

# Asymptotic Optimality of Beamforming in Multi-user MIMO-MAC with No or Partial CSI at the Transmitters

Alkan Soysal      Sennur Ulukus  
Department of Electrical and Computer Engineering  
University of Maryland, College Park, MD 20742  
alkan@eng.umd.edu      ulukus@eng.umd.edu

**Abstract**—We consider a multi-input multi-output (MIMO) multiple access channel (MAC) where the receiver has the perfect channel state information (CSI), while the transmitters have either no or partial CSI. When the transmitters have the partial CSI, it is in the form of either the covariance matrix of the channel or the mean matrix of the channel. First, we show that the optimum transmit directions of each user are the eigenvectors of its own channel covariance and mean feedback matrices, in the covariance and mean feedback models, respectively. Then, we show that in both no and partial CSI cases, beamforming is asymptotically optimal. In particular, in the case of no CSI, we show that a simple form of beamforming, which may be characterized as an arbitrary antenna selection scheme, achieves the optimum sum capacity. In the case of partial CSI, we show that beamforming in the direction of the strongest eigenvector of the channel feedback matrix achieves the optimum sum capacity.

## I. INTRODUCTION

The use of multiple antennas at both transmitters and receivers in wireless communications promises very large information rates. In [1], Telatar showed that in a single-user system, when the transmitter does not know the state of the fading channel, i.e., when the channel entries are assumed to be i.i.d., zero-mean Gaussian random variables, the optimum input covariance matrix is proportional to the identity matrix, which is full-rank. In order to achieve the capacity, either vector coding or parallel processing of scalar codes is needed. As stated in [1], vector coding will result in lower probability of error but higher complexity as compared to parallel-scalar coding, which already is very complex [2].

Beamforming is a scalar coding strategy in which input covariance matrix is unit-rank. In beamforming, the symbol stream is coded and multiplied by different coefficients at each antenna before transmission. Since the available mature scalar codec technology can be used, beamforming is highly desirable. However, in the setting of [1], the optimum input covariance matrix is full-rank, and therefore beamforming is not optimal.

Although beamforming is not optimal for no-CSI case, it is conditionally optimal, in a single-user setting, when the transmitter has the partial knowledge of the channel [3], [4], [5]. For the covariance feedback assumption, the fact that the optimal transmit covariance matrix and the channel covariance matrix have the same eigenvectors was shown in [3] for a

multi-input single-output (MISO) system, and in [4] for a MIMO system. Using this, the conditions on the *channel covariance matrix* that guarantee that the *transmit covariance matrix* is unit-rank, and therefore beamforming is optimal, are identified in [4], [5].

Reference [6] gave some results on the optimality of beamforming in a MIMO-MAC system where the channel is deterministic and fully known to the transmitters. In [6], it was shown that the number of users that have transmit covariance matrices with ranks larger than one, is less than or equal to the square of the number of receiver antennas. Therefore, if the number of users is much larger than the number of receiver antennas, then unit-rank transmission is optimal for almost all users. It is important to note that in [6] the channel is fixed and known to the transmitters. If the channel is fading, the optimality of beamforming cannot be deduced by using the techniques of [6].

In this paper, we consider a *multi-user fading* MIMO multiple access system with various assumptions on the channel state information, i.e., on the parameters of the distribution of the channel that are available to the transmitters. We show that, if there is some form of CSI at the transmitters, all users should transmit in the direction of the eigenvectors of their *own* channel parameter matrices. Therefore, we show that, the transmit directions of the users are independent of the presence of other users.

Then, motivated by the observation of [6] that beamforming is optimal asymptotically in a multi-user MIMO-MAC where the receiver and the transmitters have perfect CSI, we ask the question whether beamforming is asymptotically optimal in our case where the receiver has perfect CSI, but the transmitters have no or partial CSI. When the channel is not known at the transmitters, it is counter-intuitive to think that beamforming would be optimal unconditionally. Confirming this intuition, [1] already showed that even in the multi-user case, the transmitters should divide their powers equally over their transmit antennas, resulting in full-rank transmission for all users. However, we show that unit-rank input transmit covariance matrices are asymptotically optimal for all users when the transmitters have no knowledge of the channel. The beamforming scheme we use in this case is simpler than usual; it may be characterized as an arbitrary antenna selection scheme. Therefore, we show that, in a large system, a beamforming scheme where each user transmits from an arbitrarily selected antenna, results in optimal performance.

