Optimality of Beamforming in Fading MIMO Multiple Access Channels

Alkan Soysal, *Member, IEEE,* and Sennur Ulukus, *Member, IEEE*

Abstract—We consider the sum capacity of 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 partial CSI, it is in the form of either the covariance matrix of the channel or the mean matrix of the channel. For the covariance feedback case, we mainly consider physical models that result in single-sided correlation structures. For the mean feedback case, we consider physical models that result in in-phase received signals. Under these assumptions, we analyze the MIMO-MAC from three different viewpoints. First, we consider a finite-sized system. 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. Also, we find the conditions under which beamforming is optimal for all users. Second, in the covariance feedback case, we prove that the region where beamforming is optimal for all users gets larger with the addition of new users into the system. In the mean feedback case, we show through simulations that this is not necessarily true. Third, we consider the asymptotic case where the number of users is large. 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 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 sum capacity. Finally, we generalize our covariance feedback results to double-sided correlation structures in the Appendix.

Index Terms—Multi-user MIMO, MIMO multiple access channel, partial CSI, covariance feedback, mean feedback, optimality of beamforming, large system analysis.

I. INTRODUCTION

THE use of multiple antennas at both transmitters and receivers in wireless communications promises very large information rates. In [3], Telatar showed that in a single-user system, when the transmitter does not know the state of the fading channel, the optimum transmit 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 [3], vector coding will result in lower probability of error but higher complexity as compared to parallel scalar coding, which already is very complex [4].

Beamforming is a scalar coding strategy in which the transmit 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 [3], where there is no CSI at the transmitters and the aim is to achieve the ergodic capacity, the optimum transmit covariance matrix is full-rank, and therefore beamforming is not optimal.

Although beamforming is not optimal for the no CSI case, it is shown by [5], [6], [7] for single-sided correlation structure, and by [8] for double-sided correlation structure that beamforming is conditionally optimal, in a single-user setting, when the transmitter has the partial knowledge of the channel. For the covariance feedback case, the fact that the optimal transmit covariance matrix and the channel covariance matrix have the same eigenvectors was shown in [5] for a multi-input single-output (MISO) system, and in [6] for a MIMO system. 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 [6], [7]. 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. In [8], these conditions are generalized to the case where the receive antennas are also correlated. For the case when beamforming is not optimal, algorithmic solutions are given in [9], [10], [11] in order to find the optimal power allocation strategy in a single-user MIMO system when the channel has only transmit antenna correlations. In [12], [13], we have proved the global convergence of the single-user algorithm proposed in [10], [11]. For the mean feedback case, 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 [5] and for a MIMO system in [6]. 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 [6].

In this paper, we consider the sum capacity point of a *multi-user* MIMO multiple access capacity region for three different CSI models, namely, no CSI model, partial CSI with covariance feedback model, and partial CSI with mean

feedback model. For the partial CSI with covariance feedback model, we consider physical scenarios that result in single-sided correlation structures. We generalize the covariance feedback model to double-sided correlation structures, where the receiver correlation is the same for all users, in the Appendix. For the partial CSI with mean feedback model, we consider physical scenarios that result in in-phase received signals, while the most general version of the problem is still open.

Under these channel assumptions for the MIMO-MAC, our contributions are three-fold: the analysis of a system with a finite number of users, determining the effects of increasing the number of users, and the analysis of a large system. First, we concentrate on a finite-sized system. We show that, if there is covariance or mean feedback information at the transmitters, all users should transmit in the direction of the eigenvectors of their *own* covariance or mean feedback matrices. Consequently, we show that, the transmit directions of the users are independent of the presence of other users, and therefore, that the users maintain their single-user transmit direction strategies even in a multi-user scenario[1]. Then, we identify the necessary and sufficient conditions for the optimality of beamforming for all users. This result generalizes the single-user conditions of [6], [7] 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 these conditions are satisfied, 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, and they form a region in a space whose dimension is equal to the number of users. Similarly, if these conditions are satisfied, beamforming is optimal for all users.

