On the Capacity and Design of Limited Feedback Multiuser MIMO Uplinks *

Il Han Kim and David J. Love  
School of Electrical and Computer Engineering  
Purdue University  
EE Building  
465 Northwestern Ave.  
West Lafayette, IN 47907  
(Ph) 765-496-6797  
(Fax) 765-494-0880  

kim385@ecn.purdue.edu, djlove@ecn.purdue.edu  

February 22, 2008

Abstract

The theory of multiple-input multiple-output (MIMO) technology has been well-developed to increase fading channel capacity over single-input single-output (SISO) systems. This capacity gain can often be leveraged by utilizing channel state information at the transmitter and the receiver. Users make use of this channel state information for transmit signal adaptation. In this correspondence, we derive the capacity region for the MIMO multiple access channel (MIMO MAC) when partial channel state information is available at the transmitters, where we assume a synchronous MIMO multiuser uplink. The partial channel state information feedback has a cardinality constraint and is fed back from the basestation to the users using a limited rate feedback channel. Using this feedback information, we propose a finite codebook design method to maximize sum-rate. In this correspondence, the codebook is a set of transmit signal covariance matrices. We also derive the capacity region and codebook design methods in the case that the covariance matrix is rank-one (i.e., beamforming). This is motivated by the fact that beamforming is optimal in certain conditions. The simulation results show that when the number of feedback bits increases, the capacity also increases. Even with a small number of feedback bits, the performance of the proposed system is close to an optimal solution with the full feedback.

Index Terms-Limited feedback, MAC systems, MIMO systems, CSI at receiver (CSIR), CSI at transmitter (CSIT), Rayleigh channels, Grassmannian subspace packing

---

*This work was supported in part by the NSF under grant CCF0513916, the Korea Research Foundation Grant funded by the Korea government (MOEHRD) (KRF-2005-215-D00207), and the SBC Foundation
1 Introduction

Multiple access systems are being widely studied to effectively support multiple users at the basestation during uplink transmission. In multiple-input multiple-output multiple access channels (MIMO MACs), it is known that the channel state information at the transmitters (CSIT) helps increase system performance (e.g., sum-rate capacity) regardless of whether it is perfect or partial [1–5]. CSIT can be full channel information [1–3], channel covariance information [4,5], channel mean information [4, 5], etc. Especially in frequency division duplexing (FDD) systems, because channel state information (CSI) is not reciprocal between uplink channel and downlink channel, this information should be fed back from the receiver to the transmitters. In real communication systems, however, the feedback resources allotted for CSI feedback must be a small portion of the data being transmitted on the downlink because feedback bits are overhead that causes performance degradation.

In this correspondence, we focus on limited feedback schemes, where the receiver is assumed to have perfect knowledge of the CSI and the receiver transmits a limited number of feedback bits to the transmitter according to the CSI. It is of our interest to optimize signal covariance matrices of each user using a limited number of feedback bits. The basestation and users are assumed to have knowledge of a codebook of covariance matrices designed offline. The limited number of feedback bits allows each user to identify the covariance matrix chosen by the basestation for them to use during transmission.

The main contributions of this correspondence are given below:

1) In MIMO systems, we prove the coding theorem for MACs when a limited number of feedback bits is available at the transmitters. Based on the proved coding theorem, we give a method to design a codebook (set of covariance matrices) for each user under a finite feedback cardinality constraint to maximize the sum-rate capacity. Assuming $B$ bits of feedback, the basestation selects one codeword that maximizes the sum-rate among $2^B$ elements in a codebook. This codebook is assumed to be known to both the basestation and the users. The index corresponding to the chosen codebook matrix is fed back to the users.

2) In single user MIMO scenarios, beamforming is optimal under certain conditions [6–10] when feedback is channel mean or channel covariance information. In multiuser scenarios, beamforming is generally not universally optimal, i.e., the full capacity rate region cannot be achieved. However, if some criteria are satisfied, the sum-rate is asymptotically achieved by beamforming. Studied in [11], the optimality of beamforming was derived when channel feedback is perfect. In [4, 5], the optimality of beamforming was proved when feedback is channel covariance or channel mean information. Motivated by these works, we constrain the rank of each user’s covariance matrix to one, i.e., each user performs beamforming on his/her own array. We derive the capacity region of the beamforming
system and propose two methods to design the beamforming vectors: eigenbeamforming with a codebook designed using Lloyd’s algorithm and Grassmannian subspace packing with random power allocation of each user.

2 System Overview

In this section, we introduce the system model under consideration. A typical MIMO MAC communication system is presented in Fig. 1. Let $M_t$ be the number of transmit antennas at each user and $M_r$ be the number of receive antennas. In this setup, the received signal is given by

$$y_n = \sum_{k=1}^{K} H_n^{(k)} x_n^{(k)} + z_n = \left[ H_n^{(1)} \ H_n^{(2)} \ldots \ H_n^{(K)} \right] \left[ x_n^{(1)} \ x_n^{(2)} \ldots \ x_n^{(K)} \right] + z_n \equiv H_n x_n + z_n, n = 1, \ldots, N, \tag{1}$$

where $K$ is the number of users and $x_n^{(k)} \in \mathcal{X}^{M_t}$ is the transmitted signal from the $k$-th user with $\mathcal{X}$ denoting the transmit symbol alphabet. Here, $n$ is the time index and $N$ is the block length of each code $x_1^N = \{x_1, \ldots, x_N\}$. $H_n^{(k)}$ is the $M_r \times M_t$ channel matrix from the $k$-th user to the basestation. $z_n$ is the complex Gaussian noise with zero mean and covariance $\sigma^2 I$. Here we assume that data symbols of each user are independent of each other, i.e., $E\left[ x_n^{(k)} \ (x_n^{(l)})^* \right] = O_{M_t \times M_t}, \forall k \neq l$, where $O_{M_t \times M_t}$ is $M_t \times M_t$ all zero matrix. We assume a Rayleigh flat fading environment, where the channel independently changes from transmission to transmission. We also assume that the channel state information $U_n \in \mathcal{U}$ and $V_n \in \mathcal{V}$ are available at the transmitters and the receiver, respectively, where $\mathcal{U}$ is the space of CSIT and $\mathcal{V}$ is the space of channel state information at the receiver (CSIR). CSIT $U_n$ is fed back to the transmitters based on the current CSIR $V_n$. We assume that the channel coefficients are estimated perfectly at the receiver. Since the channel fading is assumed to be independent for each $n$, the transmitted symbol is a function of the feedback CSIT $U_n$. We assume noiseless feedback, i.e., the users know $U_n$ perfectly.

Message $W^{(k)}$ for user $k, k = 1, \ldots, K$, is encoded to $x_n^{(k)}$ for $n = 1, \ldots, N$ depending on the feedback $U_n$. These encoded signals are sent through the channel with channel transition probability $p\left(y_1^N | x_1^N, H_1^N\right)$, where $H_1^N = \{H_1, \ldots, H_N\}$ and $y_1^N = \{y_1, \ldots, y_N\}$. After receiving $y_1^N$, the receiver decodes $y_1^N$ to $(\hat{W}^{(1)}, \ldots, \hat{W}^{(K)})$.

Assuming the discrete memoryless channel model, the system is characterized by the following channel transition probability

$$p\left(y_1^N | x_1^N, H_1^N\right) = \prod_{n=1}^{N} p(y_n | x_n, H_n). \tag{2}$$

This is illustrated in Fig. 2.
3 Capacity Results

In this section, we give capacity results when complete CSIR and partial CSIT are available at the receiver and the transmitters, respectively. Partial CSIT is assumed to consist of feedback information from the basestation to the users.

The system model of Fig. 2 can be transformed into an equivalent system without CSIR and CSIT [12–16] by seeing $U_n$ as a channel state and $V_n$ as a channel output. Similarly, in this multiple access system, we first encode each $W^{(k)}$ into the matrix $T_n^{(k)} \in \mathcal{X}^{M_t \times |U|}$, $n = 1, \ldots, N$, and select one column of $T_n^{(k)}$ according to the current feedback $U_n \in \mathcal{U}$ such that $x_n^{(k)} = T_n^{(k)}(U_n)$. With the probability distribution function $p(H_n, u_n, v_n)$, the equivalent channel is characterized by the channel transition probability [15]

$$p(y_n, v_n | T_n) = \sum_{u_n, H_n} p(H_n, u_n, v_n) p(y_n | T_n(U_n), H_n),$$

(3)

where $T_n = \left[ T_n^{(1)} \cdots T_n^{(K)} \right]^T$. Fig. 3 presents this equivalent system model.

We assume that the CSIR includes both the CSIT component and the channel coefficient matrix, and the receiver feeds back the $U_n$ component of the CSIT to the transmitters. This means that the receive channel state information is given by $V_n = (H_n, U_n)$ [16]. We assume the finite rate feedback satisfies $|\mathcal{U}| = 2^B$ with $\mathcal{U} = \{u^{(1)}, \ldots, u^{(2^B)}\}$.

To derive the multiple access capacity region, let $S = \{k_1, \ldots, k_{|S|}\} \subseteq \{1, \ldots, K\}$ be a user subset. Each user covariance matrix is given by $Q^{(k)}(U_n) = E\left[ x_n^{(k)} (x_n^{(k)})^* | U_n \right]$ for $n = 1, \ldots, N, k = 1, \ldots, K$, with the power constraint $\text{Tr} \left( Q^{(k)}(U) \right) \leq P_k$ for $\forall U \in \mathcal{U}$. This means that we do not allow the average sum transmit power conditioned on some feedback information to exceed $P_k$ for any feedback possibility. In this setup, the capacity region is given by the following theorem.