When the transmitters have the partial CSI in the form of either covariance feedback or mean feedback, we show that asymptotic optimality of beamforming still holds. In these cases however, arbitrary antenna selection scheme is no longer optimal. For the covariance feedback setting, each user beamforms in the direction of the strongest eigenvector of its feedback channel covariance matrix. In [7], we found the region where beamforming is optimal for all users in a multi-user setting. In this finite-sized system, beamforming may or may not be optimal for all users depending on the eigenvalues of the channel covariance matrices. We show here that for a large system, beamforming is always optimal. For the mean feedback setting, each user beamforms in the direction of the eigenvector corresponding to the sole non-zero eigenvalue of its feedback channel mean matrix. Similar to the covariance feedback case, beamforming is optimal asymptotically irrespective of the values of the mean information.

## II. SYSTEM MODEL

We consider a multi-user multiple access channel with multiple transmit antennas at every user and multiple receive antennas at the receiver. The channel between user  $k$  and the receiver is represented by a random matrix  $\mathbf{H}_k$  with dimensions of  $n_R \times n_T$ , where  $n_R$  and  $n_T$  are the number of antennas at the receiver and at the transmitter, respectively. The receiver has perfect knowledge of the channel, while the transmitters have only the statistical model of the channel. Each transmitter sends a vector  $\mathbf{x}_k$ , and the received vector is

$$\mathbf{r} = \sum_{k=1}^K \mathbf{H}_k \mathbf{x}_k + \mathbf{n} \quad (1)$$

where  $K$  is the number of users,  $\mathbf{n}$  is a zero-mean identity-covariance circularly symmetric complex Gaussian vector, and the entries of  $\mathbf{H}_k$  are complex Gaussian random variables. Let  $\mathbf{Q}_k = E[\mathbf{x}_k \mathbf{x}_k^\dagger]$  be the transmit covariance matrix of user  $k$ , which has an average power constraint of  $P$ ,  $\text{tr}(\mathbf{Q}_k) \leq P$ .

We investigate three different statistical models at the transmitters. The first one is the “no CSI” model in which the transmitters only know the distribution of the channel while the parameters of the distribution are not known. In this case, the entries of  $\mathbf{H}_k$  are i.i.d., zero-mean, unit-variance complex Gaussian random variables. This model is used in [1], [2], and [8].

The second model is the “partial CSI with covariance feedback” model where each transmitter knows the channel covariance matrix of its own channel, in addition to the distribution of the channel. This model is used in [3], [4], and [5]. The entries in every column of  $\mathbf{H}_k$  are i.i.d., complex Gaussian random variables, and the entries in every row of  $\mathbf{H}_k$  are correlated complex Gaussian random variables with covariance  $\Sigma_k$ . In this case, the channel of user  $k$  can be written as

$$\mathbf{H}_k = \mathbf{Z}_k \Sigma_k^{1/2} \quad (2)$$

where the entries of  $\mathbf{Z}_k$  are i.i.d., zero-mean, unit-variance complex Gaussian random variables.

The third model we investigate is the “partial CSI with mean feedback” model where each transmitter knows the channel mean matrix of its own channel, in addition to the distribution of the channel. This model is used in [3], [4], [9], and [10]. In this model, the transmitters have line-of-sight component with the receiver and are assumed to be close to each other. Therefore, their signals arrive at the base station in-phase. Moreover, this Ricean channel is modeled to be of unit-rank [10]. As a result, the entries of the channel matrix are independent with a non-zero mean. In this case, the channel of user  $k$  can be written as

$$\mathbf{H}_k = \mathbf{H}_k^\mu + \mathbf{Z}_k \quad (3)$$

where the entries of  $\mathbf{Z}_k$  are i.i.d., zero-mean, unit-variance complex Gaussian random variables, and  $\mathbf{H}_k^\mu$  is the mean information representing the line-of-sight component of the channel. The mean matrix takes the form [10],

$$\mathbf{H}_k^\mu = \mathbf{a}_{Rk} \mathbf{a}_{Tk}^\dagger \quad (4)$$

where  $\mathbf{a}_{Rk}$  and  $\mathbf{a}_{Tk}$  are the specular array response vectors at the receiver and the transmitter, respectively. As a result of the in-phase assumption, the array response of the receiver will be assumed to be the same for all users, that is,  $\mathbf{a}_{Rk} = \mathbf{a}_R$ , for all  $k$ .

