has about $1.5\,\mathrm{dB}$ gain over $R^{ci-eq}$ for $N = K = 8$. When $\mathrm{SNR} > 15\,\mathrm{dB}$, $R^{ci}_{\mathrm{sha}}$ and $R^{ci*}$ have almost the same ergodic performance. Fig. 4 shows that the gain of $R^{ci}_{\mathrm{sha}}$ over $R^{ci-eq}$ also increases as the $N$ increases.

For the simulation of limited feedback, the locally optimal codebooks are obtained by training from a large set of training vectors generated according to i.i.d. Rayleigh fading. The variance of the cross user interference observed from Monte Carlo simulation is very close to the theoretical value $DP/N$. The solid curves in Fig. 5 show the ergodic performance of $R^{dpc}_{\lim}$ for $N = K = 4$ when the feedback

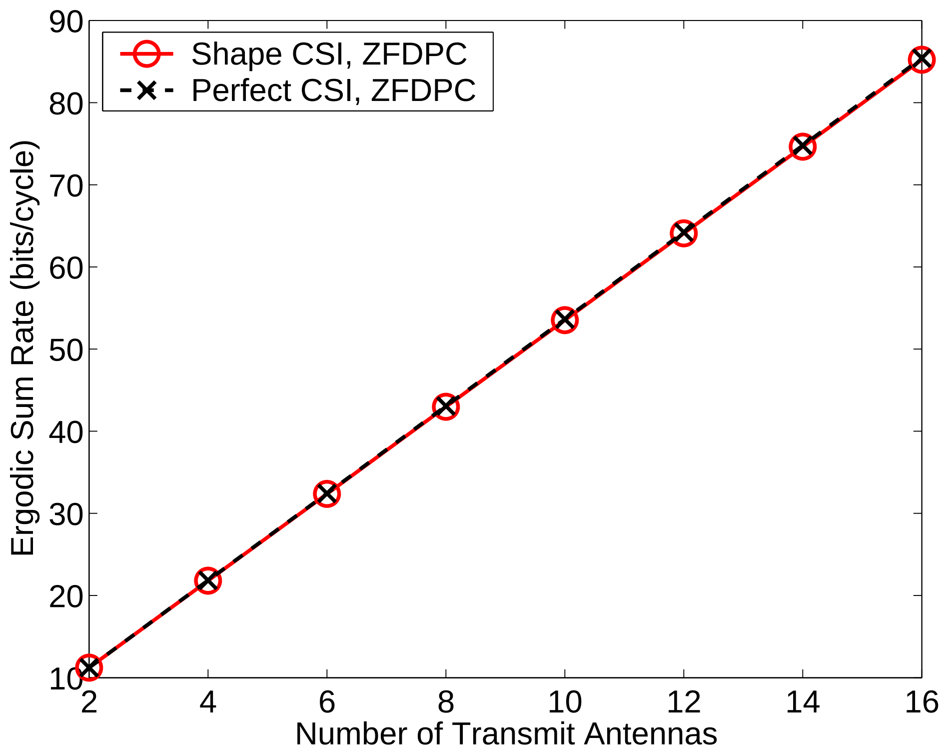


Fig. 2. Ergodic performance of $R^{dpc}_{\mathrm{sha},P>>0}$ and $R^{dpc*}$ with respect to $N$ for $\mathrm{SNR} = 20\,\mathrm{dB}$.

rate of each user is $5$bits, $10$bits and the perfect CSI case. The respective upper bounds are also plotted in dashed lines for comparison. The ceiling effect can be clearly observed when the feedback rate is finite. The sum rate curves increase linearly with $\mathrm{SNR}$ for low $\mathrm{SNR}$ and become flat for high $\mathrm{SNR}$. The upper bound curves given in Fig. 5 also indicate that a fixed overhead rate is not sufficient to obtain sum rate growth for $\mathrm{SNR} \geq 20\,\mathrm{dB}$. These results match our discussion on the approximation of necessary overhead rate.

