Transmit Directions and Optimality of Beamforming in MIMO-MAC with Partial CSI at the Transmitters[1]

Alkan Soysal Sennur Ulukus
Department of Electrical and Computer Engineering
University of Maryland, College Park, MD 20705
alkan@umd.edu ulukus@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 partial CSI, which is in the form of either the covariance matrix of the channel or the mean matrix of the channel. We show that the transmit directions of each user are the eigenvectors of the channel covariance and mean feedback matrices in covariance and mean feedback models, respectively. Then, we find the conditions under which beamforming is optimal for all users. We observe through simulations that the region where beamforming is optimal for all users gets larger with the addition of new users to the system.

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 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 multi-input single-output (MISO) system, and in [4] for MIMO system. The optimal power allocation strategy was shown to be similar to water-filling over the eigenvalues of the channel covariance matrix in [6]. 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]. This result is analogous to identifying the conditions on the channel state space and the average power in classical water-filling that guarantee that only one channel is filled as a result of having either a low power constraint or one very strong channel. Similarly, for the mean feedback assumption, the eigenvectors of the optimal transmit covariance matrix were shown to be the same as the eigenvectors of the channel mean matrix for a MISO system in [3] and for a MIMO systems in [4]. Using this, the conditions on the *channel mean matrix* that guarantee that the *transmit covariance matrix* is unit-rank, and therefore beamforming is optimal, are identified in [4].

In this paper, we consider a *multi-user* MIMO multiple access system. 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. This means that the users maintain their single-user transmit direction strategies even in a multi-user scenario. Then, we identify the necessary and sufficient conditions for the optimality of beamforming for all users. This result generalizes the single-user conditions of [4], [5] to a multi-user setting. In the case of covariance feedback, these conditions depend only on the first and second largest eigenvalues of the channel covariance matrix of each user, and they form a region in a space whose dimension is twice the number of users. If the first and second largest eigenvalues of the feedback covariance matrices of all users fall inside this region, beamforming is optimal for all users. In the case of mean feedback, these conditions depend only on the sole non-zero eigenvalue of the unit-rank channel mean matrix of each user. Similarly, if the channel mean eigenvalues of all users fall inside this region, beamforming is optimal for all users. The results of [4] and [5] can be obtained as special cases of our results. In addition to identifying the region where beamforming is optimal for all users, we observe through simulations that this region gets larger as new users are added to the system.

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

[1]This work was supported by NSF Grants ANI 02-05330 and CCR 03-11311; and ARL/CTA Grant DAAD 19-01-2-0011.

transmitter sends a vector $\mathbf{x}_k$, and the received vector is

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

where K is the number of users, $\mathbf{n}$ is a zero-mean identity-covariance 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 two different statistical models at the transmitters. The first 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], [5], and [6]. In this model, the receiver is assumed to be unobstructed and the transmitters are assumed to be obstructed by local scatterers. Therefore, the channels seen by a signal sent from a transmitter antenna of one user to all receiver antennas are uncorrelated (this corresponds to a column of the channel matrix of that user), and the channels seen by the signals sent from all transmitter antennas of one user to a single receiver antenna are correlated (this corresponds to a row of the channel matrix of that user). As a result, 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 $\mathbf{\Sigma}_k$. In this case, the channel of user k can be written as

$$\mathbf{H}_k = \mathbf{Z}_k \mathbf{\Sigma}_k^{1/2} \tag{2}$$

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

The second 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], [7], and [8]. 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 modelled to be of unit-rank [8]. As a result, the entries of the channel matrix are independent and have non-zero means. In this case, the channel of user k can be written as

$$\mathbf{H}_k = \mathbf{H}_{\mu_k} + \mathbf{Z}_k \tag{3}$$

where entries of $\mathbf{Z}_k$ are i.i.d., zero-mean, unit-variance complex Gaussian random variables, and $\mathbf{H}_{\mu_k}$ is the mean information representing the line-of-sight component of the channel. The mean matrix takes the form [8],

$$\mathbf{H}_{\mu_k} = \mathbf{a}_{R_k} \mathbf{a}_{T_k}^\dagger \tag{4}$$

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

The sum capacity optimization problem for this system is

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

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.

III. Covariance Feedback at the Transmitters

III.A Transmit Directions

In a single-user system with partial CSI in the form of the channel covariance matrix at the transmitter, let the channel covariance matrix $\mathbf{\Sigma}$ have the eigenvalue decomposition,

$$\mathbf{\Sigma} = \mathbf{U}_\Sigma \mathbf{\Lambda}_\Sigma \mathbf{U}_\Sigma^\dagger \tag{6}$$

where $\mathbf{\Lambda}_\Sigma$ is the diagonal matrix of ordered eigenvalues of $\mathbf{\Sigma}$ such that $\lambda_1^\Sigma \geq \lambda_2^\Sigma \geq \cdots \geq \lambda_{n_T}^\Sigma$, and $\mathbf{U}_\Sigma$ is a unitary matrix. The matrix $\mathbf{Q}$ has the eigenvalue decomposition