We, then, consider the effects of increasing the number of users on the region of channel parameters where beamforming is optimal. Here, in the covariance feedback case, we prove that this region gets larger as new users are added to the system. Although adding users increases the overall complexity of the system, being able to beamform for a greater range of channel values decreases the complexity. In the mean feedback case, this result does not necessarily hold. Nevertheless, we see through simulations that as the number of users gets large enough, the region where beamforming is optimal grows with the addition of new users, in the mean feedback case as well. These results raise the question of whether the region where beamforming is optimal spans the entire parameter space as the number of users grows to infinity. Therefore, next, we analyze our problem from an asymptotically large system viewpoint.

The optimality of beamforming in a MIMO-MAC system where the channel is deterministic and fully known to the transmitters is investigated in [15], where it was shown that if the number of users is much larger than the number of receive antennas, then unit-rank transmission is optimal for almost all users. Motivated by our result described above and

the result of [15] that beamforming is optimal asymptotically (with respect to the number of users) in a deterministic multi-user MIMO-MAC, we ask the question whether beamforming is unconditionally optimal asymptotically in our case as well, where the receiver has the perfect CSI, but the transmitters have no or partial CSI. When there is no CSI at the transmitters, it is counter-intuitive to think that beamforming would be optimal. Confirming this intuition, [3] already showed that in a finite-sized multi-user system with no CSI at the transmitters, the optimum transmit covariance matrices are full-rank for all users. However, we show that, in an asymptotically large system, unit-rank transmit covariance matrices are optimal for all users. The beamforming scheme we use in this case is simpler than usual; it may be characterized as an arbitrary antenna selection scheme, where for each user, only one antenna is used for transmission and that antenna is chosen arbitrarily.

When the transmitters have the partial CSI in the form of either covariance or mean feedback, we show that the asymptotic optimality of beamforming still holds. In these cases however, arbitrary antenna selection scheme is no longer optimal. In the covariance feedback setting, each user beamforms in the direction of the strongest eigenvector of its channel feedback covariance matrix. As opposed to a finite-sized system, where beamforming may or may not be optimal depending on the eigenvalues of the channel covariance matrices, we show here that for an asymptotically large system, beamforming is always optimal. In the mean feedback setting, each user beamforms in the direction of the eigenvector corresponding to the sole non-zero eigenvalue of its channel feedback mean matrix. Similar to the covariance feedback case, beamforming is optimal asymptotically irrespective of the values of the mean feedback information. Asymptotic analysis has been used in the literature before, e.g., by [16], [17], [18], where it yielded simple characterizations to complex systems. In our model, with multiple users, with multiple transmit and receive antennas, and with fading in the channel, the optimal transmit strategy turns out to be simple beamforming, when only the number of users goes to infinity.

Considering all three points of view, this paper provides a complete extension from single-user to multi-user systems with finite and infinite numbers of users, including the transient behavior of the system with increasing number of users, for MIMO systems with partial CSI at the transmitters in the form of covariance and mean information.

II. SYSTEM MODEL

We consider a 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[2]. The receiver has the perfect knowledge of the channel, while the transmitters have only the statistical model of the channel. Each transmitter

---

[1]During the review process of this paper, we were informed of [14], where for the covariance feedback case, this result is stated without proof.

[2]Although we consider the case where all transmitters have the same number of antennas, our results immediately extend to the cases where the transmitters have different number of antennas.

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_k, $\mathrm{tr}(\mathbf{Q}_k) \leq P_k$.

We investigate three different statistical models. The first one is the "no CSI" model in which the transmitters only know the distribution of the channel state 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 [3], [4], [16].

The second model is the "partial CSI with covariance feedback" model where each transmitter knows the channel covariance information of all transmitters, in addition to the distribution of the channel [19]. In this paper, we assume that the receiver (e.g., base station) does not have physical restrictions and therefore, there is sufficient spacing between the antenna elements on the receiver. If the minimum antenna spacing is sufficiently large, the correlation introduced by antenna element spacing is low enough that the fades associated with two different antenna elements can be considered independent[3] [20]. We also assume that the signals transmitted by different antenna elements are correlated, because of the lack of scatterers around the transmitters. The channel of user k is written as

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

where the entries of $\mathbf{Z}_k$ are i.i.d., zero-mean, unit-variance complex Gaussian random variables, and the channel covariance feedback matrix, $\boldsymbol{\Sigma}_k$, is the correlation between the signals transmitted from the n_T transmit antennas of user k. While writing (2), we separately apply the single-user model in [19] to every single transmitter-receiver link. Similar covariance feedback models are used in [5], [6], [7], [9] in the single-user setting. From this point on, we will refer to matrix $\boldsymbol{\Sigma}_k$ as the channel covariance feedback matrix of user k.