## III. ASYMPTOTIC ANALYSIS

The sum capacity of a multi-user MIMO-MAC system is given as,

$$C = \max_{\substack{\text{tr}(\mathbf{Q}_k) \leq P \\ k=1 \dots K}} E \left[ \log \left| \mathbf{I}_{n_R} + \sum_{k=1}^K \mathbf{H}_k \mathbf{Q}_k \mathbf{H}_k^\dagger \right| \right] \quad (5)$$

where  $E[\cdot]$  is the expectation operator over the channel matrices of all users,  $|\cdot|$  is the determinant operator, and  $C$  denotes the sum capacity. In this section, we show that for large numbers of users, even with the assumption that the transmitters have no knowledge on the channel, beamforming approaches the optimal capacity. For asymptotic analysis, we need the following lemma.

*Lemma 1:* Let  $\mathbf{x}_i$ ,  $i = 1, 2, \dots$  be a sequence of i.i.d. random vectors of length  $M$ , which have zero-mean and identity-covariance matrix, and let  $\alpha_i$ ,  $i = 1, 2, \dots$  be a sequence of real numbers. Then,

$$E \left[ \log \left| \mathbf{I}_M + \sum_{i=1}^N \alpha_i \mathbf{x}_i \mathbf{x}_i^\dagger \right| \right] \xrightarrow{a.s.} M \log \left( 1 + \sum_{i=1}^N \alpha_i \right) \quad (6)$$

The proof of Lemma 1 which uses the strong law of large numbers (SLLN) is omitted here due to space limitations.

### A. No CSI at the Transmitters

For a multi-user system with no CSI at the transmitters and identity channel covariance matrix, i.e.,  $\Sigma_k = \mathbf{I}$ , the sum capacity is [1],

$$C = E \left[ \log \left| \mathbf{I}_{n_R} + \frac{P}{n_T} \sum_{k=1}^K \mathbf{H}_k \mathbf{H}_k^\dagger \right| \right] \quad (7)$$

This means that in the multi-user setting as well, the optimal transmit covariance matrix for each user is proportional to identity, i.e.,  $\mathbf{Q}_k = (P/n_T)\mathbf{I}$ . Therefore, it is clear that, beamforming is not optimal for any user when the transmitters do not have any CSI. However, in this section, we show that when there is no CSI at the transmitters, arbitrary antenna selection is sufficient to achieve the optimum sum capacity asymptotically. This is stated in the following theorem.

*Theorem 1:* In a system where there is no CSI at the transmitters, if the number of users grows to infinity, then the sum rate achieved by unit-rank input covariance matrices approaches the optimum sum capacity. In particular, this unit-rank transmission scheme takes the form of simple antenna selection.

*Proof (Theorem 1):* The optimum sum capacity is given in (7). We define  $C^{bf}$  as the achievable sum rate by performing arbitrary antenna selection at all transmitters:

$$C^{bf} = E \left[ \log \left| \mathbf{I}_{n_R} + P\mathbf{H}_{bf}\mathbf{H}_{bf}^\dagger \right| \right] \quad (8)$$

where  $\mathbf{H}_{bf} = [\mathbf{h}_{1i}, \dots, \mathbf{h}_{Ki}]$  and  $\mathbf{h}_{ki}$  is the  $i^{th}$  column of the channel matrix of user  $k$ . The choice of the columns does not affect our result. All users may select their first antenna, i.e., the first column of their channel matrix, or they may select a column arbitrarily. Since SLLN averages out the randomness in the channel regardless of the realizations, so long as the columns of the channel matrices are independent, the transmit antenna each user selects is immaterial.

By using Lemma 1 in calculating the expectations in (7) and (8), we obtain

$$\lim_{K \rightarrow \infty} C = \lim_{K \rightarrow \infty} C^{bf} \rightarrow n_R \log(1 + KP) \quad (9)$$

Therefore, we see that the sum capacities achievable by the optimal power allocation and the arbitrary antenna selection scheme converge to the same quantity asymptotically.  $\square$

## B. Covariance Feedback at the Transmitters

1) *Transmit Directions:* In a single-user system with partial CSI in the form of channel covariance matrix at the transmitter, the capacity is no longer achieved by an identity input covariance matrix. The channel covariance matrix  $\Sigma$ , which is known at the transmitter, has the eigenvalue decomposition  $\Sigma = \mathbf{U}_\Sigma \Lambda_\Sigma \mathbf{U}_\Sigma^\dagger$  where  $\Lambda_\Sigma$  is the diagonal matrix of ordered eigenvalues of  $\Sigma$ , and  $\mathbf{U}_\Sigma$  is a unitary matrix. Similarly, the transmit covariance matrix  $\mathbf{Q}$  has the eigenvalue decomposition  $\mathbf{Q} = \mathbf{U}_Q \Lambda_Q \mathbf{U}_Q^\dagger$ . Reference [4] showed that the eigenvectors of the transmit covariance matrix must be equal to the eigenvectors of the channel covariance matrix, i.e.,  $\mathbf{U}_Q = \mathbf{U}_\Sigma$ .