$$\mathbf{Q} = \mathbf{U}_Q \mathbf{\Lambda}_Q \mathbf{U}_Q^\dagger \tag{7}$$

where $\mathbf{\Lambda}_Q$ is the diagonal matrix of ordered eigenvalues of $\mathbf{Q}$ such that $\lambda_1^Q \geq \lambda_2^Q \geq \cdots \geq \lambda_{n_T}^Q$, and $\mathbf{U}_Q$ is a unitary matrix. Reference [4] showed that the eigenvectors of the transmit covariance matrix must be equal to the eigenvectors of the channel covariance matrix, i.e., $\mathbf{U}_Q = \mathbf{U}_\Sigma$. Reference [6] proposed a numerical optimization based method to find $\mathbf{\Lambda}_Q$, the power put by the transmitter along the eigen-directions. References [4] and [5] showed that under certain conditions on the covariance feedback matrix $\mathbf{\Sigma}$, the power matrix $\mathbf{\Lambda}_Q$ has only one non-zero diagonal element, i.e., the optimal transmit covariance matrix is unit-rank, and therefore beamforming in the direction of the eigenvector corresponding to this non-zero eigenvalue is optimal.

In a multi-user setting with finite number of users, where there is covariance feedback at the transmitters, the sum capacity achieving power allocation scheme is not known. 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 1 *Let $\mathbf{\Sigma}_k = \mathbf{U}_{\Sigma_k} \mathbf{\Lambda}_{\Sigma_k} \mathbf{U}_{\Sigma_k}^\dagger$ be the eigenvalue decomposition of the channel covariance matrix of user k. Then, the optimum input covariance matrix $\mathbf{Q}_k$ of user k has the form $\mathbf{Q}_k = \mathbf{U}_{\Sigma_k} \mathbf{\Lambda}_Q \mathbf{U}_{\Sigma_k}^\dagger$, for all users.*

Proof (Theorem 1): From (2), we have the following zero-mean, identity-covariance random channel matrix representation $\mathbf{Z}_k$ for user k,

$$\mathbf{Z}_k \mathbf{U}_{\Sigma_k} \mathbf{\Lambda}_{\Sigma_k}^{1/2} \mathbf{U}_{\Sigma_k}^\dagger = \mathbf{H}_k \tag{8}$$

Then, inserting (8) into (5), we obtain

$$C = \max_{\substack{\text{tr}(\mathbf{Q}_k) \leq P \\ k=1,\dots,K}} E\left[\log\left|\mathbf{I}_{n_R} + \sum_{k=1}^{K} \mathbf{Z}_k \mathbf{\Lambda}_{\Sigma_k}^{1/2} \mathbf{U}_{\Sigma_k}^\dagger \mathbf{Q}_k \mathbf{U}_{\Sigma_k} \mathbf{\Lambda}_{\Sigma_k}^{1/2} \mathbf{Z}_k^\dagger\right|\right] \tag{9}$$

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 (9) as

$$\mathbf{\Lambda}_{\Sigma_k}^{1/2} \mathbf{U}_{\Sigma_k}^\dagger \mathbf{Q}_k \mathbf{U}_{\Sigma_k} \mathbf{\Lambda}_{\Sigma_k}^{1/2} = \mathbf{U}_k \mathbf{\Lambda}_k \mathbf{U}_k^\dagger \tag{10}$$

where $\mathbf{\Lambda}_k$ is a diagonal matrix with ordered components such that $\lambda_{k1} \geq \lambda_{k2} \geq \cdots \geq \lambda_{kn_T}$. The optimization problem in (9) may now be written as

$$C = \max_{\substack{\mathrm{tr}(\mathbf{Q}_k)\leq P\\ k=1,\ldots,K}} E\left[\log\left|\mathbf{I}_{n_R} + \sum_{k=1}^{K} \mathbf{Z}_k\mathbf{U}_k\mathbf{\Lambda}_k(\mathbf{Z}_k\mathbf{U}_k)^\dagger\right|\right] \quad (11)$$

$$= \max_{\substack{\mathrm{tr}(\mathbf{Q}_k)\leq P\\ k=1,\ldots,K}} E\left[\log\left|\mathbf{I}_{n_R} + \sum_{k=1}^{K} \mathbf{Z}_k\mathbf{\Lambda}_k\mathbf{Z}_k^\dagger\right|\right] \quad (12)$$

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 (10), the trace constraint on $\mathbf{Q}_k$ can be expressed as

$$\mathrm{tr}(\mathbf{Q}_k) = \mathrm{tr}(\mathbf{U}_{\Sigma_k}\mathbf{\Lambda}_{\Sigma_k}^{-1/2}\mathbf{U}_k\mathbf{\Lambda}_k\mathbf{U}_k^\dagger\mathbf{\Lambda}_{\Sigma_k}^{-1/2}\mathbf{U}_{\Sigma_k}^\dagger) \quad (13)$$

$$= \mathrm{tr}(\mathbf{U}_k^\dagger\mathbf{\Lambda}_{\Sigma_k}^{-1}\mathbf{U}_k\mathbf{\Lambda}_k) \quad (14)$$

where the second equality follows from $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$.