The third model we investigate is the "partial CSI with mean feedback" model where each transmitter knows the channel mean information of all transmitters, in addition to the distribution of the channel. This model is used in [5], [6], [21], [22], [23]. In this model, the transmitters have line-of-sight component with the receiver. 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

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

where the 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. This Ricean channel is modeled to be of unit-rank

---

[3]We refer the reader to the Appendix, for the extension of our results to the case where the channel has double-sided correlation structure, i.e., to the case where the signals arriving at the receiver are correlated as well.

[23], and therefore, the mean matrix takes the form

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

where $\mathbf{a}_{R_k}$ and $\mathbf{a}_{T_k}$ are the array response vectors at the receiver and the transmitter, respectively. In this most general case of the mean feedback model, the optimization problem that arises in the sum capacity calculation seems intractable. In order to simplify the mathematics and obtain a tractable optimization problem, we assume that the user signals arrive at the base station in-phase, i.e., $\mathbf{a}_{R_k} = \mathbf{a}_R$, for all k. This mathematical simplification models a physical system where the transmitters are far away from the receiver and are close to each other. This can occur, for instance, if a set of closely located transmitters have a line-of-sight "openning" with the receiver. From this point on, we will refer to matrix $\mathbf{H}_{\mu_k}$ as the channel mean matrix of user k.

III. FINITE SYSTEM ANALYSIS

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

$$C_{sum} = \max_{\substack{\mathrm{tr}(\mathbf{Q}_k) \leq P_k, \mathbf{Q}_k \succeq \mathbf{0} \\ k=1\ldots K}} E\left[\log\left|\mathbf{I}_{n_R} + \sum_{k=1}^{K} \mathbf{H}_k \mathbf{Q}_k \mathbf{H}_k^\dagger\right|\right] \tag{5}$$

where $E[\cdot]$ is the expectation operator with respect to the channel matrices of all users conditioned on the covariance or mean feedback, $|\cdot|$ is the determinant operator, $\mathrm{tr}(\cdot)$ denotes the trace of a matrix, and $\mathbf{Q}_k \succeq \mathbf{0}$ denotes positive semi-definite $\mathbf{Q}_k$. In this section, we will find the optimum transmit directions of the users, and the region where beamforming is optimal for all users, under various assumptions on the CSI available at the transmitters, for a multi-user MIMO-MAC with a finite number of users.

For a single-user system with no CSI at the transmitter and identity channel covariance matrix, i.e., $\boldsymbol{\Sigma} = \mathbf{I}$, Telatar [3] showed that the capacity is achieved when the transmitter divides its power equally over its antennas, i.e., the optimal transmitter covariance matrix, $\mathbf{Q}$, is equal to $(P/n_T)\mathbf{I}$. Clearly, in this setting, beamforming is not optimal, as the transmit covariance matrix is full-rank. For the multi-user case, [3] defines a stacked channel matrix as $\hat{\mathbf{H}} = [\mathbf{H}_1, \cdots, \mathbf{H}_K]$ and writes the sum capacity as

$$C_{sum} = E\left[\log\left|\mathbf{I}_{n_R} + \frac{KP}{Kn_T}\hat{\mathbf{H}}\hat{\mathbf{H}}^\dagger\right|\right]$$
$$= E\left[\log\left|\mathbf{I}_{n_R} + \frac{P}{n_T}\sum_{k=1}^{K} \mathbf{H}_k \mathbf{H}_k^\dagger\right|\right] \tag{6}$$

This means that in the multi-user setting as well, the sum capacity maximizing transmit covariance matrix for each user is proportional to identity, i.e., $\mathbf{Q}_k = (P/n_T)\mathbf{I}$, for all k. Therefore, it is clear that, beamforming is not optimal for any user in a finite-sized multi-user system when the transmitters do not have any CSI.