In a multi-user setting with finite number of users, where there is covariance feedback at the transmitters, we show in this paper that all users should transmit along the eigenvectors of their own channel covariance matrices, regardless of the power allocation scheme. This is stated in the following theorem.

*Theorem 2:* Let  $\Sigma_k = \mathbf{U}_{\Sigma_k} \Lambda_{\Sigma_k} \mathbf{U}_{\Sigma_k}^\dagger$  be the spectral decomposition of the channel covariance matrix of user  $k$ . Then, the optimum input covariance matrix  $\mathbf{Q}_k$  of user  $k$  has the form  $\mathbf{Q}_k = \mathbf{U}_{\Sigma_k} \Lambda_Q \mathbf{U}_{\Sigma_k}^\dagger$ , for all users.

*Proof (Theorem 2):* From (2), we have the following zero-mean, identity-covariance random channel matrix representation  $\mathbf{Z}_k$  for user  $k$ ,

$$\mathbf{Z}_k \mathbf{U}_{\Sigma_k} \Lambda_{\Sigma_k}^{1/2} \mathbf{U}_{\Sigma_k}^\dagger = \mathbf{H}_k \quad (10)$$

Then, inserting (10) into (5), we obtain

$$C = \max_{\substack{\text{tr}(\mathbf{Q}_k) \leq P \\ k=1, \dots, K}} E \left[ \log \left| \mathbf{I} + \sum_{k=1}^K \mathbf{Z}_k \Lambda_{\Sigma_k}^{1/2} \mathbf{U}_{\Sigma_k}^\dagger \mathbf{Q}_k \mathbf{U}_{\Sigma_k} \Lambda_{\Sigma_k}^{1/2} \mathbf{Z}_k^\dagger \right| \right] \quad (11)$$

where the random matrices  $\mathbf{Z}_k \mathbf{U}_{\Sigma_k}$  and  $\mathbf{Z}_k$  have the same distribution for zero-mean identity-covariance Gaussian  $\mathbf{Z}_k$  and unitary  $\mathbf{U}_{\Sigma_k}$  [1]. We may spectrally decompose the expression sandwiched between the equivalent channel matrix and its conjugate transpose in (11) as

$$\Lambda_{\Sigma_k}^{1/2} \mathbf{U}_{\Sigma_k}^\dagger \mathbf{Q}_k \mathbf{U}_{\Sigma_k} \Lambda_{\Sigma_k}^{1/2} = \mathbf{U}_k \Lambda_k \mathbf{U}_k^\dagger \quad (12)$$

where  $\Lambda_k$  is a diagonal matrix with ordered components such that  $\lambda_{k1} \geq \lambda_{k2} \geq \dots \geq \lambda_{kn_T}$ . The optimization problem in (5) and (11) may now be written as

$$C = \max_{\substack{\text{tr}(\mathbf{Q}_k) \leq P \\ k=1, \dots, K}} E \left[ \log \left| \mathbf{I}_{n_R} + \sum_{k=1}^K \mathbf{Z}_k \Lambda_k \mathbf{Z}_k^\dagger \right| \right] \quad (13)$$

where we again used the fact that the random matrices  $\mathbf{Z}_k \mathbf{U}_k$  and  $\mathbf{Z}_k$  have the same distribution. Using (12), the trace constraint on  $\mathbf{Q}_k$  can be expressed as

$$\text{tr}(\mathbf{Q}_k) = \text{tr}(\mathbf{U}_{\Sigma_k} \Lambda_{\Sigma_k}^{-1/2} \mathbf{U}_k \Lambda_k \mathbf{U}_k^\dagger \Lambda_{\Sigma_k}^{-1/2} \mathbf{U}_{\Sigma_k}^\dagger) \quad (14)$$

$$= \text{tr}(\mathbf{U}_k^\dagger \Lambda_{\Sigma_k}^{-1} \mathbf{U}_k \Lambda_k) \quad (15)$$

where the second equality follows from  $\text{tr}(\mathbf{A}\mathbf{B}) = \text{tr}(\mathbf{B}\mathbf{A})$ .