**Theorem 1** For a fixed mapping from $U$ to $\left\{Q^{(k)}(U)\right\}_{k=1}^K$ such that $\text{Tr} \left( Q^{(k)}(U) \right) \leq P_k$ for $k = 1, \ldots, K$, the capacity region is the $K$-dimensional polyhedron

$$\left\{ R : \sum_{k \in S} R^{(k)} \leq E_V \left[ \log_2 \det \left( I + \sigma^{-2} \sum_{k \in S} H^{(k)} Q^{(k)}(U) \left( H^{(k)} \right)^* \right) \right], \forall S \subseteq \{1, \ldots, K\} \right\},$$

(4)

where $R = [R^{(1)} \cdots R^{(K)}]^T$ is the rate vector. Additionally, for individual power constraints $\text{Tr} \left( Q^{(k)}(U) \right) \leq P_k$ for $k = 1, \ldots, K$, the capacity region is given by

$$C_{MAC}(P_1, \ldots, P_K) = \text{Cl} \left( \bigcup_{k=1}^K \left\{ R : \sum_{k \in S} R^{(k)} \leq E_V \left[ \log_2 \det \left( I + \sigma^{-2} \sum_{k \in S} H^{(k)} Q^{(k)}(U) \left( H^{(k)} \right)^* \right) \right], \forall S \subseteq \{1, \ldots, K\} \right\} \right),$$

(5)

where $\text{Cl}$ denotes the convex closure operation.
We prove this theorem in the Appendix A.

Fig. 4 shows an example for the capacity regions of two user systems corresponding to \( \mathcal{U}_2 = \{ u^{(1)}, u^{(2)} \} \) and \( \mathcal{U}_4 = \{ u^{(1)}, u^{(2)}, u^{(3)}, u^{(4)} \} \) for fixed power allocation \( P_1 \) and \( P_2 \). The capacity region is a union of capacity pentagons for fixed covariance matrices that are given by (4).

Remark \[17–19\] With a sum power constraint \( \sum_{k=1}^{K} P_k \leq P \), the capacity region is given by

\[
\mathcal{C}_{\text{union}}(P) = Cl \left( \bigcup_{\sum_{k=1}^{K} P_k \leq P} \mathcal{C}_{\text{MAC}}(P_1, \ldots, P_K) \right).
\]

using the duality between the MIMO MAC and MIMO BC (broadcasting channel).

4 Codebook Design Procedure Under a Sum Power Constraint

In this section, we give a method to design a codebook in a finite rate feedback environment to maximize the sum-rate.

The sum-rate can be represented as

\[
\sum_{k=1}^{K} R^{(k)} \leq E \left[ \log_2 \det \left( I + \sigma^{-2} \sum_{k=1}^{K} H^{(k)} Q^{(k)}(U) \left( H^{(k)} \right)^* \right) \right] \\
\leq E \left[ \log_2 \det \left( I + \sigma^{-2} \sum_{k=1}^{K} H^{(k)} Q^{(k)}(U) \left( H^{(k)} \right)^* \right) \right] \\
= \sum_{q=1}^{2^B} \max_{\mathcal{H}_q, \Tr(Q_q) \leq P} E \left[ \log_2 \det \left( I + \sigma^{-2} \sum_{k=1}^{K} H^{(k)} Q_q^{(k)} \left( H^{(k)} \right)^* \right) \right] \left[ H \in \mathcal{H}_q \right] P_q,
\]

where \( H = [H^{(1)} \cdots H^{(K)}] \), \( P_q = P(H \in \mathcal{H}_q) \), and \( \mathcal{H} = \{ \mathcal{H}_1, \ldots, \mathcal{H}_{2^B} \} \) is the partition on the channel matrix space \( \mathcal{H} \). We let \( Q_q^{(k)} = Q^{(k)}(u^{(q)}) \) and \( Q_q = \text{diag} \left( Q_q^{(1)}, \ldots, Q_q^{(K)} \right) \) to simplify the notation, where \( u^{(q)} \) is a realization of \( U \) and \( \text{diag}(\cdot, \ldots, \cdot) \) is a block diagonal operator. In this setup, the basestation can simply feed back the covariance index \( q \in \{1, 2, \ldots, 2^B\} \) to notify which covariance matrix \( Q_q \) should be used to maximize the achievable sum-rate. This means that we can convey this CSIT using a finite rate feedback channel that sends the same \( B \) bits of feedback to all users.

We apply Lloyd’s algorithm [20] to solve the maximization problem. Because the detailed procedure of Lloyd’s algorithm is similar to [20], we only emphasize the difference. In the process of designing the codebook using Lloyd’s algorithm, one of the steps is to solve the optimization problem

\[
\bar{Q}_q = \arg \max_{Q_q : \Tr(Q_q) \leq P} E \left[ \log_2 \det \left( I + \sigma^{-2} \sum_{k=1}^{K} H^{(k)} Q_q^{(k)} \left( H^{(k)} \right)^* \right) \right] \left[ H \in \mathcal{H}_q \right] \\
\Delta = \arg \max_{Q_q : \Tr(Q_q) \leq P} f(Q_q)
\]

\footnote{As described in [20], Lloyd’s algorithm converges to local optimum. As in [20], to obtain a point close to the globally optimal point, we run the same algorithm several times with different initial test channels and initial codebooks. Among the local optimum points, we choose the best point that gives the best capacity value.}
for a fixed partition $\mathcal{H} = \{\mathcal{H}_1, \ldots, \mathcal{H}_{2^n}\}$. We illustrate the approximate procedure of solving this optimization problem. By the heuristic approximation of [20],

$$
\begin{align*}
    f(Q_q) &= E \left[ \log_2 \det(I + \sigma^{-2} HQ_q H^*) \mid H \in \mathcal{H}_q \right] P_q \\
    &\approx \log_2 \det(I + \sigma^{-2} E[H^*H \mid H \in \mathcal{H}_q] Q_q) P_q \\
    &= \log_2 \det(I + \sigma^{-2} R_q Q_q) P_q \\
    &= \log_2 \det(I + \sigma^{-2} S_q Q_q S_q^*) P_q.
\end{align*}
$$

Here $R_q = E[H^*H \mid H \in \mathcal{H}_q] = S_q^*S_q$ with $S_q = [S_q^{(1)} \ldots S_q^{(K)}]$, where $S_q^{(k)} \in \mathbb{C}^{M_k \times M_t}$. Thus, the optimization problem becomes

$$
\bar{Q}_q = \arg\max_{Q_q : \text{Tr}(Q_q) \leq P} \log_2 \det \left( I + \sigma^{-2} \sum_{k=1}^K S_q^{(k)} Q_q^{(k)} (S_q^{(k)})^* \right) P_q.
$$

Therefore, the sum power iterative waterfilling algorithm [2] can be applied to this kind of optimization problem by treating the locally averaged $S_q^{(1)}, \ldots, S_q^{(K)}$ as the channel matrices. Solving for $q = 1, \ldots, 2^B$, we can obtain the codebook $Q = \{ \bar{Q}_1, \ldots, \bar{Q}_{2^B} \}$. Note that (13) is an approximation of the optimal solution of the original problem.

**Remark** In a typical MAC system, a suboptimal solution can be found using individual power constraints. Indeed, the sum-rate maximization problem under individual power constraints is a special case of the sum-rate maximization problem under the sum power constraint. We can also apply Lloyd’s algorithm to design a codebook under individual power constraints. An intermediate step in applying Lloyd’s algorithm is to solve the optimization problem

$$
\begin{align*}
    \left( \bar{Q}_q^{(1)}, \ldots, \bar{Q}_q^{(K)} \right) &= \arg\max_{\{Q_q^{(k)}\}_{k=1}^K : \text{Tr}(Q_q^{(k)}) \leq P_k} E \left[ \log_2 \det \left( I + \sigma^{-2} \sum_{k=1}^K H_q^{(k)} Q_q^{(k)} (H_q^{(k)})^* \right) \mid H \in \mathcal{H}_q \right].
\end{align*}
$$

The same procedure as the sum power constraint case gives

$$
\begin{align*}
    \left( \bar{Q}_q^{(1)}, \ldots, \bar{Q}_q^{(K)} \right) &= \arg\max_{\{Q_q^{(k)}\}_{k=1}^K : \text{Tr}(Q_q^{(k)}) \leq P_k} \log_2 \det \left( I + \sigma^{-2} \sum_{k=1}^K S_q^{(k)} Q_q^{(k)} (S_q^{(k)})^* \right) P_q,
\end{align*}
$$

and this problem can be solved by the individual power iterative waterfilling algorithm [1] with the effective channels $S_q^{(1)}, \ldots, S_q^{(K)}$.