A. 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 transmit

covariance matrix. In this case, the problem becomes that of choosing a transmit covariance matrix $\mathbf{Q}$, which is subject to a trace constraint representing the average transmit power constraint,

$$C = \max_{\mathrm{tr}(\mathbf{Q}) \le P, \mathbf{Q} \succeq \mathbf{0}} E\left[\log | \mathbf{I}_{n_R} + \mathbf{H}\mathbf{Q}\mathbf{H}^{\dagger}|\right] \qquad (7)$$

The channel covariance matrix $\mathbf{\Sigma}$, which is known at the transmitter, and the transmit covariance matrix $\mathbf{Q}$ have the eigenvalue decompositions $\mathbf{\Sigma} = \mathbf{U}_{\Sigma}\mathbf{\Lambda}_{\Sigma}\mathbf{U}_{\Sigma}^{\dagger}$, and $\mathbf{Q} = \mathbf{U}_{Q}\mathbf{\Lambda}_{Q}\mathbf{U}_{Q}^{\dagger}$, respectively. Here, $\mathbf{\Lambda}_{\Sigma}$ and $\mathbf{\Lambda}_{Q}$ are the diagonal matrices of ordered eigenvalues of $\mathbf{\Sigma}$, and $\mathbf{Q}$, and $\mathbf{U}_{\Sigma}$, and $\mathbf{U}_{Q}$ are unitary matrices.

Reference [6], [24] showed that the eigenvectors of the transmit covariance matrix must be equal to the eigenvectors of the channel covariance matrix, i.e., $\mathbf{U}_{Q} = \mathbf{U}_{\Sigma}$. References [9], [10], [13] proposed numerical optimization based methods to find $\mathbf{\Lambda}_{Q}$, the power put by the transmitter along the eigen-directions. References [6] and [7] 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 this paper, for a multi-user setting with a finite number of users, where there is covariance feedback at the transmitters, we prove 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 spectral decomposition of the channel covariance matrix of user k. Then, the optimum transmit covariance matrix $\mathbf{Q}_k$ of user k has the form $\mathbf{Q}_k = \mathbf{U}_{\Sigma_k}\mathbf{\Lambda}_{Q_k}\mathbf{U}_{\Sigma_k}^{\dagger}$, for all users.

Single-user versions of this theorem can be found in [6], [24], and the multi-user version is stated without proof in [14]. The proof follows [6] closely, and therefore is omitted here. It can be found in [25].

The optimization problem in (5) with channel matrices $\mathbf{H}_k = \mathbf{Z}_k\mathbf{\Sigma}_k^{1/2} = \mathbf{Z}_k\mathbf{U}_{\Sigma_k}\mathbf{\Lambda}_{\Sigma_k}^{1/2}\mathbf{U}_{\Sigma_k}^{\dagger}$ can be written as

$$C_{sum} = \max_{\substack{\mathrm{tr}(\mathbf{Q}_k) \le P_k, \mathbf{Q}_k \succeq \mathbf{0} \\ k=1,\ldots,K}} E\left[\log \left|\mathbf{I}_{n_R} + \sum_{k=1}^{K} \mathbf{Z}_k\mathbf{U}_{\Sigma_k}\mathbf{\Lambda}_{\Sigma_k}^{1/2}\mathbf{U}_{\Sigma_k}^{\dagger}\right.\right.$$
$$\left.\left. \mathbf{Q}_k\mathbf{U}_{\Sigma_k}\mathbf{\Lambda}_{\Sigma_k}^{1/2}\mathbf{U}_{\Sigma_k}^{\dagger}\mathbf{Z}_k^{\dagger}\right|\right] \qquad (8)$$

Using Theorem 1, we can write the optimization problem in (8) as,

$$C_{sum} = \max_{\substack{\mathrm{tr}(\mathbf{\Lambda}_{Q_k}) \le P_k, \mathbf{\Lambda}_{Q_k} \succeq \mathbf{0} \\ k=1,\ldots,K}} E\left[\log \left|\mathbf{I}_{n_R} + \sum_{k=1}^{K} \mathbf{Z}_k\mathbf{\Lambda}_{Q_k}\mathbf{\Lambda}_{\Sigma_k}\mathbf{Z}_k^{\dagger}\right|\right] \qquad (9)$$

$$= \max_{\substack{\sum_{i=1}^{n_T} \lambda_{ki}^{Q} \le P_k, \lambda_{ki}^{Q} \ge 0 \\ k=1,\ldots,K}} E\left[\log \left|\mathbf{I}_{n_R} + \sum_{k=1}^{K}\sum_{i=1}^{n_T} \lambda_{ki}^{Q}\lambda_{ki}^{\Sigma}\mathbf{z}_{ki}\mathbf{z}_{ki}^{\dagger}\right|\right] \qquad (10)$$

where $\mathbf{z}_{ki}$ is the ith column of $\mathbf{Z}_k$, i.e., $\{\mathbf{z}_{ki}, k = 1,\ldots,K, i = 1,\ldots,n_T\}$ is a set of $n_R \times 1$ dimensional i.i.d., zero-mean, identity-covariance Gaussian random vectors.