From [11, Theorem H.1.h],  $\text{tr}(\Lambda_{\Sigma_k}^{-1} \Lambda_k) \leq \text{tr}(\mathbf{U}_k^\dagger \Lambda_{\Sigma_k}^{-1} \mathbf{U}_k \Lambda_k) \leq P$ , for all unitary  $\mathbf{U}_k$ . This means that, if we choose  $\mathbf{U}_k = \mathbf{I}$ , the trace constraint will still be satisfied. This choice will not affect the objective function since it does not involve the matrix  $\mathbf{U}_k$ . Then, from (12), we have the desired result:

$$\mathbf{Q}_k = \mathbf{U}_{\Sigma_k} \Lambda_{\Sigma_k}^{-1} \Lambda_k \mathbf{U}_{\Sigma_k}^\dagger \quad (16)$$

with  $\Lambda_{Q_k} = \Lambda_{\Sigma_k}^{-1} \Lambda_k$ .  $\square$

Therefore, using Theorem 2, we can write the optimization problem in (5) as,

$$C = \max_{\substack{\sum_{i=1}^{n_T} \lambda_{ki}^Q \leq P \\ k=1, \dots, K}} E \left[ \log \left| \mathbf{I}_{n_R} + \sum_{k=1}^K \sum_{i=1}^{n_T} \lambda_{ki}^Q \lambda_{ki}^\Sigma \mathbf{z}_{ki} \mathbf{z}_{ki}^\dagger \right| \right] \quad (17)$$

where  $\mathbf{z}_{ki}$  is the  $i^{th}$  column of  $\mathbf{Z}_k$ .

2) *Asymptotic Optimality of Beamforming*: When the transmitters have the partial CSI in the form of covariance feedback, Theorem 2 shows that for any number of users, the transmit directions of a user are the eigenvectors of its channel covariance feedback matrix. For sufficiently large number of users, the asymptotic optimality of beamforming in achieving the optimum sum capacity is proved in the following theorem.

*Theorem 3*: In a system where there is partial CSI at the transmitters in the form of covariance feedback, if the number of users grows to infinity, then the sum rate achieved by unit-rank input covariance matrices approaches the optimum sum capacity. In particular, this beamforming, for each user, is in the direction of the strongest eigenvector of the channel covariance matrix of that user.

*Proof (Theorem 3)*: Applying Lemma 1 to (17),

$$C \xrightarrow{a.s.} \max_{\substack{\sum_{i=1}^{n_T} \lambda_{ki}^Q \leq P \\ k=1 \dots K}} n_R \log \left( 1 + \sum_{k=1}^K \sum_{i=1}^{n_T} \lambda_{ki}^Q \lambda_{ki}^\Sigma \right) \quad (18)$$

In order to solve the above optimization problem, we form the Lagrangian with  $\mu_k$ 's as the Lagrangian multipliers,

$$n_R \log \left( 1 + \sum_{k=1}^K \sum_{i=1}^{n_T} \lambda_{ki}^Q \lambda_{ki}^\Sigma \right) - \sum_{k=1}^K \mu_k \left( \sum_{i=1}^{n_T} \lambda_{ki}^Q - P \right) \quad (19)$$

The KKT optimality conditions for user  $k$  are,

$$\frac{n_R \lambda_{ki}^\Sigma}{1 + \sum_{k=1}^K \sum_{i=1}^{n_T} \lambda_{ki}^Q \lambda_{ki}^\Sigma} \leq \mu_k, \quad i = 1, \dots, n_T \quad (20)$$

where (20) is satisfied with equality if  $\lambda_{ki}^Q > 0$ . Note that the denominators on the left hand side of all the KKT conditions are the same. Without loss of generality, let  $\lambda_{kn_T}^\Sigma < \dots < \lambda_{k1}^\Sigma$  for user  $k$ . Assume that  $\lambda_{kj}^Q > 0$  and  $\lambda_{ki}^Q > 0$ . Then, we must have  $\lambda_{kj}^\Sigma = \lambda_{ki}^\Sigma$ , which is a contradiction. Therefore, for user  $k$ , only one  $\lambda_{kj}^Q$ ,  $j = 1, \dots, n_T$  can be non-zero. From the cost function in (18), we observe that the non-zero  $\lambda_{kj}^Q$  must correspond to the largest eigenvalue of the channel covariance feedback matrix. Hence, the only non-zero power component in  $\Lambda_{Q_k}$  is the first diagonal element. Finally, from the trace constraint, we have  $\lambda_{k1}^Q = P$ , for all  $k$ .