For all unitary $\mathbf{U}_k$, $\mathrm{tr}(\mathbf{\Lambda}_{\Sigma_k}^{-1}\mathbf{\Lambda}_k) \leq \mathrm{tr}(\mathbf{U}_k^\dagger\mathbf{\Lambda}_{\Sigma_k}^{-1}\mathbf{U}_k\mathbf{\Lambda}_k) \leq P$ [9, Theorem H.1.h]. 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 (10), we have the desired result:

$$\mathbf{Q}_k = \mathbf{U}_{\Sigma_k}\mathbf{\Lambda}_{\Sigma_k}^{-1}\mathbf{\Lambda}_k\mathbf{U}_{\Sigma_k}^\dagger \quad (15)$$

with $\mathbf{\Lambda}_{Q_k} = \mathbf{\Lambda}_{\Sigma_k}^{-1}\mathbf{\Lambda}_k$. □

Therefore, using Theorem 1, we can write the optimization problem in (5) as,

$$C = \max_{\substack{\mathrm{tr}\left(\mathbf{\Lambda}_{Q_k}\right)\leq P\\ k=1,\ldots,K}} E\left[\log\left|\mathbf{I}_{n_R} + \sum_{k=1}^{K} \mathbf{Z}_k\mathbf{\Lambda}_{Q_k}\mathbf{\Lambda}_{\Sigma_k}\mathbf{Z}_k^\dagger\right|\right] \quad (16)$$

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

where $\mathbf{z}_{ki}$ is the i^{th} column of $\mathbf{Z}_k$, i.e., $\mathbf{z}_{ki}$, for $k = 1,\ldots,K$ and $i = 1,\ldots,n_T$, are a set of $n_R \times 1$ dimensional i.i.d. zero-mean, unit-variance random Gaussian vectors.

III.B Conditions for the Optimality of Beamforming

In this section, we identify the conditions for the optimality of beamforming in a multi-user system with a finite number of users. References [4] and [5] found these conditions in a single-user system. For a single user system, let λ_1^Σ and λ_2^Σ denote the largest and second largest eigenvalues of the channel covariance matrix $\mathbf{\Sigma}$, respectively. Then, the necessary and sufficient condition for the optimality of beamforming is [4],

$$P\lambda_2^\Sigma \leq \frac{1 - E\left[\frac{1}{1+P\lambda_1^\Sigma\mathbf{z}^\dagger\mathbf{z}}\right]}{n_R - 1 + E\left[\frac{1}{1+P\lambda_1^\Sigma\mathbf{z}^\dagger\mathbf{z}}\right]} \quad (18)$$

where $\mathbf{z}$ is an $n_R \times 1$ dimensional Gaussian random vector with zero-mean and identity-covariance.

In this paper, we find the conditions for the optimality of beamforming for all users in a multi-user setting. Our method is somewhat different than that of [4]. Inserting $K = 1$ in our results provides an alternative proof for (18). In our results, the number of conditions equals to the number of users. The condition corresponding to user k depends on the two largest eigenvalues of the channel covariance matrix of that user, and the largest eigenvalues of the channel covariance matrices of all other users. We have the following theorem.

Theorem 2 *In a MIMO-MAC system where the transmitters have partial CSI in the form of covariance feedback, the transmit covariance matrices of all users that maximize (17) have unit-rank (i.e., beamforming is optimal for all users) if and only if*

$$P\lambda_{k2}^\Sigma \leq \frac{1 - E\left[\frac{1}{1+P\lambda_{k1}^\Sigma\mathbf{z}_k\mathbf{A}_k^{-1}\mathbf{z}_k}\right]}{n_R - K + \sum_{l=1}^{K} E\left[\frac{1}{1+P\lambda_{l1}^\Sigma\mathbf{z}_l\mathbf{A}_l^{-1}\mathbf{z}_l}\right]}, \quad k = 1,\ldots,K \quad (19)$$

where $\mathbf{A} = \mathbf{I}_{n_R} + P\sum_{l=1}^{K}\lambda_{l1}^\Sigma\mathbf{z}_l\mathbf{z}_l^\dagger$, $\mathbf{A}_k = \mathbf{A} - P\lambda_{k1}^\Sigma\mathbf{z}_k\mathbf{z}_k^\dagger$, λ_{ki} *is the* i^{th} *largest eigenvalue of the channel covariance matrix of user* k, *and* $\mathbf{z}_l$ *are* $n_R \times 1$ *dimensional i.i.d. Gaussian random vectors with zero-mean and identity-covariance.*

Proof (Theorem 2): The Lagrangian for the optimization problem in (17), with μ_k as the Lagrange multiplier of user k corresponding to its power constraint, is

$$E\left[\log\left|\mathbf{I}_{n_R} + \sum_{k=1}^{K}\sum_{i=1}^{n_T}\lambda_{ki}^Q\lambda_{ki}^\Sigma\mathbf{z}_{ki}\mathbf{z}_{ki}^\dagger\right|\right] - \sum_{k=1}^{K}\mu_k\left(\sum_{i=1}^{n_T}\lambda_{ki}^Q - P\right) \quad (20)$$