In a MIMO system, a transmit strategy is a combination of a transmit direction strategy and a transmit power allocation strategy. A result of Theorem 1 is that the optimal multi-user transmit direction strategies are decoupled into a set of single-user transmit direction strategies. However, in general, this is not true for the optimal transmit power allocation strategies. The amount of power each user allocates in each direction depends on both the transmit directions and the power allocations of other users [13]. Because of this, finding the conditions under which beamforming is optimal becomes even more critical in the multi-user case. When beamforming is optimal, the optimal transmit power allocation strategy for each user reduces to allocating all of its power to its strongest eigen-direction, and this strategy does not require the user to know the channel covariance matrices of the other users.

2) 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 [6] and [7] 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 [6]:

$$P\lambda_2^{\Sigma} < \frac{1 - E\left[\frac{1}{1 + P\lambda_1^{\Sigma}\mathbf{z}^T\mathbf{z}}\right]}{n_R - 1 + E\left[\frac{1}{1 + P\lambda_1^{\Sigma}\mathbf{z}^T\mathbf{z}}\right]} \qquad (11)$$

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 necessary and sufficient conditions for the optimality of beamforming for all users in a multi-user setting. Inserting $K = 1$ in our results would reduce them to (11). In our results, the number of conditions equals 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. Before stating our theorem in this section, we need the following lemma.

Lemma 1: When $\mathbf{A}$ and $\mathbf{A}_k$ are defined as in Theorem 2, the following identities hold

$$E_{k1} = \lambda_{k1}^{\Sigma} E\left[\mathbf{z}_{k1}^{\dagger}\mathbf{A}^{-1}\mathbf{z}_{k1}\right]$$
$$= \frac{1}{P_k}\left(1 - E\left[\frac{1}{1 + P_k\lambda_{k1}^{\Sigma}\mathbf{z}_{k1}^T\mathbf{A}_k^{-1}\mathbf{z}_{k1}}\right]\right) \qquad (12)$$
$$E_{ki} = \lambda_{ki}^{\Sigma} E\left[\mathbf{z}_{ki}^{\dagger}\mathbf{A}^{-1}\mathbf{z}_{ki}\right], \quad i \ne 1$$
$$= \lambda_{ki}^{\Sigma}\left(n_R - K + \sum_{l=1}^{K} E\left[\frac{1}{1 + P_l\lambda_{l1}^{\Sigma}\mathbf{z}_{l1}^T\mathbf{A}_l^{-1}\mathbf{z}_{l1}}\right]\right) \qquad (13)$$

A proof of Lemma 1 is given in Appendix A.

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 (10) have unit-rank (i.e., beamforming is optimal for all users) if

and only if

$$P_k \lambda_{k2}^\Sigma < \frac{1 - E\left[\frac{1}{1 + P_k \lambda_{k1}^\Sigma \mathbf{z}_{k1} \mathbf{A}_k^{-1} \mathbf{z}_{k1}}\right]}{n_R - K + \sum_{l=1}^K E\left[\frac{1}{1 + P_l \lambda_{l1}^\Sigma \mathbf{z}_{l1} \mathbf{A}_l^{-1} \mathbf{z}_{l1}}\right]} \quad (14)$$

for $k = 1, \ldots, K$, and where $\mathbf{A} = \mathbf{I}_{n_R} + \sum_{l=1}^K P_l \lambda_{l1}^\Sigma \mathbf{z}_{l1} \mathbf{z}_{l1}^\dagger$, $\mathbf{A}_k = \mathbf{A} - P_k \lambda_{k1}^\Sigma \mathbf{z}_{k1} \mathbf{z}_{k1}^\dagger$, λ_{ki}^Σ is the ith largest eigenvalue of the channel covariance matrix of user k, and $\mathbf{z}_{l1}$ 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 (10), with μ_k as the Lagrange multiplier of user k corresponding to its power constraint, is