The results for the ergodic sum rate of limited feedback with channel inversion are presented in Fig. 6 for $N = K = 4$ and feedback rate equals $5$bits and $10$bits of each user. $R^{ci-eq}$ is also plotted for comparison reason. The sum rate ceiling can be observed, and their performance is worse than limited feedback with ZFDPC with the same feedback rate.

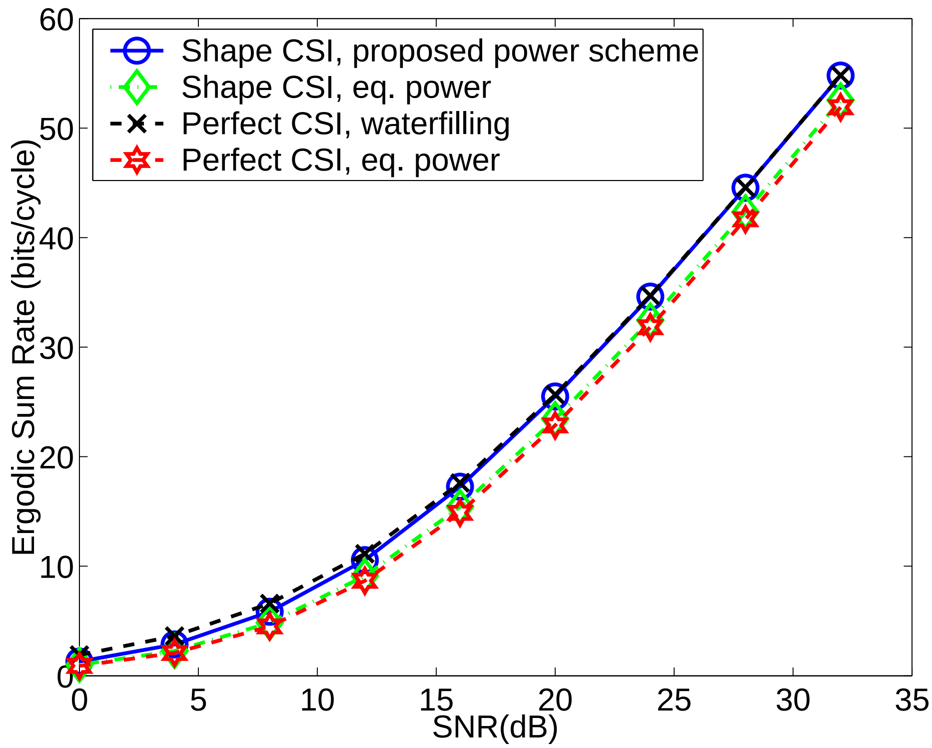


Fig. 3. Ergodic performance of $R^{ci}_{\text{sha}}$, $R^{ci-eq}_{\text{sha}}$, $R^{ci*}$, and $R^{ci-eq}$ with respect to SNR for $N = 8$ and $K = 8$.

## VI. Conclusion

In this paper, we considered two kinds of partial CSI models for the basestation in multiple antenna broadcast systems. We showed that the shape feedback achieves the optimal sum rate for asymptotically high SNR when combined with the ZFDPC scheme. When using CI, shape feedback can obtain a sum rate larger than CI with perfect CSI under equal power allocation by using the proposed power allocation strategy. For the limited feedback model, we proposed a limited feedback scheme that avoids the ill-conditioning of basestation CSI by randomly rotating a general codebook known at each receiver. We derived upper bounds for the ergodic sum rate of limited feedback under both ZFDPC and CI. The bound gives critical insight about the sum rate performance of limited feedback. It shows that the systems experience a ceiling effect on the sum rate for a fixed feedback rate.