In order to derive the KKT conditions, we need the following identity which is proved in [4],

$$\frac{\partial}{\partial x}\log[\det(A + xB)] = \mathrm{tr}\left\{(A + xB)^{-1}B\right\} \quad (21)$$

Using this identity, the KKT conditions for user k are

$$\lambda_{ki}^\Sigma E\left[\mathbf{z}_{ki}^\dagger\left(\mathbf{I} + \sum_{l=1}^{K}\sum_{i=1}^{n_T}\lambda_{li}^Q\lambda_{li}^\Sigma\mathbf{z}_{li}\mathbf{z}_{li}^\dagger\right)^{-1}\mathbf{z}_{ki}\right] \leq \mu_k, i = 1,\ldots,n_T$$

where the conditions are satisfied with equality if the corresponding eigenvalue of the transmit covariance matrix is non-zero. Beamforming is optimal for all users, if the inequalities corresponding to $i = 1$ for $k = 1,\ldots,K$ are satisfied with equality, and the rest of the inequalities remain as strict inequalities. In this case, $\lambda_{k1}^Q = P$, for $k = 1,\ldots,K$, and all other eigenvalues of the transmit covariance matrices are zero. We have for user k,

$$E_{k1} = \lambda_{k1}^\Sigma E\left[\mathbf{z}_{k1}^\dagger\left(\mathbf{I} + P\sum_{l=1}^{K}\lambda_{l1}^\Sigma\mathbf{z}_{l1}\mathbf{z}_{l1}^\dagger\right)^{-1}\mathbf{z}_{k1}\right] = \mu_k \quad (22)$$

$$E_{ki} = \lambda_{ki}^\Sigma E\left[\mathbf{z}_{ki}^\dagger\left(\mathbf{I} + P\sum_{l=1}^{K}\lambda_{l1}^\Sigma\mathbf{z}_{l1}\mathbf{z}_{l1}^\dagger\right)^{-1}\mathbf{z}_{ki}\right] < \mu_k, \quad \forall i \neq 1$$

Equivalently, the conditions for the optimality of beamforming for all users are

$$E_{ki} - E_{k1} < 0, \quad \forall i \neq 1, \quad k = 1,\ldots,K \quad (23)$$

Due to the symmetry in these conditions, we will derive the condition for user k only. To proceed, we need the following lemma.

Lemma 1 *Let* $\mathbf{A} = \mathbf{I}_{n_R} + P\sum_{l=1}^{K} \lambda_{l1}^{\Sigma}\mathbf{z}_{l1}\mathbf{z}_{l1}^{\dagger}$*, and* $\mathbf{A}_k = \mathbf{A} - P\lambda_{k1}^{\Sigma}\mathbf{z}_{k1}\mathbf{z}_{k1}^{\dagger}$*, then the following identities hold*

$$\lambda_{k1}^{\Sigma} E\left[\mathbf{z}_{k1}^{\dagger}\mathbf{A}^{-1}\mathbf{z}_{k1}\right] = \frac{1}{P}\left(1 - E\left[\frac{1}{1+\gamma_k}\right]\right) \quad (24)$$

$$\lambda_{ki}^{\Sigma} E\left[\mathbf{z}_{ki}^{\dagger}\mathbf{A}^{-1}\mathbf{z}_{ki}\right] = \lambda_{ki}^{\Sigma}\left(n_R - K + \sum_{l=1}^{K} E\left[\frac{1}{1+\gamma_l}\right]\right), i \neq 1$$

where $\gamma_k = P\lambda_{k1}^{\Sigma}\mathbf{z}_{k1}^{T}\mathbf{A}_k^{-1}\mathbf{z}_{k1}$ *for* $k = 1, \ldots, K$.

A proof of Lemma 1 can be found in [11]. Using Lemma 1 and (23) for user k, we have

$$P\lambda_{ki}^{\Sigma} \leq \frac{1 - E\left[\frac{1}{1+\gamma_k}\right]}{n_R - K + \sum_{l=1}^{K} E\left[\frac{1}{1+\gamma_l}\right]}, \quad i = 2, \ldots, n_T \quad (25)$$

The left hand side is maximized for $i = 2$. Therefore, inserting $i = 2$ in (25) gives the condition in (19) for user k. Note that the Gaussian random vectors $\mathbf{z}_{k1}$ in (25) are denoted as $\mathbf{z}_k$ in (19) in order to simplify the notation. □

In order to have the optimality of beamforming, a combination of the largest eigenvalues of all users induce an upper bound on the second largest eigenvalues of all users. If the second largest eigenvalues of all users satisfy (19), then beamforming is optimal for all users. Inserting $K = 1$ in (19), we obtain the condition in (18), which is derived in [4].