$$E\left[\log\left|\mathbf{I}_{n_R} + \sum_{k=1}^K \sum_{i=1}^{n_T} \lambda_{ki}^Q \lambda_{ki}^\Sigma \mathbf{z}_{ki} \mathbf{z}_{ki}^\dagger\right|\right] - \sum_{k=1}^K \mu_k \left(\sum_{i=1}^{n_T} \lambda_{ki}^Q - P_k\right) \quad (15)$$

In order to derive the Karush-Kuhn-Tucker (KKT) conditions, we need the following identity which is proved in [6],

$$\frac{\partial}{\partial x} \log|\mathbf{A} + x\mathbf{B}| = \operatorname{tr}\left[(\mathbf{A} + x\mathbf{B})^{-1}\mathbf{B}\right] \quad (16)$$

Using this identity, the KKT conditions for user k are

$$\lambda_{ki}^\Sigma E\left[\mathbf{z}_{ki}^\dagger \left(\mathbf{I} + \sum_{l=1}^K \sum_{i=1}^{n_T} \lambda_{li}^Q \lambda_{li}^\Sigma \mathbf{z}_{li} \mathbf{z}_{li}^\dagger\right)^{-1} \mathbf{z}_{ki}\right] \leq \mu_k \quad (17)$$

for $i = 1, \ldots, n_T$. 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, $\lambda_{k1}^Q = P_k$, for $k = 1, \ldots, K$, and all other eigenvalues of the transmit covariance matrices are zero. We have the following for user k,

$$E_{k1} = \lambda_{k1}^\Sigma E\left[\mathbf{z}_{k1}^\dagger \mathbf{A}^{-1} \mathbf{z}_{k1}\right] = \mu_k \quad (18)$$

$$E_{ki} = \lambda_{ki}^\Sigma E\left[\mathbf{z}_{ki}^\dagger \mathbf{A}^{-1} \mathbf{z}_{ki}\right] < \mu_k, \quad \forall i \neq 1 \quad (19)$$

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

$$\frac{E_{k1}}{E_{ki}} > 1, \quad \forall i \neq 1, \quad k = 1, \ldots, K \quad (20)$$

Due to the symmetry in these conditions, we will derive the condition for user k only. Using Lemma 1 and (20) for user k, we have

$$P_k \lambda_{ki}^\Sigma < \frac{1 - E\left[\frac{1}{1 + P_k \lambda_{k1}^\Sigma \mathbf{z}_{k1}^T \mathbf{A}_k^{-1} \mathbf{z}_{k1}}\right]}{n_R - K + \sum_{l=1}^K E\left[\frac{1}{1 + P_l \lambda_{l1}^\Sigma \mathbf{z}_{l1}^T \mathbf{A}_l^{-1} \mathbf{z}_{l1}}\right]} \quad (21)$$

for $i = 2, \ldots, n_T$. Note that the left hand side is maximized for $i = 2$, that is, if the condition for $i = 2$ holds, then it holds for all other i, as well. Therefore, inserting $i = 2$ in (21) gives the condition in (14) for user k. $\square$

Note that inserting $K = 1$ in (14), we obtain the condition in (11), which is derived in [6]. In our case, the right hand side of (14) depends only on the largest eigenvalues of all users.

Therefore, 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 (14), then beamforming is optimal for all users.

One important issue in the analysis of the region where beamforming is optimal, is the change in the region with varying numbers of users. In the next theorem, we show that the region where beamforming is optimal grows with the addition of new users into the system.

Theorem 3: In a MIMO-MAC system where the transmitters have partial CSI in the form of covariance feedback, the region where beamforming is optimal gets larger by the addition of new users.