This means that, for a sufficiently large number of users, the transmit covariance matrix of each user is unit-rank, i.e., there is only one non-zero eigenvalue in each of the transmit covariance matrices. Moreover, this non-zero element corresponds to the dominant eigenvector of the channel covariance matrix. This proves that beamforming is optimal and in particular, this beamforming, for each user, is in the direction of the strongest eigenvector of the channel covariance matrix of that user.  $\square$

### C. Mean Feedback at the Transmitters

1) *Transmit Directions*: Similar to the covariance feedback case, for a single-user system with partial CSI in the form of channel mean matrix at the transmitter, the capacity is no longer achieved by an identity input covariance matrix. In this setting, the channel covariance matrix is identity, i.e.,

$\Sigma = \mathbf{I}$ , and the channel mean matrix  $\mathbf{H}_\mu$  is fed back to the transmitter. With the assumption that  $\mathbf{H}_\mu$  is unit-rank, [4] showed that the optimal input covariance matrix can be written as  $\mathbf{Q} = \mathbf{U}_\mu \Lambda_Q \mathbf{U}_\mu^\dagger$ , where the first column of the unitary matrix  $\mathbf{U}_\mu$  is the eigenvector corresponding to the non-zero eigenvalue of  $\mathbf{H}_\mu$ , and the remaining columns are arbitrary, with the restriction that the columns of  $\mathbf{U}_\mu$  are orthonormal.

In this paper, we show that, in a multi-user setting, every user should transmit along the eigenvectors of its own channel mean matrix. Let the singular value decomposition of the channel mean matrix of user  $k$  be

$$\mathbf{H}_k^\mu = \mathbf{U}_{\mu_k} \Lambda_{\mu_k}^{1/2} \mathbf{V}_{\mu_k}^\dagger \quad (21)$$

Since  $\mathbf{H}_k^\mu$  is a unit-rank matrix, the first column of  $\mathbf{U}_{\mu_k}$  can be chosen as  $\frac{\mathbf{a}_R}{|\mathbf{a}_R|}$ ; and the rest of the columns can be chosen arbitrarily as long as  $\mathbf{U}_{\mu_k}$  has orthonormal columns. Also, note that  $\mathbf{U}_{\mu_k} = \mathbf{U}_\mu$ , for  $k = 1, \dots, K$ . Similarly, the first column of  $\mathbf{V}_{\mu_k}$  can be chosen as  $\frac{\mathbf{a}_{T_k}}{|\mathbf{a}_{T_k}|}$  and the rest of the columns can be chosen arbitrarily as long as  $\mathbf{V}_{\mu_k}$  has orthonormal columns. However,  $\mathbf{V}_{\mu_k}$  is different for different users. The diagonal matrix  $\Lambda_{\mu_k}^{1/2}$  has only one non-zero element, which is  $|\mathbf{a}_R| |\mathbf{a}_{T_k}|$ .

In the following theorem, we identify the optimum transmit directions of all users. The single-user version of this theorem was proved in [9]. Here, we provide and prove a general version, valid for a multi-user system.

*Theorem 4*: Let  $\mathbf{H}_k^\mu = \mathbf{U}_{\mu_k} \Lambda_{\mu_k}^{1/2} \mathbf{V}_{\mu_k}^\dagger$  be the singular value decomposition of the channel mean matrix of user  $k$ . Then, the optimum input covariance matrix  $\mathbf{Q}_k$  of user  $k$  may be expressed as  $\mathbf{Q}_k = \mathbf{V}_{\mu_k} \Lambda_k \mathbf{V}_{\mu_k}^\dagger$ , for all users.

*Proof (Theorem 4)*: We prove the theorem in two steps. In the first step, we show that the sum capacity resulting from  $\{\mathbf{H}_k^\mu\}_{k=1}^K$  as the channel mean matrices and the sum capacity resulting from  $\{\Lambda_{\mu_k}^{1/2}\}_{k=1}^K$  as the channel mean matrices are the same.