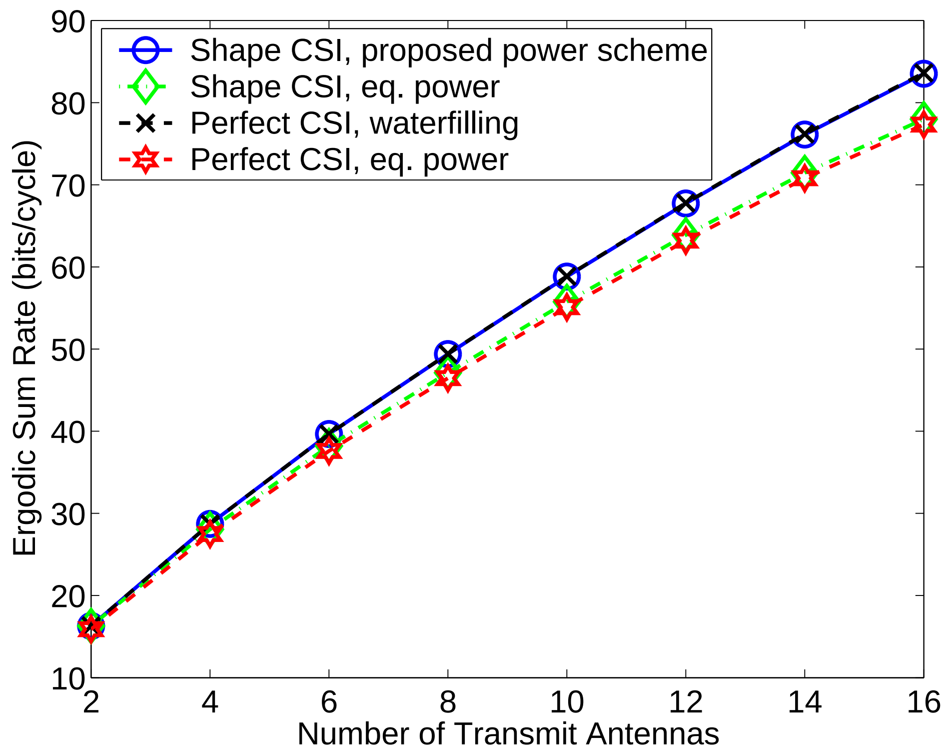


Fig. 4. Ergodic performance of $R_{\rm sha}^{ci}$, $R_{\rm sha}^{ci-eq}$, $R^{ci*}$, and $R^{ci-eq}$ with respect to $N$ for $\mathrm{SNR} = 30\,\mathrm{dB}$.

## REFERENCES


[1] T. Cover, "Broadcast channels," *IEEE Trans. Inform. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.

[2] Q. H. Spencer, C.B. Peel, A. L. Swindlehurst, and M. Haardt, "An introduction to the multi-user MIMO downlink," *IEEE Comm. Magazine*, vol. 42, no. 10, pp. 60–67, Oct. 2004.

[3] G. Caire and S. Shamai, "On the achievable throughput of a multi-antenna Gaussian broadcast channel," *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1691–1706, July 2003.

[4] P. Viswanath and D. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inform. Theory*, vol. 49, no. 8, pp. 1912–1921, August 2003.

[5] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates and sum capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inform. Theory*, vol. 43, no. 8, pp. 2658–2668, August 2003.

[6] W. Yu and J. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 1875–1892, Sept. 2004.

[7] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, no. 5, pp. 439–441, May 1983.