### 5 Limited Feedback MAC Beamforming

In this section, we consider beamforming transmission for each user. We obtain the capacity region and derive codebook design methods for the limited feedback environment.
5.1 Capacity Analysis

In the beamforming transmission, the received signal is given by

\[ y_n = \sum_{k=1}^{K} H_n^{(k)} w_n^{(k)} x_n^{(k)} + z_n = \begin{bmatrix} H_n^{(1)} & H_n^{(2)} & \cdots & H_n^{(K)} \end{bmatrix} \begin{bmatrix} w_n^{(1)} \ x_n^{(1)} \\ w_n^{(2)} \ x_n^{(2)} \\ \vdots \\ w_n^{(K)} \ x_n^{(K)} \end{bmatrix} + z_n, n = 1, \ldots, N, \]  

where \( w_n^{(k)} \in \mathbb{C}^{M_k} \) is the beamforming vector of the \( k \)-th users. We assume that \( E [ |x_n^{(k)}|^2 ] = 1 \), which means that the transmission power of the system is given by \( \sum_{k=1}^{K} \| w_n^{(k)} \|^2 \).

In the limited feedback system with \( B \) bits of feedback, the codebook is given by \( W = \{ w_1, \ldots, w_{2^B} \} \), where \( w_q = \left( w_1^{(1)} \right)^T \cdots \left( w_1^{(K)} \right)^T \right)^T \). Based on the current feedback \( u(q) \), the beamforming vector is selected by \( w(u(q)) = w_q \) for some \( 1 \leq q \leq 2^B \). The capacity region in the limited feedback system is given in the following corollary.

**Corollary 1** For sum power constraint \( \sum_{k=1}^{K} P_k \leq P \) with \( \| w^{(k)}(u(q)) \|^2 \leq P_k \) for \( k = 1, \ldots, K \), the capacity region subject to a beamforming constraint is

\[
C_{\text{beamforming}}(P) = \text{Cl} \left( \bigcup_{k=1}^{K} \text{Cl} \left( \bigcup_{q=1}^{2^B} \text{Cl} \left( \bigcup_{s \in S} \text{Cl} \left( \bigcup_{k=1}^{K} \text{Cl} \left( \bigcup_{q=1}^{2^B} \text{Cl} \left( \bigcup_{s \in S} \text{Cl} \left( \bigcup_{k=1}^{K} \text{Cl} \left( \bigcup_{q=1}^{2^B} \sum_{k=1}^{K} P_k \leq P \right) \sum_{q=1}^{2^B} \max_{H_k, \| w_q \|^2 \leq P} E \left[ \log_2 \det \left( I + \sigma^{-2} \sum_{k=1}^{K} H_k^{(k)} w_q^{(k)} \left( w_q^{(k)} \right)^* \left( H_k^{(k)} \right)^* \right) | H \in \mathcal{H}_q \right] \right) \right) \right) \right) \right) \right) \right) \right) \right) \right).
\]

The proof immediately follows from Theorem 1.

5.2 Eigenbeamforming with Lloyd’s Algorithm

In this section, we derive the beamforming codebook design method using eigenbeamforming with Lloyd’s algorithm. As in Section 4, the sum-rate can be represented as

\[
\sum_{k=1}^{K} R^{(k)} \leq \sum_{q=1}^{2^B} \max_{H_k, \| w_q \|^2 \leq P} E \left[ \log_2 \det \left( I + \sigma^{-2} \sum_{k=1}^{K} H_k^{(k)} w_q^{(k)} \left( w_q^{(k)} \right)^* \left( H_k^{(k)} \right)^* \right) | H \in \mathcal{H}_q \right] \mathcal{P}_q.
\]

We also let \( w_q^{(k)} = w^{(k)}(u(q)) \) for the notational simplicity.

We can also apply Lloyd’s algorithm to solve this kind of problem [20]. Similar to Section 4, in the process of solving for the codebook using Lloyd’s algorithm, one of the steps is to solve

\[
\tilde{w}_q = \arg \max_{w_q : \| w_q \|^2 \leq P} E \left[ \log_2 \det \left( I + \sigma^{-2} \sum_{k=1}^{K} H_k^{(k)} w_q^{(k)} \left( w_q^{(k)} \right)^* \left( H_k^{(k)} \right)^* \right) | H \in \mathcal{H}_q \right] \mathcal{P}_q
\]

\[
\triangleq \arg \max_{w_q : \| w_q \|^2 \leq P} f(w_q).
\]
This is not a convex optimization problem and is therefore difficult to solve. Thus we approximate this problem by using a lower bound and employing the heuristic approximation from [20].

First, we claim that

\[
\log_2 \det \left( I + \sigma^{-2} \sum_{k=1}^{K} H^{(k)} w^{(k)}_q \left( w^{(k)}_q \right)^* \left( H^{(k)} \right)^* \right) \geq \log_2 \left( 1 + \frac{\sigma^{-2}}{K} \sum_{k=1}^{K} \left( w^{(k)}_q \right)^* \left( H^{(k)} \right)^* H^{(k)} w^{(k)}_q \right).
\]

This follows from the fact that

\[
\log_2 \det \left( I + \sigma^{-2} \sum_{k=1}^{K} H^{(k)} w^{(k)}_q \left( w^{(k)}_q \right)^* \left( H^{(k)} \right)^* \right) \geq \log_2 \det \left( I + \sigma^{-2} H^{(l)} w^{(l)}_q \left( w^{(l)}_q \right)^* \left( H^{(l)} \right)^* \right)
= \log_2 \left( 1 + \sigma^{-2} \left( w^{(l)}_q \right)^* \left( H^{(l)} \right)^* H^{(l)} w^{(l)}_q \right)
\]

for \(1 \leq l \leq K\) because \(\log_2 \det(\cdot)\) is monotonically increasing in the semi-definite cone. Because \(\log_2\) is monotonically increasing function, it is satisfied that

\[
\det \left( I + \sigma^{-2} \sum_{k=1}^{K} H^{(k)} w^{(k)}_q \left( w^{(k)}_q \right)^* \left( H^{(k)} \right)^* \right) \geq 1 + \sigma^{-2} \left( w^{(k)}_q \right)^* \left( H^{(k)} \right)^* H^{(k)} w^{(k)}_q.
\]

Adding \(21\) to both sides for \(1 \leq l \leq K\) and dividing by \(K\) results in

\[
\det \left( I + \sigma^{-2} \sum_{k=1}^{K} H^{(k)} w^{(k)}_q \left( w^{(k)}_q \right)^* \left( H^{(k)} \right)^* \right) \geq 1 + \sigma^{-2} \sum_{k=1}^{K} \left( w^{(k)}_q \right)^* \left( H^{(k)} \right)^* H^{(k)} w^{(k)}_q.
\]

Then by the heuristic approximation of [20],

\[
f(w_q) \geq E \left[ \log_2 \left( 1 + \frac{\sigma^{-2}}{K} \sum_{k=1}^{K} \left( w^{(k)}_q \right)^* \left( H^{(k)} \right)^* H^{(k)} w^{(k)}_q \right) \right] \mathcal{P}_q
\]

\[
= \log_2 \left( 1 + \frac{\sigma^{-2}}{K} \sum_{k=1}^{K} \left( w^{(k)}_q \right)^* \mathcal{P}_q \right)
\]

\[
= \log_2 \left( 1 + \frac{\sigma^{-2}}{K} \sum_{k=1}^{K} \left( R^{(k)}_q \right)^* \mathcal{P}_q \right)
\]

\[
\equiv g(w_q),
\]

where \(R^{(k)}_q = E \left[ \left( H^{(k)} \right)^* H^{(k)} | H \in \mathcal{H}_q \right].
\]

Instead of maximizing \(f(w_q)\), we utilize the lower bound \(g(w_q)\) because it is easier to obtain an analytical solution.

Let \(w^{(k)}_q = \alpha^{(k)}_q v^{(k)}_q, w^{(k)}_q \in \mathbb{C}^{M_q}\), where \(\alpha^{(k)}_q = \|w^{(k)}_q\|\) and \(\|v^{(k)}_q\| = 1\). Then

\[
g(w_q) = \log_2 \left( 1 + \frac{\sigma^{-2}}{K} \sum_{k=1}^{K} \left( w^{(k)}_q \right)^* R^{(k)}_q w^{(k)}_q \right)
\]

\[
= \log_2 \left( 1 + \frac{\sigma^{-2}}{K} \sum_{k=1}^{K} \left( \alpha^{(k)}_q \right)^2 \left( v^{(k)}_q \right)^* R^{(k)}_q v^{(k)}_q \right).
\]

We can select \(v^{(k)}_q\) as the unit-norm eigenvector associated with the maximum eigenvalue of \(R^{(k)}_q\). Then,

\[
g(w_q) = \log_2 \left( 1 + \frac{\sigma^{-2}}{K} \sum_{k=1}^{K} \left( \alpha^{(k)}_q \right)^2 \left( \lambda^{(k)}_q \right)^2 \right),
\]
where \( (\lambda_q^{(k)})^2 \) is the maximum eigenvalue of \( R_q^{(k)} \). By the Schwartz inequality,

\[
g(w_q) \leq \log_2 \left( 1 + \frac{\sigma^{-2}}{K} \left[ \sum_{k=1}^{K} \left( \alpha_q^{(k)} \right)^2 \right]^2 \sqrt{\sum_{k=1}^{K} \left( (\lambda_q^{(k)})^2 \right)^2} \right) .
\] (28)

Equality holds if and only if

\[
\left( \alpha_q^{(k)} \right)^2 = a \left( \lambda_q^{(k)} \right)^2
\] (29)

for some positive constant \( a \). Since \( \sum_{k=1}^{K} \left( \alpha_q^{(k)} \right)^2 = P \),

\[
a = \frac{P}{\sum_{k=1}^{K} \left( (\lambda_q^{(k)})^2 \right)} .
\] (30)

Therefore

\[
\alpha_q^{(k)} = \frac{\sqrt{P} \lambda_q^{(k)}}{\sqrt{\sum_{k=1}^{K} \left( (\lambda_q^{(k)})^2 \right)}}
\] (31)

and

\[
w_q^{(k)} = \frac{\sqrt{P} \lambda_q^{(k)}}{\sqrt{\sum_{k=1}^{K} \left( (\lambda_q^{(k)})^2 \right)}} v_q^{(k)} ,
\] (32)

where \( (\lambda_q^{(k)})^2 \) is the maximum eigenvalue of \( R_q^{(k)} \) and \( v_q^{(k)} \) is the corresponding unit-norm eigenvector.