In Figure 1, we plot two dimensional slices from the region corresponding to $K = 1, 2, 3, 5, 10$. These slices give the maximum possible λ_{12}^{Σ} for a range of λ_{11}^{Σ}. The largest eigenvalues of other users are kept constant. The number of receive antennas is $n_R = 2$. We see that the region where beamforming is optimal gets larger with the increasing number of users. Note that these curves have to lie below the $\lambda_{12}^{\Sigma} = \lambda_{11}^{\Sigma}$ line, because of the assumption that λ_{11}^{Σ} is the largest eigenvalue. However, we observe that the curves get closer to the $\lambda_{12}^{\Sigma} = \lambda_{11}^{\Sigma}$ line as K increases. This figure shows that with the addition of more and more users to the system, beamforming becomes optimal for more and more channel covariance matrices. This observation motivates us to investigate whether the growth of the region where beamforming is optimal is bounded, or beamforming is unconditionally optimal for very large numbers of users. This issue is addressed in [10].

IV. Mean Feedback at the Transmitters

V. Transmit Directions

In a single-user system with partial CSI in the form of the channel mean matrix at the transmitter, with the assumption that $\mathbf{H}_\mu$ is unit-rank, [4] showed that the optimal input covariance matrix $\mathbf{Q}$ can be written as

$$\mathbf{Q} = \mathbf{U}_\mu \mathbf{\Lambda}_Q \mathbf{U}_\mu^{\dagger} \quad (26)$$

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. Reference [4] proves this by using the results from [3].

In this paper, we show that, in a multi-user setting, every user should transmit along the eigenvectors of its own channel mean matrix. In showing this, we will follow another direction, and therefore the single-user case of our result will provide an alternative proof for Theorem 3 in [4]. In the multi-user setting, let the singular value decomposition of the channel mean matrix of user k be

$$\mathbf{H}_{\mu_k} = \mathbf{U}_{\mu_k}\mathbf{\Lambda}_{\mu_k}^{1/2}\mathbf{V}_{\mu_k}^{\dagger} \quad (27)$$

Since $\mathbf{H}_{\mu_k}$ 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 $k = 1, \ldots, K$. 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 $\mathbf{\Lambda}_{\mu_k}^{1/2}$ has only one non-zero element, which is $|\mathbf{a}_R||\mathbf{a}_{T_k}|$.

The main result of this section is contained in the following theorem which identifies the optimum transmit directions for all users. The single-user version of this theorem was proved in [7]. Here, we provide and prove a general version, valid for a multi-user system.

Theorem 3 *Let* $\mathbf{H}_{\mu_k} = \mathbf{U}_\mu\mathbf{\Lambda}_{\mu_k}^{1/2}\mathbf{V}_{\mu_k}^{\dagger}$ *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* $\mathbf{Q}_k = \mathbf{V}_{\mu_k}\mathbf{\Lambda}_k\mathbf{V}_{\mu_k}^{\dagger}$*, for all users.*

Proof (Theorem 3): We prove the theorem in two steps. In the first step, we show that the sum capacity resulting from $\{\mathbf{H}_{\mu_k}\}_{k=1}^{K}$ as the channel mean matrices and the sum capacity resulting from $\{\mathbf{\Lambda}_{\mu_k}^{1/2}\}_{k=1}^{K}$ as the channel mean matrices are the same.

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

$$C\left(\{\mathbf{H}_{\mu_k}\}_{k=1}^{K}\right) =$$

$$= \max_{\substack{\mathrm{tr}(\mathbf{Q}_k)\leq P\\ k=1\ldots K}} E\left[\log\left|\mathbf{I} + \sum_{k=1}^{K}(\mathbf{H}_{\mu_k}+\mathbf{Z}_k)\mathbf{Q}_k(\mathbf{H}_{\mu_k}+\mathbf{Z}_k)^{\dagger}\right|\right] \quad (28)$$

$$= \max_{\substack{\mathrm{tr}(\mathbf{Q}_k)\leq P\\ k=1\ldots K}} E\left[\log\left|\mathbf{I} + \sum_{k=1}^{K}(\mathbf{U}_\mu\mathbf{\Lambda}_{\mu_k}^{1/2}\mathbf{V}_{\mu_k}^{\dagger}+\mathbf{Z}_k)\mathbf{Q}_k (\mathbf{U}_\mu\mathbf{\Lambda}_{\mu_k}^{1/2}\mathbf{V}_{\mu_k}^{\dagger}+\mathbf{Z}_k)^{\dagger}\right|\right] \quad (29)$$

$$= \max_{\substack{\mathrm{tr}(\tilde{\mathbf{Q}}_k)\leq P\\ k=1\ldots K}} E\left[\log\left|\mathbf{I} + \sum_{k=1}^{K}(\mathbf{\Lambda}_{\mu_k}^{1/2}+\mathbf{Z}_k)\tilde{\mathbf{Q}}_k(\mathbf{\Lambda}_{\mu_k}^{1/2}+\mathbf{Z}_k)^{\dagger}\right|\right] \quad (30)$$

$$= C\left(\{\mathbf{\Lambda}_{\mu_k}^{1/2}\}_{k=1}^{K}\right) \quad (31)$$

where we inserted the singular value decomposition of the channel mean matrix of user k in (29), and used $\det(\mathbf{I}+\mathbf{AB}) = \det(\mathbf{I}+\mathbf{BA})$ to cancel $\mathbf{U}_\mu$, and the invariance of the distribution of zero-mean, identity-covariance matrix $\mathbf{Z}_k$ under unitary transformations in (30).