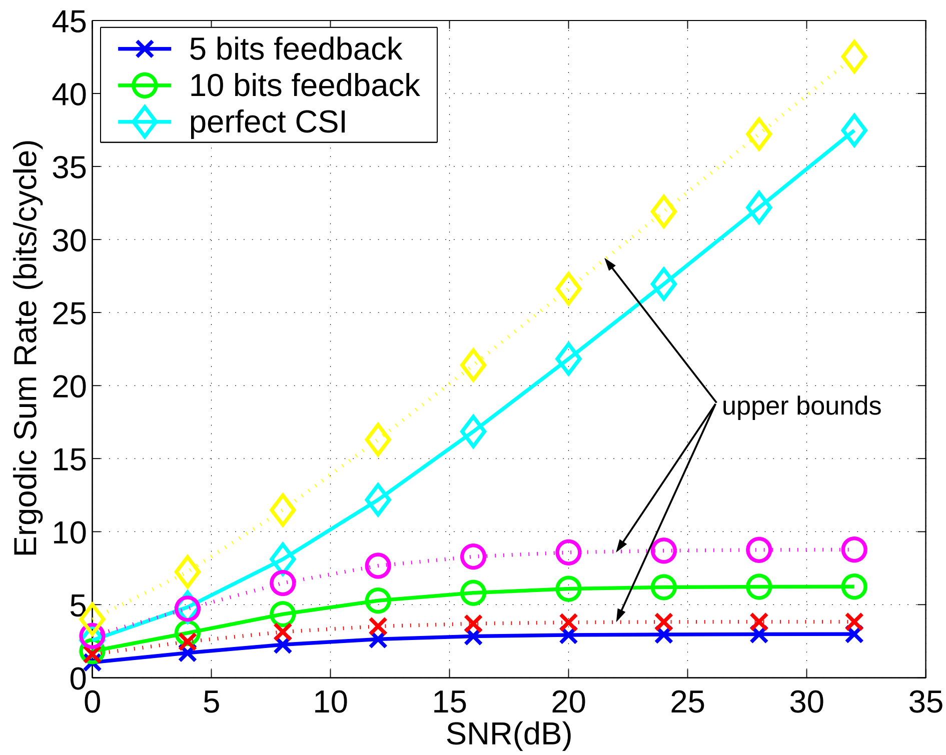


Fig. 5. Ergodic sum rate of limited feedback with ZFDPC (in solid lines) and the upper bound (in dashed lines) for $N = 4$ and $K = 4$.

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

[9] C. Peel, B. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multi-antenna multi-user communication – Part I: Channel inversion and regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195–202, Jan. 2005.

[10] B. Hochwald, C. Peel, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multi-antenna multi-user communication – Part II: Perturbation," *IEEE Trans. Commun.*, vol. 54, no. 2, Feb. 2005.

[11] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channels with partial side information," *IEEE Trans. Inform. Theory*, vol. 51, no. 2, pp. 506–522, Feb. 2005.

[12] T. Yoo and A. J. Goldsmith, "On the optimality of multi-antenna broadcast scheduling using zero-forcing beamforming," *To appear IEEE Journal of Selected Areas of Commun. Special Issue on 4G Wireless systems*.

[13] C. Swannack, E. Uysal-Biyikoglu, and G. W. Wornell, "Low-complexity multi-user scheduling for maximizing throughput