The proposed method can additionally be validated by the papers [4, 5, 9], which state that when the covariance feedback is available at the transmitters the optimum beamforming vectors can be found as the eigenvector associated with the maximum eigenvalue of the covariance channel matrix of each user.

### 5.3 Grassmannian Design

In this section, we develop the Grassmannian codebook by solving the Grassmannian subspace packing problem [21] with random power allocation. References [22–25] dealt with a random beamforming codebook by employing random vector quantization (RVQ) in single user systems. This is based on the idea that the right singular vector of the channel which maximizes the capacity is random and isotropically distributed [24] under the assumption of an independent and identically distributed (i.i.d.) complex Gaussian channel matrix. In the same sense, in the MAC beamforming problem, each user’s channel is i.i.d., and therefore we can assume that the power allocation of each user is randomly distributed according to sum power constraint. In this correspondence, the beamforming direction of each user will be generated in the Grassmannian sense [21]. The codebook design procedure will be described in the following.
5.3.1 Codebook design

Let \( \mathcal{W} = \{\mathbf{w}_1, \ldots, \mathbf{w}_{2^B}\} \) be the beamforming codebook, where \( \mathbf{w}_q = \left[ (\mathbf{w}_q^{(1)})^T \cdots (\mathbf{w}_q^{(K)})^T \right]^T \). Here the sum power constraint is given by \( \|\mathbf{w}_q\|^2 \leq P \) for \( q = 1, \ldots, 2^B \). Then the sum-rate capacity is

\[
C_{\text{sum}}^{\text{beamforming}} = E_{\mathbf{H}, \mathcal{W}} \left[ \max_{\mathbf{w}_q \in \mathcal{W}} \log_2 \det \left( \mathbf{I} + \sigma^{-2} \sum_{k=1}^{K} \mathbf{H}_q^{(k)} \mathbf{w}_q^{(k)} \left( \mathbf{w}_q^{(k)} \right)^* \left( \mathbf{H}_q^{(k)} \right)^* \right) \right]
\]

\[
= E_{\mathbf{H}, \mathcal{W}} \left[ \max_{\mathbf{w}_q \in \mathcal{W}} \log_2 \det \left( \mathbf{I} + \sigma^{-2} \mathbf{H}_q \mathbf{W}_q^* \mathbf{H}_q^* \right) \right], \tag{33}
\]

where

\[
\mathbf{W}_q = \begin{bmatrix}
\mathbf{w}_q^{(1)} & \mathbf{0}_{M_t} & \cdots & \mathbf{0}_{M_t} \\
\mathbf{0}_{M_t} & \mathbf{w}_q^{(2)} & \cdots & \mathbf{0}_{M_t} \\
\vdots & \vdots & \ddots & \vdots \\
\mathbf{0}_{M_t} & \mathbf{0}_{M_t} & \cdots & \mathbf{w}_q^{(K)}
\end{bmatrix} \tag{34}
\]

is a beamforming matrix obtained by augmenting with each user’s beamforming vector with the zero column vector of length \( M_t \), \( \mathbf{0}_{M_t} \). We can see that \( (33) \) is the capacity formula for the single user MIMO system with the covariance matrix \( \mathbf{W}_q \mathbf{W}_q^* \) of the constraint that \( \mathbf{W}_q \) is block diagonal when \( \{\mathbf{W}_1, \ldots, \mathbf{W}_{2^B}\} \) is a codebook. Note that \( \mathbf{W}_q \in \mathbb{C}^{M_t K \times K} \) has orthogonal columns, but it is not unitary because \( \mathbf{W}_q^* \mathbf{W}_q \neq \alpha \mathbf{I} \), where \( \alpha \) is some non-negative constant chosen according to sum power constraint. In the unitary case of single user MIMO systems, references [22, 24] designed random codebooks by RVQ and reference [21] designed Grassmannian codebooks by Grassmannian subspace packing with the Fubini-Study distance. However, these codebook design methods are not directly applicable to our problem because our augmented beamforming matrix is not unitary.

In this section, we will assume that \( \mathbf{w}_q^{(k)} = \alpha_q^{(k)} \mathbf{v}_q^{(k)} \), where \( \alpha_q^{(k)} \) is the power allocation factor for the user \( k \) and \( \mathbf{v}_q^{(k)} \) is the normalized unit-norm beamforming vector for the user \( k \). Here we assume that \( \sum_{k=1}^{K} \left| \alpha_q^{(k)} \right|^2 = P \) for sum power constraint \( P \). The augmented beamforming matrix is given by

\[
\mathbf{W}_q = \mathbf{B}_q \mathbf{V}_q, \tag{35}
\]

where \( \mathbf{B}_q = \mathbf{A}_q \otimes \mathbf{I} \) with \( \mathbf{A}_q = \text{diag} \left( \alpha_1^{(1)}, \ldots, \alpha_q^{(K)} \right) \) and

\[
\mathbf{V}_q = \begin{bmatrix}
\mathbf{v}_q^{(1)} & \mathbf{0}_{M_t} & \cdots & \mathbf{0}_{M_t} \\
\mathbf{0}_{M_t} & \mathbf{v}_q^{(2)} & \cdots & \mathbf{0}_{M_t} \\
\vdots & \vdots & \ddots & \vdots \\
\mathbf{0}_{M_t} & \mathbf{0}_{M_t} & \cdots & \mathbf{v}_q^{(K)}
\end{bmatrix}. \tag{36}
\]
To maximize the sum-rate, it suffices to maximize the average determinant

$$C_{\text{sum}}^{\text{beamforming}} = E_{\mathcal{H},\mathcal{V}} \left[ \max_{\mathbf{w}_q \in \mathcal{W}} \det \left( \mathbf{I} + \sigma^{-2} \sum_{k=1}^{K} \mathbf{H}^{(k)} \mathbf{w}_q^{(k)} \left( \mathbf{w}_q^{(k)} \right)^* \left( \mathbf{H}^{(k)} \right)^* \right) \right]$$

$$= E_{\mathcal{H},\mathcal{V}} \left[ \max_{\mathbf{w}_q \in \mathcal{W}} \det \left( \mathbf{I} + \sigma^{-2} \mathbf{H}_q \mathbf{V}_q \mathbf{V}_q^* \mathbf{B}_q^* \mathbf{H}_q^* \right) \right]$$

$$= E_{\mathcal{H},\mathcal{V}} \left[ \max_{\mathbf{w}_q \in \mathcal{W}} \det \left( \mathbf{I} + \sigma^{-2} \mathbf{H}_q \mathbf{V}_q \mathbf{V}_q^* \mathbf{H}_q^* \right) \right], \quad (37)$$

where \( \sim \) means the equivalence of the optimization and \( \mathbf{H}_q = \mathbf{H} \). Thus, the codebook design problem can be viewed as the unitary codebook design of \( \mathcal{V} \sim \{ \mathbf{V}_1, \ldots, \mathbf{V}_{2^B} \} \) with the equivalent channel \( \mathbf{H}_q \) [21]. This problem can be solved by Grassmannian subspace packing using the Fubini-Study distance. The Fubini-Study distance is given by [21]

$$d_{FS}(\mathbf{V}_q, \mathbf{V}_r) = \arccos \left| \det(\mathbf{V}_q^* \mathbf{V}_r) \right|$$

$$= \arccos \left( \prod_{k=1}^{K} \left| \left( \mathbf{w}_q^{(k)} \right)^* \mathbf{v}_r^{(k)} \right| \right). \quad (38)$$

Let \( \delta_{FS}(\mathcal{V}) = \min_{1 \leq q < r \leq 2^B} d_{FS}(\mathbf{V}_q, \mathbf{V}_r) \). By [21], the codebook \( \mathcal{V} \) can be constructed by maximizing the minimum distance \( \delta_{FS}(\mathcal{V}) \). A detailed procedure to maximize \( \delta_{FS}(\mathcal{V}) \) is given in the Appendix B using Lloyd’s algorithm. \( \mathbf{B}_q \) can be randomly generated according to the sum power constraint, and the final beamforming vector can be obtained by \( \mathbf{w}_q^{(k)} = \alpha_q^{(k)} \mathbf{v}_q^{(k)}, \quad k = 1, \ldots, K, \quad q = 1, \ldots, 2^B \). The codebook can be constructed as \( \mathcal{W} = \{ \mathbf{w}_1, \ldots, \mathbf{w}_{2^B} \} \), where \( \mathbf{w}_q = \left[ \left( \mathbf{w}_q^{(1)} \right)^T \cdots \left( \mathbf{w}_q^{(K)} \right)^T \right]^T \).