Note that $\mathrm{tr}(\mathbf{Q}_k) = \mathrm{tr}(\tilde{\mathbf{Q}}_k)$, since $\tilde{\mathbf{Q}}_k = \mathbf{V}_{\mu_k}^{\dagger}\mathbf{Q}_k\mathbf{V}_{\mu_k}$. By comparing (28) and (30), we see that the diagonal eigenvalue matrices of the channel mean matrices result in the same sum 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 (30) is diagonal. In order to prove this, we use the technique

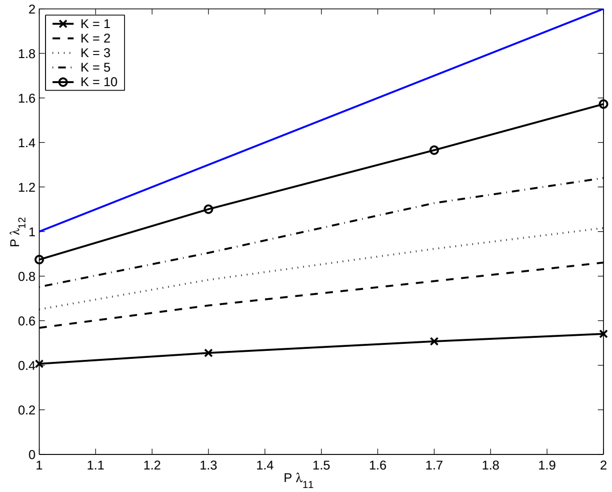

Fig. 1: The region where beamforming is optimal for different numbers of users in the covariance feedback model.

presented in [7]. Let $\mathbf{\Xi}$ be an $n_T \times n_T$ diagonal matrix, whose i^{th} diagonal entry is -1, and all other diagonal entries are 1. Let $\tilde{\mathbf{\Xi}}$ be an $n_R \times n_R$ diagonal matrix such that if $n_R < n_T$, then $\tilde{\mathbf{\Xi}} = \mathbf{I}_{n_R}$ and if $n_R > n_T$, then the i^{th} diagonal entry of $\tilde{\mathbf{\Xi}}$ is -1, and all other diagonal entries are 1. Then, we have

$$\tilde{\mathbf{\Xi}}\mathbf{\Lambda}_{\mu_k}^{1/2}\mathbf{\Xi} = \mathbf{\Lambda}_{\mu_k}^{1/2}, \quad k = 1, \ldots, K \tag{32}$$

Let us consider now a set of arbitrary transmit covariance matrices $\{\tilde{\mathbf{Q}}_k\}_{k=1}^K$, and define another set of transmit covariance matrices as $\hat{\mathbf{Q}}_k = \mathbf{\Xi}^\dagger \tilde{\mathbf{Q}}_k \mathbf{\Xi}$, for $k = 1, \ldots, K$. Note that the entries of $\hat{\mathbf{Q}}_k$ are equal to the entries of $\tilde{\mathbf{Q}}_k$ except that the off-diagonal entries in the i^{th} row and the i^{th} column have opposite signs. We can rewrite the optimization problem in (30) for a set of arbitrary transmit covariance matrices $\{\tilde{\mathbf{Q}}_k\}_{k=1}^K$ as

$$C(\{\tilde{\mathbf{Q}}_k\}_{k=1}^K) = \tag{33}$$

$$= E\left[\log\left|\mathbf{I}_{n_R} + \sum_{k=1}^K (\mathbf{\Lambda}_{\mu_k}^{1/2} + \mathbf{Z}_k)\tilde{\mathbf{Q}}_k(\mathbf{\Lambda}_{\mu_k}^{1/2} + \mathbf{Z}_k)^\dagger\right|\right] \tag{34}$$

$$= E\left[\log\left|\mathbf{I}_{n_R} + \sum_{k=1}^K (\mathbf{\Lambda}_{\mu_k}^{1/2} + \mathbf{Z}_k)\mathbf{\Xi}\hat{\mathbf{Q}}_k\mathbf{\Xi}^\dagger(\mathbf{\Lambda}_{\mu_k}^{1/2} + \mathbf{Z}_k)^\dagger\right|\right] \tag{35}$$

$$= E\left[\log\left|\mathbf{I}_{n_R} + \sum_{k=1}^K (\tilde{\mathbf{\Xi}}\mathbf{\Lambda}_{\mu_k}^{1/2}\mathbf{\Xi} + \mathbf{Z}_k)\hat{\mathbf{Q}}_k(\tilde{\mathbf{\Xi}}\mathbf{\Lambda}_{\mu_k}^{1/2}\mathbf{\Xi} + \mathbf{Z}_k)^\dagger\right|\right] \tag{36}$$

$$= E\left[\log\left|\mathbf{I}_{n_R} + \sum_{k=1}^K (\mathbf{\Lambda}_{\mu_k}^{1/2} + \mathbf{Z}_k)\hat{\mathbf{Q}}_k(\mathbf{\Lambda}_{\mu_k}^{1/2} + \mathbf{Z}_k)^\dagger\right|\right] \tag{37}$$

$$= C(\{\hat{\mathbf{Q}}_k\}_{k=1}^K) \tag{38}$$

where we again used the fact that $\mathbf{Z}_k$ and $\tilde{\mathbf{\Xi}}\mathbf{Z}_k\mathbf{\Xi}$ have the same distribution, and inserted (32) into (36) to obtain (42).