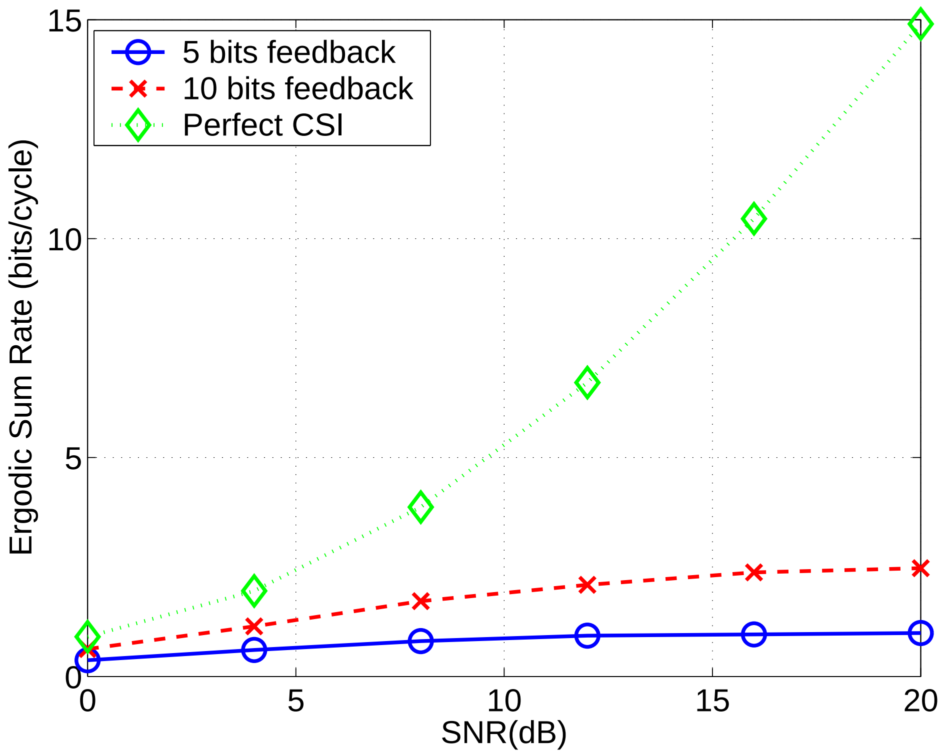


Fig. 6. Ergodic sum rate of limited feedback with channel inversion for $N = 4$ and $K = 4$.

in the MIMO broadcast channel," in *Proc. Allerton Conf. Commun., Control and Computing*, Oct. 2004.

[14] M. Fakhereddin, M. Sharif, and B. Hassibi, "Throughput analysis in wideband MIMO broadcast channels with partial feedback," in *Proc. IEEE SPAWC*, 2005, pp. 771–775.

[15] H. Lee, M. Shin, and C. Lee, "An eigen-based MIMO multiuser scheduler with partial feedback information," *IEEE Commun. Letter*, vol. 9, no. 4, pp. 328–330, Apr. 2005.

[16] U. Madhow and E. Visotsky, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inform. Theory*, vol. 47, no. 9, pp. 2632–2639, Sept. 2001.

[17] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2599–2613, Oct. 2002.

[18] S. A. Jafar, S. Vishwanath, and A. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback," in *Proc. IEEE ICC*, 2001, vol. 7, pp. 2266–2270.

[19] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel correlations," *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1673–1690, July 2003.

[20] J. C. Roh and B. D. Rao, "Multiple antenna channels with partial channel state information at the transmitter," *IEEE Trans. Wireless Commun.*, vol. 3, no. 2, pp. 677–688, Mar. 2004.

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

[22] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple antenna systems," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2562–2579, Oct. 2003.

[23] M. Skoglund and G. Jongren, "Quantized feedback information in orthogonal space-time block coding," *IEEE Trans. Inform. Theory*, vol. 50, no. 10, pp. 2473–2486, Oct. 2004.

[24] D. J. Love and R. W. Heath Jr., "Limited feedback unitary precoding for orthogonal space-time block codes," *IEEE Trans. Signal Processing*, vol. 53, no. 1, pp. 64–73, Jan. 2005.

[25] J. Akhtar and D. Gesbert, "Extending orthogonal block codes with partial feedback," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1959–1962, Nov. 2004.

[26] D. J. Love and R. W. Heath Jr., "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Trans. Inform. Theory*, vol. 51, no. 8, pp. 2967–2976, Aug. 2005.

[27] J. C. Roh and B. D. Rao, "Channel feedback quantization methods for MISO and MIMO systems," in *Proc. IEEE Intl. Symp. on Personal, Indoor and Mobile Radio Commun.(PIMRC)*, 2004, vol. 2, pp. 805–809.

[28] M. A. Sadrabadi, A. K. Khandani, and F. Lahouti, "A new method of channel feedback quantization for high data rate MIMO systems," in *Proc. IEEE GLOBECOM*, 2004, vol. 1, pp. 91– 95.

[29] R. S. Blum, "MIMO with limited feedback of channel state information," in *Proc. IEEE ICASSP*, 2003, vol. 4, pp. 89–92.

[30] V. K. N. Lau, Y. Liu, and T.-A. Chen, "On the design of MIMO block-fading channels with feedback-link capacity constraint," Jan. 2004, vol. 52, pp. 62–70.