### 5.3.2 Codebook Design in Correlated Fading

In this section, we discuss how the codebook design can be modified for a correlated fading channel. We assume that the receiver experiences an uncorrelated fading environment [4, 5]. Under a Kronecker correlated channel model, the general channel model of the user \( k \) is given by [4, 5, 26–28]

$$\mathbf{H}^{(k)} = \mathbf{H}_w^{(k)} \left( \mathbf{R}_t^{(k)} \right)^{1/2}, \quad (39)$$

where \( \mathbf{R}_t^{(k)} \) is covariance matrix of transmit antennas for user \( k \). The \( M_r \times M_t \) random matrix \( \mathbf{H}_w^{(k)} \) is i.i.d. circular symmetric complex Gaussian with zero-mean and unit-variance.

In this channel correlation environment, the statistical beamforming is introduced to achieve the sum-rate capacity [4, 5]. The statistical beamforming matrix under the assumption that equal power is allocated for each user is given by [4, 5]

$$\mathbf{V}_{stat} = \begin{bmatrix} \mathbf{v}_{stat}^{(1)} & 0_{M_t} & \cdots & 0_{M_t} \\ 0_{M_t} & \mathbf{v}_{stat}^{(2)} & \cdots & 0_{M_t} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{M_t} & 0_{M_t} & \cdots & \mathbf{v}_{stat}^{(K)} \end{bmatrix}, \quad (40)$$
where $\mathbf{v}_{\text{stat}}^{(k)}$ is the eigenvector associated with the maximum eigenvalue of $\mathbf{R}_t^{(k)}$. The statistical beamforming approach considered in [4, 5] has excellent performance in highly correlated channel, but suffers when correlation is moderate. We can improve the sum-rate performance of statistical beamforming in a limited feedback scenario by leveraging a modified codebook design [29]. This follows from the fact that the minimum distance of the codebook $\mathcal{V} = \{\mathbf{V}_1, \ldots, \mathbf{V}_{2^B}\}$ from Lloyd’s algorithm is the same as the the minimum distance of the codebook $\tilde{\mathcal{V}} = \{\mathbf{U} \mathbf{V}_1, \ldots, \mathbf{U} \mathbf{V}_{2^B}\}$, where $\mathbf{U}$ is an arbitrary unitary matrix. We can design our codebook such that $\mathbf{V}_{\text{stat}} = \mathbf{U} \mathbf{V}_1$ and the corresponding codebook is $\tilde{\mathcal{V}} = \{\mathbf{U} \mathbf{V}_1, \ldots, \mathbf{U} \mathbf{V}_{2^B}\}$.

In highly correlated channel environments, the proposed rotated codebook in the limited feedback environment has similar performance to the statistical beamforming. However, in moderate correlated channel environment, the proposed codebook has the better sum-rate performance than the statistical beamforming. We will see this increase of sum-rate in the numerical simulation section.

6 Numerical Simulation

In this section, we present numerical simulation results that show the sum-rate capacity performance of the proposed codebooks using Monte Carlo simulations. We call the multiple access system that has $K$ users, $M_t$ transmit antennas, and $M_r$ receive antennas a $(K, M_t, M_r)$ system.

Covariance Codebook

In this correspondence, we will compare the performance of the proposed system against a system with full CSI feedback and a system with no feedback. By no feedback, we mean that we allocate some scaled identity matrix to each data covariance matrix according to the power constraint. By full CSI feedback, we mean that we use the well-known iterative waterfilling [2] under a sum power constraint when the channel state information is known perfectly at all transmitters.

Fig. 5 shows the simulation result for a $(2, 2, 4)$ system under sum power constraint. The upper bound is the sum-rate performance when the users know the full CSI. The lower bound is the sum-rate performance when the users have no knowledge of the CSI. As we can see in this figure, as the number of feedback bits increases, the sum-rate capacity gain is also increased.

Beamforming Codebook

We show the performance of the eigenbeamforming with Lloyd’s algorithm and the Grassmannian codebook with random power allocation. For the Grassmannian packing, to compare the performance, we simulated the random codebook with random power allocation.
First, we show the performance of eigenbeamforming with Lloyd’s algorithm under sum power constraint. We simulated the system assuming each user receives 1 to 4 feedback bits to identify its beamforming vector. Fig. 6 shows the performance of the proposed eigenbeamforming of a (5,3,3) system. As the number of feedback bits increases, the sum-rate capacity also increases. Using 4 bits instead of 1 bit provides around a 1 dB gain. Compared to no feedback system, using 4 bits provides around a 2 dB SNR gain.

Fig. 7 shows the sum-rate performance of the Grassmannian codebook and the random codebook for a (2,2,3) system. From the figure, we can see that the two beamforming methods work well. Using 3 bits instead of 1 bit provides around a 1.5 dB SNR gain. We can also see that the sum-rate performance of the Grassmannian codebook is better than that of the random codebook. This is because the minimum distance of the Grassmannian codebook is larger than that of the random codebook. We can expect that the larger the minimum distance, the higher the obtained performance gain.

Fig. 8 shows the sum-rate performance of our Grassmannian codebook and statistical beamforming in correlated channel for a (2,2,3) system. We used the correlated channel model such that the eigenvalues of $R_t$ are made be 1.2 and 0.8 for each user. We can see that using 3 bit feedback information instead of statistical beamforming has an SNR advantage of around 1.5 dB. Note that the performance gain would be reduced if the transmitter spatial correlation increased.

Fig. 9 shows the relative performance comparison between the covariance codebook, the eigenbeamforming codebook and the Grassmannian beamforming codebook for a (2,2,4) system with 2 bit feedback. The performance of the full CSI is included as a reference. As expected, the covariance codebook has the best performance of all the codebook schemes and especially at high SNR the performance is close to the optimal iterative waterfilling solution which has the same slope of increasing according to SNR. Also for the beamforming schemes, as expected we do not have a multiplexing gain compared to the covariance codebook.

Fig. 10 shows the performance between the covariance codebook and time-division multiple access (TDMA) for a (2,2,4) system with 2 bit feedback. We can see that the proposed covariance codebook has similar performance compared to the TDMA scheme at low SNR. However, at high SNR, the performance gap becomes large. The second figure shows the ratio defined in [30, eq. (8)] between the covariance codebook and TDMA scheme. We can also verify the performance increases at high SNR. We can also observe the same performance trend as [30, Fig. 3]
7 Conclusion

In this correspondence, we proved a coding theorem for the MIMO MAC with finite cardinality feedback and proposed a codebook (a set of covariance matrices) design algorithm with the spatial waterfilling design using Lloyd’s algorithm. Compared to a no feedback system, a performance enhancement is observed even for a small number of feedback bits. Compared to full feedback, the proposed covariance codebook achieves comparable performance. We also derived the capacity region and two codebook design methods based on eigenbeamforming and Grassmannian subspace packing when we use beamforming for each user. Simulation results show that we can achieve more sum-rate capacity gains compared to the random beamforming codebook. Also for correlated channels, the proposed Grassmannian beamforming system has better sum-rate performance than the statistical beamforming codebook.

Appendix A: Proof of Theorem 1

We will prove this theorem in several steps.

Lemma 1 Let \( \mathbf{x}_n(S) = \left[ \left( \mathbf{x}_n^{(k_1)} \right)^T \cdots \left( \mathbf{x}_n^{(k_{|S|})} \right)^T \right]^T \) with \( S = \{k_1, \ldots, k_{|S|} \} \subseteq \{1, \ldots, K\} \). \( \mathbf{x}_n(S^C) \) can be defined in the same way with \( S^C \). Then

\[
I\left( \mathbf{x}_n(S); \mathbf{y}_n|V_n, \mathbf{x}_n(S^C) \right) \leq E_{V_n} \left[ \log_2 \det \left( I + \sigma^{-2} \sum_{k \in S} \mathbf{H}_n^{(k)} \mathbf{Q}(U_n) \left( \mathbf{H}_n^{(k)} \right)^* \right) \right] \equiv C(S). \tag{41}
\]

Proof

\[
I\left( \mathbf{x}_n(S); \mathbf{y}_n|V_n, \mathbf{x}_n(S^C) \right) = h\left( \mathbf{y}_n|V_n, \mathbf{x}_n(S^C) \right) - h\left( \mathbf{y}_n|\mathbf{x}_n(S), V_n, \mathbf{x}_n(S^C) \right)
= h\left( \mathbf{y}_n|V_n, \mathbf{x}_n(S^C) \right) - h\left( \mathbf{z}_n \right)
\leq E_{V_n} \left[ \log_2 \det \left( I + \sigma^{-2} E\left[ \mathbf{H}_n(S) \mathbf{x}_n(S) \mathbf{x}_n^*(S) \mathbf{H}_n^H(S)|V_n) \right] \right) \right] \tag{42}
= E_{V_n} \left[ \log_2 \det \left( I + \sigma^{-2} \sum_{k \in S} \mathbf{H}_n^{(k)} E\left[ \mathbf{x}_n^{(k)} \mathbf{x}_n^{(k)*}|U_n \right] \left( \mathbf{H}_n^{(k)*} \right) \right) \right]
= E_{V_n} \left[ \log_2 \det \left( I + \sigma^{-2} \sum_{k \in S} \mathbf{H}_n^{(k)} \mathbf{Q}(U_n) \left( \mathbf{H}_n^{(k)*} \right) \right) \right], \tag{43}
\]

where \( \mathbf{H}_n(S) = \left[ \mathbf{H}_n^{(k_1)} \cdots \mathbf{H}_n^{(k_{|S|})} \right] \) and \( E\left[ \mathbf{x}_n^{(k)} \mathbf{x}_n^{(k)*}|U_n \right] = \mathbf{Q}^{(k)}(U_n) \). In (42), we use the maximum entropy property because the covariance matrix of \( \mathbf{y}_n \) given \( V_n \) and \( \mathbf{x}_n(S^C) \) is \( E\left[ \mathbf{H}_n(S) \mathbf{x}_n(S) \mathbf{x}_n^*(S) \mathbf{H}_n^H(S)|V_n \right] + \sigma^2 I \) [16].