The optimization problem in (5) with channel mean matrices  $\{\mathbf{H}_k^\mu\}_{k=1}^K$  can be written as

$$C(\{\mathbf{H}_k^\mu\}_{k=1}^K) = \max_{\substack{\text{tr}(\mathbf{Q}_k) \leq P \\ k=1 \dots K}} E \left[ \log \left| \mathbf{I} + \sum_{k=1}^K (\mathbf{H}_k^\mu + \mathbf{Z}_k) \mathbf{Q}_k (\mathbf{H}_k^\mu + \mathbf{Z}_k)^\dagger \right| \right] \quad (22)$$

$$= \max_{\substack{\text{tr}(\mathbf{Q}_k) \leq P \\ k=1 \dots K}} E \left[ \log \left| \mathbf{I} + \sum_{k=1}^K (\mathbf{U}_{\mu_k} \Lambda_{\mu_k}^{1/2} \mathbf{V}_{\mu_k}^\dagger + \mathbf{Z}_k) \mathbf{Q}_k (\mathbf{U}_{\mu_k} \Lambda_{\mu_k}^{1/2} \mathbf{V}_{\mu_k}^\dagger + \mathbf{Z}_k)^\dagger \right| \right] \quad (23)$$

$$= \max_{\substack{\text{tr}(\mathbf{Q}_k) \leq P \\ k=1 \dots K}} E \left[ \log \left| \mathbf{I} + \sum_{k=1}^K (\Lambda_{\mu_k}^{1/2} + \mathbf{Z}_k) \tilde{\mathbf{Q}}_k (\Lambda_{\mu_k}^{1/2} + \mathbf{Z}_k)^\dagger \right| \right] \quad (24)$$

$$= C(\{\Lambda_{\mu_k}^{1/2}\}_{k=1}^K) \quad (25)$$

where we use the singular value decomposition of the channel mean matrix of user  $k$  in (23), and the invariance of the distribution of zero-mean, identity-covariance matrix  $\mathbf{Z}_k$  under unitary transformations in (24).

Note that  $\text{tr}(\mathbf{Q}_k) = \text{tr}(\tilde{\mathbf{Q}}_k)$ , since  $\tilde{\mathbf{Q}}_k = \mathbf{V}_{\mu_k}^\dagger \mathbf{Q}_k \mathbf{V}_{\mu_k}$ . By comparing (22) and (24), we see that the diagonal eigenvalue matrices of the channel mean matrices result in the same capacity as the channel mean matrices themselves except that we changed the input covariance matrices accordingly.

In the second step, our goal is to show that the optimal  $\tilde{\mathbf{Q}}_k$  in (24) is diagonal. In order to prove this, we use the technique presented in [9]. This technique depends on the cancellation of the off-diagonal elements of an arbitrary transmit covariance matrix. The optimality of a diagonal covariance matrix is then proved using the concavity of the mutual information with respect to the input covariance matrices. We generalize the technique in [9] to a multi-user setting in [12]. The details are omitted here due to space limitations. Since  $\tilde{\mathbf{Q}}_k$  is diagonal, we have the desired result, which is  $\mathbf{Q}_k = \mathbf{V}_{\mu_k} \mathbf{\Lambda}_k \mathbf{V}_{\mu_k}^\dagger$ .  $\square$

Therefore, using Theorem 4, we can write the optimization problem in (5) as,

$$C = \max_{\substack{\sum_{k=1}^{n_T} \lambda_{ki}^Q \leq P \\ k=1, \dots, K}} E \left[ \log \left| \mathbf{I}_{n_R} + \sum_{k=1}^K \sum_{i=1}^{n_T} \lambda_{ki}^Q \hat{\mathbf{z}}_{ki} \hat{\mathbf{z}}_{ki}^\dagger \right| \right] \quad (26)$$

where  $\hat{\mathbf{z}}_{ki}$  is the  $i^{\text{th}}$  column of  $\mathbf{\Lambda}_{\mu_k}^{1/2} + \mathbf{Z}_k$ .

2) *Asymptotic Optimality of Beamforming*: When the transmitters have the partial CSI in the form of mean feedback, Theorem 4 shows that for any number of users, the transmit directions of a user are the eigenvectors of its channel mean feedback matrix. For sufficiently large number of users, the asymptotic optimality of beamforming in achieving the optimum sum capacity is proved in the following theorem.

*Theorem 5*: In a system where there is partial CSI at the transmitters in the form of the mean feedback information, if the number of users grows to infinity, then the sum rate achieved by unit-rank input covariance matrices approaches the optimum sum capacity. In particular, this beamforming, for each user, is in the direction of the eigenvector corresponding to the sole non-zero eigenvalue of the channel mean matrix of that user.