[31] J. S. Kim, H. Kim, and K. B. Lee, "Limited feedback precoding for wireless MIMO broadcast channels," in *Proc. IST: Mobile and Wireless Commun. Summit*, 2005.

[32] J. Vicario and C. Anton-Haro, "Robust exploitation of spatial and multi-user diversity in limited feedback systems," in *Proc. IEEE ICASSP*, 2005, vol. 3, pp. 417–420.

[33] N. Jindal and A. Goldsmith, "DPC vs. TDMA for MIMO broadcast channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 5, pp. 1783–1794, May. 2005.

[34] A. P. Liavas, "Tomlinson-Harashima precoding with partial channel knowledge," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 5–9, Jan. 2005.

[35] F. W. Vook, X. Zhuang, K. L. Baum, T. A. Thomas, and P. Sartori, "Improvements to the uplink channel sounding signaling for OFDMA," Contribution to *IEEE 802.16 Broadband Wireless Access Working Group*, Nov. 2004.

[36] T. L. Marzetta and B. M. Hochwald, “Fast transfer of channel state information in wireless systems,” *submitted to IEEE Trans. Comm.*, June 2004.

[37] T. A. Thomas, K. L. Baum, and P. Sartori, “Obtaining channel knowledge for closed-loop multi-stream broadband MIMO-OFDM communications using direct channel feedback,” in *Proc. IEEE Globecom*, Nov. 2005.

[38] O. Simeone and U. Spagnolini, “Lower bound on training-based channel estimation error for frequency-selective block-fading Rayleigh MIMO channels,” *IEEE Trans. Signal Processing*, vol. 52, no. 11, pp. 3265–3277, Nov. 2004.

[39] L. Tong and S. Perreau, “Blind channel estimation: From subspace to maximum likelihood methods,” *IEEE Proceedings*, vol. 86, no. 10, pp. 1951–1968, Oct. 1998.

[40] N. Jindal, “MIMO broadcast channels with finite rate feedback,” in *Proc. IEEE GLOBECOM*, 2005, pp. 771–775.

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

[42] A.T. James, “Distributions of matrix variates and latent roots derived from normal samples,” *Ann. Math. Statist.*, vol. 35, pp. 475–501, 1964.

[43] A. Edelman, *Eigenvalues and Condition Numbers of Random Matrices*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, May 1989.

[44] M. S. Alouini and A. J. Goldsmith, “Capacity of Rayleigh fading channels under different adaptive transmission and diversity-combining techniques,” *IEEE Trans. Vehicular Technology*, vol. 48, no. 4, pp. 1165–1181, July 1999.

[45] H. Shin and J. H. Lee, “Capacity of multiple-antenna fading channels: spatial fading correlation, double scattering, and keyhole,” *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2636–2647, Oct. 2003.

[46] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, San Diego, CA, 1994.

[47] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, 1992.

[48] J. K. Su, J. J. Eggers, and B. Girod, “Illustration of the duality between channel coding and rate distortion with side information,” in *Proc. of the 34th Asilomar Conference on Signals, Systems and Computers*, Oct.-Nov. 2000, vol. 2, pp. 1841–1845.

[49] A. Lapidoth and S. Shamai, “Fading channels: How perfect need “perfect side information” be?,” *IEEE Trans. Inform. Theory*, vol. 48, no. 5, pp. 1118–1134, May. 2002.

[50] A. Lapidoth, “Nearest neighbor decoding for additive non-Gaussian noise channels,” *IEEE Trans. Inform. Theory*, vol. 42, no. 5, pp. 1520–1529, Sept. 1996.

[51] H. Weingarten, Y. Steinberg, and S. Shamai, “Gaussian codes and weighted nearest neighbor decoding in fading multiple-antenna channels,” *IEEE Trans. Inform. Theory*, vol. 50, no. 8, pp. 1665–1686, Aug. 2004.