Lemma 2 Let \( \mathbf{T} = \left[ \left( \mathbf{T}^{(1)} \right)^T \cdots \left( \mathbf{T}^{(K)} \right)^T \right]^T \) and \( \mathbf{T}(S) = \left[ \left( \mathbf{T}^{(k_1)} \right)^T \cdots \left( \mathbf{T}^{(k_{|S|})} \right)^T \right]^T \). \( \mathbf{T}(S^C) \) is defined in the same way. Then

\[
I\left( \mathbf{T}(S); \mathbf{T}(S^C)|\mathbf{y}, V \right) = I\left( \mathbf{x}(S); \mathbf{x}(S^C)|\mathbf{y}, V \right). \tag{44}
\]
Proof We know that \( \mathbf{T}^{(k)} = [\mathbf{t}_1^{(k)} \cdots \mathbf{t}_{|U|}^{(k)}] \), where \( \mathbf{t}_i^{(k)} \in \mathcal{X}^{M_i} \). \( \mathbf{t}_j(S) \) is defined as \( \mathbf{t}_j(S) = \left[ (\mathbf{t}_j^{(k)})^T \cdots (\mathbf{t}_j^{(k)})^T \right]^T \), i.e. the \( j \)-th column of \( \mathbf{T}(S) \). \( \mathbf{t}_j(S^C) \) is defined by the same way. We define a function \( f: U \to \{1, \ldots, |U|\} \), where \( f(U) \) is a bijection. We also define \( \mathbf{T}(S, U) = \mathbf{t}_{f(U)}(S) \) and

\[
\mathbf{T}(S, U^C) = [\mathbf{t}_1(S) \cdots \mathbf{t}_{f(U)-1}(S) \mathbf{t}_{f(U)+1}(S) \cdots \mathbf{t}_{|U|}(S)].
\] (45)

\( \mathbf{T}(S^C, U) \) and \( \mathbf{T}(S^C, U^C) \) can be defined in the same way.

Then

\[
I(\mathbf{T}(S); \mathbf{T}(S^C) | \mathbf{y}, \mathbf{V}) = I(\mathbf{T}(S, U^C), \mathbf{T}(S, U); \mathbf{T}(S^C, U^C), \mathbf{T}(S^C, U) | \mathbf{y}, \mathbf{V})
\]
\[= I(\mathbf{T}(S^C, U); \mathbf{T}(S, U) | \mathbf{y}, \mathbf{V})
\]
\[+ I(\mathbf{T}(S^C, U); \mathbf{T}(S, U^C) | \mathbf{T}(S, U), \mathbf{y}, \mathbf{V})
\]
\[+ I(\mathbf{T}(S^C, U^C); \mathbf{T}(S, U) | \mathbf{T}(S^C, U), \mathbf{y}, \mathbf{V})
\]
\[+ I(\mathbf{T}(S^C, U^C); \mathbf{T}(S^C, U) | \mathbf{T}(S^C, U), \mathbf{T}(S, U), \mathbf{y}, \mathbf{V})
\]
\[= I(\mathbf{T}(S^C, U); \mathbf{T}(S, U) | \mathbf{y}, \mathbf{V})
\]
\[= I(\mathbf{x}(S); \mathbf{x}(S^C) | \mathbf{y}, \mathbf{V}).
\] (46) (47) (48)

Eqn. (47) follows from independence of data for the set \( S \) and \( S^C \) according to \( U \) and \( U^C \). Eqn. (48) follows from \( \mathbf{x}(S) = \mathbf{T}(S, U) \) and \( \mathbf{x}(S^C) = \mathbf{T}(S^C, U) \). Here, we used \( I(A; B; C, D) = I(D; B) + I(D; A|B) + I(C; B|D) + I(C; A|D, B) \). This can easily be seen by the chain rule.

Now we prove Theorem 1.

Proof We first prove (4).

Achievability: We follow the procedure of [31] with the equivalent channel in Fig. 3.

Codebook generation: Fix \( p(\mathbf{T}^{(1)}, \ldots, \mathbf{T}^{(K)}) = \prod_{k=1}^K p(\mathbf{T}^{(k)}) \). We generate \( 2^{N_R(k)} \) independent codewords

\[
\{\mathbf{T}^{(k)}(\omega^{(k)})\}_{i=1}^N, \omega^{(k)} \in \{1, 2, \ldots, 2^{N_R(k)}\}
\]

with respect to \( \prod_{n=1}^N p(\mathbf{T}^{(n)}(\omega^{(n)}) \), where

\[
\left\{\mathbf{T}^{(k)}(\omega^{(k)})\right\}_{i=1}^N = \left\{\mathbf{T}^{(k)}_1(\omega^{(k)}), \ldots, \mathbf{T}^{(k)}_N(\omega^{(k)})\right\}.
\]

(49)

These codewords form the codebook, which is revealed to the transmitters and the receiver.

Encoding: To send index \( \omega^{(k)} \), sender \( k \) sends the codeword \( \mathbf{T}^{(k)}_1(\omega^{(k)}), \ldots, \mathbf{T}^{(k)}_N(\omega^{(k)}) \) through the equivalent channel \( p(y^N_1, v^N_1 | \mathbf{T}^{(k)}_1) \).
Decoding: Let $A^{(N)}_t$ denote the set of typical $\left(\{T^{(1)}\}_{1}^{N}, \ldots, \{T^{(K)}\}_{1}^{N}, y_{1}^{N}, v_{1}^{N}\)$ sequences. The properties of $A^{(N)}_t$ are given in [31]. The receiver chooses the $K$-tuple $\{\omega^{(1)}, \ldots, \omega^{(K)}\}$ such that

$$\left(\left\{T^{(1)}(\omega^{(1)})\right\}_{1}^{N}, \ldots, \left\{T^{(K)}(\omega^{(K)})\right\}_{1}^{N}, y_{1}^{N}, v_{1}^{N}\right) \in A^{(N)}_t$$

if such a $K$-tuple $\{\omega^{(1)}, \ldots, \omega^{(K)}\}$ exists and is unique; otherwise, an error is declared.

Analysis of the probability error: We can assume that $\{\omega^{(1)}, \ldots, \omega^{(K)}\} = (1, \ldots, 1)$. We have an error if either the correct codewords are not typical with the received sequence or there is a pair of incorrect codewords that are typical with the received sequence. Define the events

$$E_{\omega^{(1)}, \ldots, \omega^{(K)}} = \left\{\left(\left\{T^{(1)}(\omega^{(1)})\right\}_{1}^{N}, \ldots, \left\{T^{(K)}(\omega^{(K)})\right\}_{1}^{N}, y_{1}^{N}, v_{1}^{N}\right) \in A^{(N)}_t\right\}$$

and

$$E_{0}^{(k)} = \left\{\left\{T^{(k)}(1)\right\}_{1}^{N} : \left|\operatorname{Tr}\left(\frac{1}{N} \sum_{n=1}^{N} t^{(k)}_{f(U_{n})(1)} \left(t^{(k)}_{f(U_{n})(1)}\right)^{*}\right) - P_{k}\right| > \epsilon_{k}\right\}$$

for $U_{n} \in \mathcal{U}$. Then

$$P_{\epsilon}^{(N)} = P\left(\bigcup_{k=1}^{K} E_{0}^{(k)} \bigcup E_{\omega^{(1)}, \ldots, \omega^{(K)}} \bigcup \left(\bigcup_{(\omega^{(1)}, \ldots, \omega^{(K)}) \neq (1, \ldots, 1)} E_{\omega^{(1)}, \ldots, \omega^{(K)}}\right)\right)$$

$$\leq \sum_{k=1}^{K} P\left(E_{0}^{(k)}\right) + P\left(E_{\omega^{(1)}, \ldots, \omega^{(K)}}\right) + \sum_{S \subseteq \{1, \ldots, K\}} P\left(E_{\omega(S) \neq (1, \ldots, 1), \omega(S') = (1, \ldots, 1)}\right),$$

where $\omega(S) = \{\omega^{(k_{1}), \ldots, \omega^{(k_{|S|})}}\}$. We can prove that $P\left(E_{\omega^{(1)}, \ldots, \omega^{(K)}}\right) \leq \epsilon$ by the property of $\epsilon$-typical sequence. We can also prove that $P\left(E_{0}^{(k)}\right) \leq \epsilon$ by the law of large numbers. If we define

$$T^{N}_{1}\left(S, \omega^{(k_{1})}, \ldots, \omega^{(k_{|S|})}\right) = \left\{\left\{T^{(k_{1})}(\omega^{(k_{1})})\right\}_{1}^{N}, \ldots, \left\{T^{(k_{|S|})}(\omega^{(k_{|S|})})\right\}_{1}^{N}\right\},$$

then

$$P\left(E_{\omega(S) = (\omega^{(k_{1}), \ldots, \omega^{(k_{|S|})})}, \omega(S') = (1, \ldots, 1)}\right)$$