*Proof (Theorem 5)*: Applying Lemma 1 to (26), we have

$$C \xrightarrow{a.s.} \max_{\substack{\sum_{k=1}^{n_T} \lambda_{ki}^Q \leq P \\ k=1, \dots, K}} \sum_{i=1}^{n_R} \log \left( 1 + \sum_{k=1}^K \lambda_{ki}^\mu \lambda_{ki}^Q + KP \right) \quad (27)$$

Note that, since the channel mean matrices are unit-rank, for any  $k$ , only  $\lambda_{k1}^\mu$  is non-zero. Therefore, (27) becomes,

$$C \xrightarrow{a.s.} \max_{\substack{\sum_{k=1}^{n_T} \lambda_{k1}^Q \leq P \\ k=1, \dots, K}} L + \log \left( 1 + \sum_{k=1}^K \lambda_{k1}^\mu \lambda_{k1}^Q + KP \right) \quad (28)$$

where  $L = (n_R - 1) \log(1 + KP)$ . Since only one eigenvalue from the input covariance matrix of each user appears in (28), the optimum choice is to allocate all of the power of each user to that eigenvalue, i.e.,  $\lambda_{k1}^Q = P$ , for all  $k$ .  $\square$

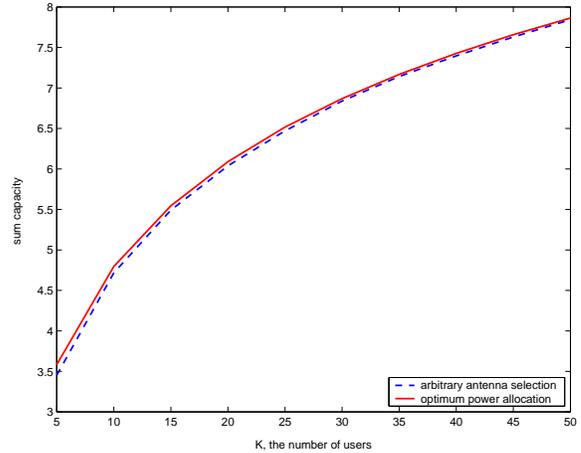


Fig. 1. The sum rates resulting from optimum power allocation and arbitrary antenna selection for different numbers of users in the no-CSI model.

#### IV. NUMERICAL ANALYSIS

In Figure 1, we show a simple example where the number of receiver and transmitter antennas is equal to 2. We plot the sum rates resulting from optimal power allocation and arbitrary antenna selection schemes for the no-CSI model. We observe from the figure that for this instance, even for a small number of users, arbitrary antenna selection performs very close to the optimum power allocation scheme.

#### REFERENCES

- [1] I. E. Telatar. Capacity of multi-antenna Gaussian channels. *European Transactions on Telecommunication*, 10(6):585–596, November 1999.
- [2] G. J. Foschini and M. J. Gans. On limits of wireless communication in a fading environment when using multiple antennas. *Wireless Personal Communications*, 6:311–335, 1998.
- [3] E. Visotsky and U. Madhow. Space-time transmit precoding with imperfect feedback. *IEEE Transactions on Information Theory*, 47(6):2632–2639, September 2001.
- [4] S. A. Jafar and A. Goldsmith. Transmitter optimization and optimality of beamforming for multiple antenna systems. *IEEE Transactions on Wireless Communications*, 3(4):1165–1175, July 2004.
- [5] H. Boche and E. Jorswieck. On the optimality range of beamforming for MIMO systems with covariance feedback. *IEICE Trans. Commun.*, E85-A(11):2521–2528, November 2002.
- [6] W. Rhee, W. Yu, and J. M. Cioffi. The optimality of beamforming in uplink multiuser wireless systems. *IEEE Transactions on Wireless Communications*, 3(1):86–96, January 2004.
- [7] A. Soysal and S. Ulukus. Transmit directions and optimality of beamforming in MIMO-MAC with partial CSI at the transmitters. In *Conference on Information Sciences and Systems*, March 2005. Submitted.
- [8] B. M. Hochwald, T. L. Marzetta, and V. Tarokh. Multiple-antenna channel hardening and its implications for rate feedback and scheduling. *IEEE Transactions on Information Theory*, 50(9):1893–1909, September 2004.
- [9] D. Höslı and A. Lapidoth. The capacity of a MIMO Ricean channel is monotonic in the singular values of the mean. In *5<sup>th</sup> International ITG Conference on Source and Channel Coding*, January 2004.
- [10] A. Lozano, F. R. Farrokhi, and R. A. Valenzuela. Asymptotically optimal open-loop space-time architecture adaptive to scattering conditions. In *Vehicular Technology Conference*, Spring 2001.
- [11] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications*. New York:Academic, 1979.
- [12] A. Soysal and S. Ulukus. Optimality of beamforming in fading MIMO multiple access channels. *To be submitted for journal publication*.