Now, let us define the set of transmit covariance matrices as $\mathbf{Q}_k^* = \frac{\tilde{\mathbf{Q}}_k + \hat{\mathbf{Q}}_k}{2}$, for $k = 1, \ldots, K$. The entries of $\mathbf{Q}_k^*$ are equal to the entries of $\tilde{\mathbf{Q}}_k$ except that the off-diagonal entries in the i^{th} row and the i^{th} column are zero. By the concavity of mutual information, it follows that the mutual information achieved by $\{\mathbf{Q}_k^*\}_{k=1}^K$ is greater than or equal to the mutual information achieved by $\{\tilde{\mathbf{Q}}_k\}_{k=1}^K$.

$$C(\{\mathbf{Q}_k^*\}_{k=1}^K) \geq \frac{1}{2}\left(C(\{\tilde{\mathbf{Q}}_k\}_{k=1}^K) + C(\{\hat{\mathbf{Q}}_k\}_{k=1}^K)\right) \tag{39}$$

$$= C(\{\tilde{\mathbf{Q}}_k\}_{k=1}^K) \tag{40}$$

Applying this procedure to every i for $1 \leq i \leq n_T$, we have shown that nulling the off-diagonal elements of transmit covariance matrices increases the capacity. This proves that the optimal $\tilde{\mathbf{Q}}_k$ is diagonal. This also proves the theorem since we have

$$\mathbf{Q}_k = \mathbf{V}_{\mu_k}\tilde{\mathbf{Q}}_k\mathbf{V}_{\mu_k}^\dagger = \mathbf{V}_{\mu_k}\mathbf{\Lambda}_k\mathbf{V}_{\mu_k}^\dagger \tag{41}$$

□

Therefore, using Theorem 3, we can write the optimization problem in (5) as,

$$C = \max_{\substack{\sum_{i=1}^{n_T}\lambda_{ki}^Q \leq P \\ k=1,\ldots,K}} E\left[\log\left|\mathbf{I}_{n_R} + \sum_{k=1}^K\sum_{i=1}^{n_T}\lambda_{ki}^Q\hat{\mathbf{z}}_{ki}\hat{\mathbf{z}}_{ki}\right|\right] \tag{42}$$

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

V.A Conditions for the Optimality of Beamforming

In this section, we analyze the conditions for the optimality of beamforming in a multi-user system with a finite number of users, when the partial CSI available at the transmitters is in the form of mean feedback. Reference [4] identified these conditions in a single-user system. Similar to the covariance feedback case, we find the conditions for the optimality of beamforming for all users in a multi-user setting. Inserting $K = 1$ in our results would provide an alternative proof for the condition in [4]. In our results, the number of conditions equals the number of users. The condition corresponding to user k depends on the non-zero eigenvalues of the channel mean matrices of all users.

Using Theorem 3, we can write the optimization problem in (28) and equivalently in (42) as

$$C_{sum} = \max_{\substack{\mathrm{tr}(\mathbf{\Lambda}_k) \leq P \\ k=1\ldots K}} E\left[\log\det\left(\mathbf{I}_{n_R} + \sum_{k=1}^K \hat{\mathbf{Z}}_k\mathbf{\Lambda}_k\hat{\mathbf{Z}}_k^\dagger\right)\right] \tag{43}$$

where $\hat{\mathbf{Z}}_k = \mathbf{\Lambda}_{\mu_k}^{1/2} + \mathbf{Z}_k$ is the equivalent channel matrix. Note that while the first column of the equivalent channel matrix is a non-zero mean Gaussian vector, all of the remaining columns are zero-mean Gaussian vectors. For this optimization problem, we have the following theorem.

Theorem 4 *In a MIMO-MAC system where the transmitters have partial CSI in the form of mean feedback, the transmit covariance matrices of all users that maximize (43) have unit-rank (i.e., beamforming is optimal for all users) if and only if*

$$P \leq \frac{1 - E\left[\frac{1}{1 + P\hat{\mathbf{z}}_k^\dagger\mathbf{B}_k^{-1}\hat{\mathbf{z}}_k}\right]}{n_R - K + \sum_{l=1}^K E\left[\frac{1}{1 + P\hat{\mathbf{z}}_l^\dagger\mathbf{B}_l^{-1}\hat{\mathbf{z}}_l}\right]}, \quad k = 1, \ldots, K \tag{44}$$

where $\mathbf{B} = \mathbf{I}_{n_R} + P\sum_{l=1}^K\hat{\mathbf{z}}_l\hat{\mathbf{z}}_l^\dagger$, $\mathbf{B}_k = \mathbf{B} - P\hat{\mathbf{z}}_k\hat{\mathbf{z}}_k^\dagger$, *and* $\hat{\mathbf{z}}_k = (\lambda_{k1}^\mu)^{1/2}\mathbf{e}_1 + \mathbf{z}_k$ *is the first column of the equivalent channel matrix,* $\hat{\mathbf{Z}}_k$.