$$= P\left(T^{N}_{1}\left(S, \omega^{(k_{1})}, \ldots, \omega^{(k_{|S|})}\right), T^{N}_{1}\left(S^{C}, 1, \ldots, 1\right), y_{1}^{N}, v_{1}^{N}\right) \in A^{(N)}_t\right)$$

$$= \sum_{\left\{T^{(1)}\right\}_{1}^{N} \ldots \left\{T^{(K)}\right\}_{1}^{N}, y_{1}^{N}, v_{1}^{N} \in A^{(N)}_t} p\left(T^{N}_{1}\left(S\right)\right) p\left(T^{N}_{1}\left(S^{C}\right), y_{1}^{N}, v_{1}^{N}\right)$$

$$\leq |A^{(N)}_t| 2^{-N(h(T(S)) - \epsilon)} 2^{-N(h(T(S^{C})) y, V) - \epsilon})$$

$$\leq 2^{N(h(T(S), T(S^{C}), y, V) + \epsilon) - h(T(S), T(S^{C}), y, V) - \epsilon}) 2^{-N(h(T(S^{C})) y, V) - \epsilon})$$

$$= 2^{-N(h(T(S)) + h(T(S^{C})) y, V) - h(T(S), T(S^{C}), y, V) - 3\epsilon})$$

$$= 2^{-N(I(T(S); T(S^{C}), y, V) - 3\epsilon})}$$

$$= 2^{-N(I(T(S); y(T(S^{C}), V) - 3\epsilon),}$$

(55)
where \( T^N_i(S) = \{ T_i(S), \ldots, T_N(S) \} \). The last equality follows from
\[
I(T(S) ; T(SC), y, V) = I(T(S) ; T(SC) , V) + I(T(S) ; y|T(SC) , V) \tag{56}
\]
and \( I(T(S) ; T(SC), V) = 0 \) since \( T(S) \) and \( T(SC) \) are independent.

Therefore
\[
P_e^{(N)} \leq (K + 1) \epsilon + \sum_{S \subseteq \{1, \ldots, K \}} 2^{N(R(S) - I(T(S) ; y|T(SC), V) + 3\epsilon)}, \tag{57}
\]
where \( R(S) = \sum_{k \in S} R^{(k)} \). Therefore, if \( R(S) < I(T(S) ; y|T(SC), V) \), \( P_e^{(N)} \to 0 \) as \( N \to \infty \).

To complete the proof, it suffices to show that
\[
I(T(S) ; y|T(SC), V) \leq I(x(S) ; y|x(SC), V). \tag{58}
\]
This is because
\[
I(T(S) ; y|T(SC), V) = I(T(S) ; y|V) - I(T(S) ; T(SC) | V) + I(T(S) ; T(SC) | y, V)
\]
\[
= I(T(S) ; y|V) + I(T(S) ; T(SC) | y, V)
\]
\[
\leq I(x(S) ; y|V) + I(x(S) ; x(SC) | y, V) \tag{59}
\]
\[
= I(x(S) ; y|V) - I(x(S) ; x(SC) | V) + I(x(S) ; x(SC) | y, V)
\]
\[
= I(x(S) ; y|x(SC), V). \tag{60}
\]
We also used the fact that
\[
I(A; B|C, D) = I(A; B|D) - I(A; C|D) + I(A; C|B, D) \tag{62}
\]
and \( I(T(S) ; T(SC) | V) = 0 \) since \( T(S) \) and \( T(SC) \) are independent. Eqn. \((59)\) follows from the data processing inequality. We proved that \( I(T(S) ; T(SC) | y, V) = I(x(S) ; x(SC) | y, V) \) by Lemma 2.

Therefore \( I(T(S) ; y|T(SC), V) \leq I(x(S) ; y|x(SC), V) \) and the theorem follows.

Converse Proof: We now prove the converse theorem.

Let \( W^{(k)} \in \{ 1, \ldots, 2^{NR(S)} \} \) be a random variable, where \( k \in \{ 1, \ldots, K \} \). Then,
\[
NR(S) = H(W(S))
\]
\[
= H(W(S)|y^N, V_1^N) + I(W(S) ; y^N_1, V_1^N)
\]
\[
\leq H(W^{(1)}, \ldots, W^{(K)}|y^N_1, V_1^N) + I(W(S) ; y^N_1, V_1^N)
\]
\[
\leq 1 + NR_P^{(N)} + I(W(S) ; y^N_1, V_1^N) \tag{63}
\]
\[
= 1 + NR_P^{(N)} + I(W(S) ; y^N_1|V_1^N). \tag{64}
\]
where \( R = \sum_{k=1}^{K} R^{(k)} \) and \( W(S) = \{ W^{(k_1)}, \ldots, W^{(k_S)} \} \). Eqn. (63) follows from Fano’s inequality and (64) follows from the independence between \( W(S) \) and \( V_n \).

\[
I (W(S); y_1^N | V_1^N) = \sum_{n=1}^{N} I (W(S); y_n | V_1^N, y_1^{n-1})
\]

Eqn. (65) follows from the chain rule, and (66) follows from the data processing inequality. Eqn. (67) follows from the fact that \( T_n(S) \) is independent of the other variables except for \( y_n \), and (68) follows from the fact that the condition does not increase entropy and \( T_n(S) \) and \( T_n(S^C) \) are independent.

We showed that \( I (T_n(S); y_n | T_n(S^C), V_n) \leq I (x_n(S); y_n | T_n(S^C), V_n) \), and therefore,

\[
R(S) \leq \frac{1}{N} + RP_c^{(N)} + C(S).
\]

If \( P_c^{(N)} \to 0 \) as \( N \to \infty \), then \( R(S) \leq C(S) \).

Now we prove (5). By [28], the capacity of any MAC can be written as the convex closure of the union of rate regions corresponding to the input product distribution satisfying the individual power constraint. Therefore to obtain the capacity region of our MIMO MAC, we can unionize all capacity polyhedrons about all covariance matrices satisfying the power constraint and take the convex closure. For a fixed set of covariance matrices, the capacity polyhedron is given by (4). The capacity region must be closed by Theorem 3 in [32]. We can show that this fixed capacity polyhedron is closed and that the union of these polyhedrons is also closed. This completes the proof.

Appendix B: Lloyd’s Algorithm

We give Lloyd’s algorithm to find the codebook that maximizes the minimum distance. The Grassmannian line packing problem approximately solved using Lloyd’s algorithm for single user MIMO beamforming systems has been
studied in [33, 34]. Based on the \textit{Fubini-Study distance}, we define the distortion measure

\[ d_1 (G, V_q) = \arccos(|\det (G^* V_q)|), \]  

(70)

where

\[ G = \begin{bmatrix} g^{(1)} & \cdots & 0_{M_t} \\ \vdots & \ddots & \vdots \\ 0_{M_t} & \cdots & g^{(K)} \end{bmatrix}, \]  

(71)

and each \( g^{(k)} \in \mathbb{C}^{M_t} \) is a unit-norm vector. To find an optimal codebook, we should solve the following optimization problem

\[ V = \arg\min_{V'} E_G \left[d_1 (G, V(G))\right], \]  

(72)

where \( V(G) = \arg\min_{1 \leq q \leq 2^B} d_1 (G, V_q), \) \( V_q \in V' \), and \( V' \) is a codebook of cardinality \( 2^B \). The optimization problem of (72) is difficult to solve because of the \( \arccos \) function. Instead we define a modified distortion measure

\[ d_2 (G, V_q) = -\prod_{k=1}^{K} (v_q^{(k)})^* g^{(k)} (g^{(k)})^* v_q^{(k)}, \]  

(73)

which is mathematically tractable to solve analytically. This is because (72) becomes

\[ V = \arg\min_{V'} E_G \left[\arccos \left(|\det (G^* V(G))|\right)\right] \]  

\[ = \arg\min_{V'} E_G \left[-|\det (G^* V(G))|\right] \]  

\[ = \arg\min_{V'} E_G \left[-|\det (G^* V(G))|^2\right] \]  

\[ = \arg\min_{V'} E_G \left[-\prod_{k=1}^{K} (v_q^{(k)})^* g^{(k)} (g^{(k)})^* v_q^{(k)}\right], \]  

(74)

(75)

where \( v_q^{(k)} \) is the \( k \)-th column of \( V(G) \). (74) follows from the fact that \( 0 \leq |\det (G^* V(G))| \leq 1 \) and \( \arccos(\cdot) \) is a monotonically decreasing function on \([0, 1] \). Optimization of (75) is also easier than the optimization of (72) to obtain the analytical solution.

To minimize the average distortion, we apply Lloyd’s algorithm. An intermediate step in applying Lloyd’s algorithm is to solve the optimization problem

\[ V = \arg\min_{V'} E_G \left[d_2 (G, V)\right] \]  

\[ = \arg\min_{V'} E_G \left[-\prod_{k=1}^{K} (v_q^{(k)})^* g^{(k)} (g^{(k)})^* v_q^{(k)}\right] \]  

\[ = \arg\min_{V'} \sum_{q=1}^{2^B} \left[-\prod_{k=1}^{K} (v_q^{(k)})^* E \left[g^{(k)} (g^{(k)})^* G \in G_q \right] v_q^{(k)}\right] P(G \in G_q) \]  

\[ = \arg\min_{V'} \sum_{q=1}^{2^B} \left[-\prod_{k=1}^{K} (v_q^{(k)})^* R_q^{(k)} v_q^{(k)}\right] P(G \in G_q), \]  

(76)
where $\mathcal{G} = \{ \mathcal{G}_1, \ldots, \mathcal{G}_{2^n} \}$ is a specific Voronoi regions and $\mathbf{R}_q^{(k)} = E \left[ \mathbf{g}_q^{(k)} (\mathbf{g}_q^{(k)})^* \mid \mathbf{G} \in \mathcal{G}_q \right]$. To minimize the distortion measure, $\mathbf{v}_q^{(k)}$ can be chosen by the eigenvector associated with the maximum eigenvalue of $\mathbf{R}_q^{(k)}$. Therefore we can obtain $\mathbf{V} = \{ \mathbf{V}_1, \ldots, \mathbf{V}_{2^n} \}$.

In general, the average distortion converges, but the minimum distance $\delta_{FS}$ does not converge as iteration proceeds. $\delta_{FS}$ increases during the first few step, and has tendency to increase as iteration proceeds but may oscillate slightly after some iteration [33]. In this case, as indicated in [33], we save all intermediate minimum distances during the iterations and at some fixed iteration number, we obtain the maximum of the minimum distances from the previous saved minimum distances. We also obtain the codebook that corresponds to the index that maximizes the minimum distance.

**References**

[1] W. Yu, W. Rhee, S. Boyd, and J. Cioffi, “Iterative water-filling for Gaussian vector multiple-access channels,” IEEE Trans. Info. Th., vol. 50, no. 1, pp. 145–152, Jan. 2004.

[2] N. Jindal, W. Rhee, S. Vishwanath, S. Jafar, and A. Goldsmith, “Sum power iterative water-filling for multi-antenna Gaussian broadcast channels,” IEEE Trans. Info. Th., vol. 51, no. 4, pp. 1570–1580, Apr. 2005.

[3] E. Jorswieck and H. Boche, “Transmission strategies for the MIMO MAC with MMSE receiver: average MSE optimization and achievable individual MSE region,” IEEE Trans. Sig. Proc., vol. 51, no. 11, pp. 2872–2881, Nov. 2003.

[4] A. Soysal and S. Ulukus, “Optimality of beamforming in fading multi-user MIMO multiple access channels,” IEEE Trans. Commun., submitted for review.

[5] ———, “Asymptotic optimality of beamforming in multi-user MIMO-MAC with no or partial CSI at the transmitters,” in Proc. IEEE Veh. Technol. Conf., 2005-Spring, pp. 1619–1623.

[6] E. Visotsky and U. Madhow, “Space-time precoding with imperfect feedback,” IEEE Trans. Info. Th., vol. 47, pp. 2632–2639, Sept. 2001.

[7] S. Jafar and A. Goldsmith, “Multiple-antenna capacity in correlated rayleigh fading with channel covariance information,” IEEE Trans. Wireless Comm., vol. 4, no. 3, pp. 990–997, May 2005.

[8] ———, “Transmitter optimization and optimality of beamforming for multiple antenna systems,” IEEE Trans. Wireless Comm., vol. 3, pp. 1165–1175, July 2004.

[9] E. A. Jorswieck and H. Boche, “Channel capacity and capacity-range of beamforming in MIMO wireless systems under correlated fading with covariance feedback,” IEEE Trans. Wireless Comm., vol. 3, no. 5, pp. 1543–1553, Sept. 2004.

[10] H. Boche and E. Jorswieck, “On the optimality range of beamforming for MIMO systems with covariance feedback,” IEICE Trans. Commun., vol. E85-A, pp. 2521–2528, Nov. 2002.

[11] W. Rhee, W. Yu, and J. M. Cioffi, “The optimality of beamforming in uplink multiuser wireless systems,” IEEE Trans. Wireless Comm., vol. 3, pp. 86–96, Jan. 2004.

[12] C. Shannon, “Channels with side information at the transmitter,” IBM J. Res. Develop., vol. 2, pp. 289–293, 1958.

[13] M. Salehi, “Capacity and coding for memories with real-time noisy defect information at encoder and decoder,” in IEE Proc. Commun., Speech and Vision, vol. 139, Apr. 1992, pp. 113–117.

[14] G. Caire and S. Shamai, “On the capacity of some channels with channel state information,” IEEE Trans. Info. Th., vol. 45, no. 6, pp. 2007–2019, July 1999.

[15] V. Lau, Y. Liu, and T. Chen, “Capacity of memoryless channels and block-fading channels with designable cardinality-constrained channel state feedback,” IEEE Trans. Info. Th., vol. 50, no. 9, pp. 2038–2049, Sept. 2004.

[16] M. Skoglund and G. Jöngren, “On the capacity of a multiple-antenna communication link with channel side information,” IEEE Jour. Select. Areas in Commun., vol. 21, no. 3, pp. 395–405, Apr. 2003.

[17] S. Vishwanath, N. Jindal, and A. Goldsmith, “Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels,” IEEE Trans. Info. Th., vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[18] N. Jindal, S. Vishwanath, and A. Goldsmith, “On the duality of Gaussian multiple-access and broadcast channels,” IEEE Trans. Info. Th., vol. 50, no. 5, pp. 768–783, May 2004.

[19] ———, “On the duality of multiple-access and broadcast channels,” in Proc. IEEE Int. Symp. Info. Th., Jun. 30-Jul. 5 2002, p. 500.

[20] V. Lau, Y. Liu, and T. Chen, “On the design of MIMO block-fading channels with feedback-link cardinality constraint,” IEEE Trans. Commun., vol. 52, no. 1, pp. 62–70, Jan. 2004.

[21] D. J. Love and R. W. Heath Jr., “Limited feedback unitary precoding for spatial multiplexing systems,” IEEE Trans. Info. Th., vol. 51, pp. 2967–2976, Aug. 2005.
[22] W. Santipach and M. L. Honig, “Asymptotic performance of MIMO wireless channels with limited feedback,” in *Proc. IEEE Mil. Comm. Conf.*, vol. 1, Oct. 2003, pp. 141–146.

[23] ——, “Asymptotic capacity of beamforming with limited feedback,” in *Proc. IEEE Int. Symp. Info. Th.*, Jun. 27-Jul. 2 2004, p. 289.

[24] ——, “Achievable rates for MIMO fading channels with limited feedback and linear receivers,” in *Proc. IEEE Int. Symp. Spread Spectrum Techniques and Applications*, Aug. 30-Sept. 2 2004, pp. 1–6.

[25] C. K. Au-Yeung and D. J. Love, “On the performance of random vector quantization limited feedback beamforming in a MISO system,” *IEEE Trans. Wireless Comm.*, accepted for publication.

[26] G. J. Foschini and M. J. Gans, “On limits of wireless communications in a fading environment when using multiple antennas,” *Wireless Personal Communications* 6, pp. 311–335, 1998.

[27] D. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, “Fading correlation and its effect on the capacity of multielement antenna systems,” *IEEE Trans. Commun.*, vol. 48, pp. 502–513, Mar. 2000.

[28] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath, “Capacity limits of MIMO channels,” *IEEE Jour. Select. Areas in Commun.*, vol. 21, no. 5, pp. 684–702, June 2003.

[29] D. J. Love and R. W. Heath Jr., “Limited feedback diversity techniques for correlated channels,” *IEEE Trans. on Veh. Technol.*, vol. 55, no. 2, pp. 718–722, Mar. 2006.

[30] N. Jindal and A. Goldsmith, “Dirty-paper coding versus TDMA for MIMO broadcast channels,” *IEEE Trans. Info. Th.*, vol. 51, no. 5, pp. 1783–1794, May 2005.

[31] T. Cover and J. Thomas, *Elements of Information Theory*. New York, NY: Wiley, 1991.

[32] S. Verdú, “Multiple-access channels with memory with and without frame synchronism,” *IEEE Trans. Info. Th.*, vol. 35, no. 3, pp. 605–619, May 1989.

[33] P. Xia, S. Zhou, and G. B. Giannakis, “Achieving the Welch bound with difference sets,” *IEEE Trans. Info. Th.*, vol. 51, no. 5, pp. 1900–1907, May 2005.

[34] P. Xia and G. B. Giannakis, “Design and analysis of transmit-beamforming based on limited-rate feedback,” *IEEE Trans. Sig. Proc.*, vol. 54, no. 5, pp. 1853–1863, May 2006.
Figure 1: MIMO MAC model with CSIT and CSIR.

Figure 2: Channel model of the communication system with CSIT and CSIR.

Figure 3: Equivalent channel model of the communication system with CSIT and CSIR.
Figure 4: Capacity region for MIMO MAC according to the cardinality of $\mathcal{U}$.

Figure 5: Sum-rate capacity of a (2,2,4) system using a covariance codebook under a sum power constraint.
Figure 6: Sum-rate capacity of a (5,3,3) system using an eigenbeamforming codebook under a sum power constraint.

Figure 7: Sum-rate capacity of a (2,2,3) system using a Grassmannian beamforming codebook and random beamforming codebook with random power allocation.
Figure 8: Sum-rate capacity of (2,2,3) systems using Grassmannian beamforming and statistical beamforming.

Figure 9: Sum-rate of (2,2,4) systems for the covariance codebook, the eigenbeamforming codebook, and the Grassmannian beamforming codebook.
Figure 10: Sum-rate of (2,2,4) systems for the covariance codebook and TDMA scheme, and their comparison.