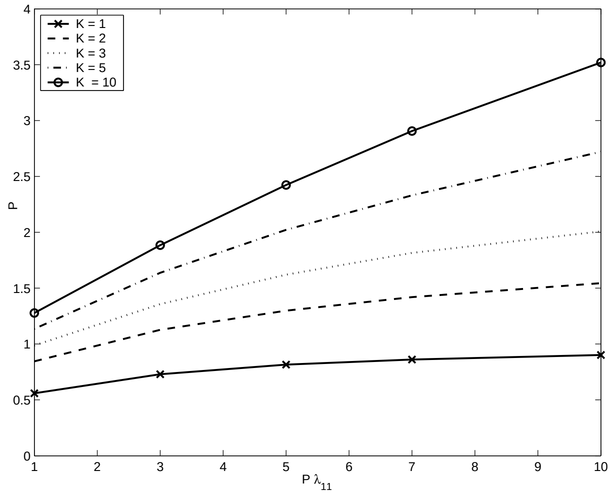

Fig. 2: The region where beamforming is optimal for different numbers of users in the mean feedback model.

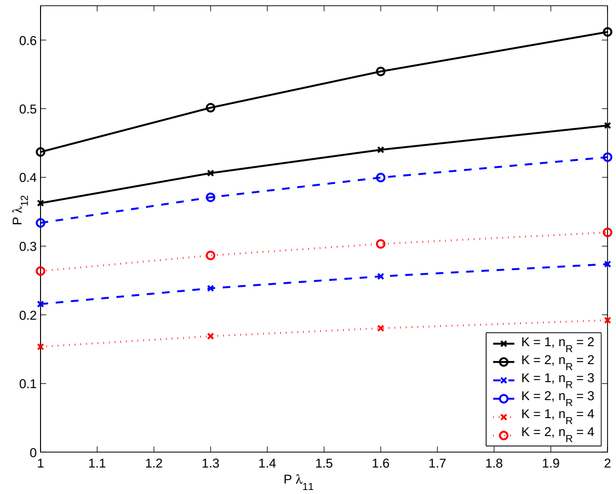

Fig. 3: The region where beamforming is optimal for different number of receive antennas in the covariance feedback model.

The proof of Theorem 4 is similar to the proof of Theorem 2, and is omitted here due to space restrictions [11].

The condition in (44) depends only on the non-zero eigenvalues of all users. If the powers and the eigenvalues of the feedback mean matrices of all users are such that they satisfy the inequality in (44), then beamforming is optimal for all users. Inserting $K = 1$ in (44), we obtain the condition proved in [4].

In Figure 2, we plot two dimensional slices from the region corresponding to $K = 1, 2, 3, 5, 10$. These slices give the possible power and λ_{11}^{μ} values for beamforming to be optimal. The largest eigenvalues of all other users are kept constant. The number of receive antennas is $n_R = 2$. Similar to the covariance feedback case, in the mean feedback case as well, we see that the region where beamforming is optimal gets larger with the increasing number of users. That is, with the addition of more and more users to the system, beamforming becomes optimal for more and more channel mean matrices [10].

We have analyzed the optimality of beamforming for different channel state assumptions for a finite sized system. We have shown that the region, in the channel state space, in which beamforming is optimal enlarges with the number of users. However, it is not immediate from these results that this region covers the entire channel state space of all users while the number of users grows to infinity. In [10], we show that for a large number of users, beamforming achieves a sum rate which approaches the optimal sum capacity. For asymptotic analysis, we utilize the strong law of large numbers. In a large system, even with the assumption that the transmitters have no knowledge of the channel, beamforming strategy turns out to be optimal [10].

In Figure 3, we illustrate the change in the region where beamforming is optimal with the number of receive antennas. We observe that the region gets smaller as the number of receive antennas is increased. However, for a fixed finite number of receive antennas, the region grows to the entire parameter region asymptotically as the number of users increases.

References

[1] İ. 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. Tranmitter 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] E. Jorswieck and H. Boche. Optimal transmission with imperfect channel state information at the transmit antenna array. *Wireless Personal Communications*, 27(1):33–56, 2003.

[7] D. Hösli and A. Lapidoth. The capacity of a MIMO Ricean channel is monotonic in the singular values of the mean. In 5^{th} *International ITG Conference on Source and Channel Coding*, January 2004.

[8] A. Lozano, F. R. Farrokhi, and R. A Valenzuela. Asymptotically optimal open-loop space-time architecture adaptive to scattering conditions. In *Proceedings of Vehicular Technology Conference*, Spring 2001.

[9] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications*. New York:Academic, 1979.

[10] A. Soysal and S. Ulukus. Asymptotic optimality of beamforming in MIMO-MAC with no or partial CSI at the transmitters. In *Vehicular Technology Conference*, May 2005.

[11] A. Soysal and S. Ulukus. Optimality of beamforming in fading MIMO multiple access channels. *To be submitted for journal publication.*