Proof (Theorem 3): From (20), beamforming is optimal for all users if and only if

$$\lambda_{k2}^\Sigma < \frac{\lambda_{k1}^\Sigma E\left[\mathbf{z}_{k1}^\dagger \mathbf{A}^{-1} \mathbf{z}_{k1}\right]}{E\left[\mathbf{z}_{k2}^\dagger \mathbf{A}^{-1} \mathbf{z}_{k2}\right]}, \quad k = 1, \ldots, K \quad (22)$$

Note that $\mathbf{z}_{k2}$ is independent of $\mathbf{A}$, and has identity covariance. Therefore, the denominator of the right hand side of (22) becomes $E[\operatorname{tr}(\mathbf{A}^{-1})]$. Let us define the "boundary function" $f_k(\boldsymbol{\lambda})$ as

$$f_k(\boldsymbol{\lambda}) = \frac{\lambda_{k1}^\Sigma E[\mathbf{z}_{k1}^\dagger \mathbf{A}^{-1} \mathbf{z}_{k1}]}{E[\operatorname{tr}(\mathbf{A}^{-1})]}, \quad k = 1, \ldots, K \quad (23)$$

where $\boldsymbol{\lambda} = [\lambda_{11}^\Sigma, \lambda_{21}^\Sigma, \ldots, \lambda_{K1}^\Sigma]^T$ contains the largest eigenvalues of the covariance feedback matrices of all users. Then, beamforming is optimal for all users if and only if

$$\lambda_{k2}^\Sigma < f_k(\boldsymbol{\lambda}), \quad k = 1, \ldots, K \quad (24)$$

We will show that, $f_k(\boldsymbol{\lambda})$ increases in every component of the vector $\boldsymbol{\lambda}$, for all k. This will prove that, when a user is added to the system, i.e., the eigenvalue of the corresponding user is increased to a positive number from zero, the region in which beamforming is optimal for all users increases as long as the the condition for the new user is also satisfied. In order to prove that each $f_k(\boldsymbol{\lambda})$ increases in $\boldsymbol{\lambda}$, we will prove that every component of the vector of boundary functions, $\mathbf{f}(\boldsymbol{\lambda}) = [f_1(\boldsymbol{\lambda}), \ldots, f_K(\boldsymbol{\lambda})]^T$, increases in $\boldsymbol{\lambda}$. Let us define $\bar{\mathbf{Z}} = [\mathbf{z}_{11}, \mathbf{z}_{21}, \ldots, \mathbf{z}_{K1}]$, and $\bar{\boldsymbol{\Lambda}}$ and $\bar{\mathbf{P}}$ as diagonal matrices having $\{\lambda_{11}^\Sigma, \lambda_{21}^\Sigma, \ldots, \lambda_{K1}^\Sigma\}$ and $\{P_1, P_2, \ldots, P_K\}$ along their diagonals, respectively. Then,

$$\mathbf{f}(\boldsymbol{\lambda}) = \frac{\operatorname{diag}\left\{E\left[\bar{\boldsymbol{\Lambda}}^{1/2} \bar{\mathbf{Z}}^\dagger \mathbf{A}^{-1} \bar{\mathbf{Z}} \bar{\boldsymbol{\Lambda}}^{1/2}\right]\right\}}{E\left[\operatorname{tr}(\mathbf{A}^{-1})\right]} \quad (25)$$

where $\operatorname{diag}\{\cdot\}$ is the vector composed of the diagonal elements of its argument, and $\mathbf{A}$ can be expressed in terms of $\bar{\mathbf{Z}}$, $\bar{\mathbf{P}}$, and $\bar{\boldsymbol{\Lambda}}$ as $\mathbf{A} = \mathbf{I} + \bar{\mathbf{Z}} \bar{\mathbf{P}} \bar{\boldsymbol{\Lambda}} \bar{\mathbf{Z}}^\dagger$. Note that the expectation of the $(k, l)^{th}$ off-diagonal element of the random matrix $\bar{\boldsymbol{\Lambda}}^{1/2} \bar{\mathbf{Z}}^\dagger \mathbf{A}^{-1} \bar{\mathbf{Z}} \bar{\boldsymbol{\Lambda}}^{1/2}$ is zero. The reason for this is that when the expectation is expressed as an integral, the contribution to the integral at $\mathbf{z}_{l1}$ is cancelled by the contribution at $-\mathbf{z}_{l1}$, due to the odd function property of $\lambda_{k1}^\Sigma \mathbf{z}_{k1}^\dagger \mathbf{A}^{-1} \mathbf{z}_{l1}$. Note also that, since $\lambda_{l1}^\Sigma \mathbf{z}_{l1} \mathbf{z}_{l1}^\dagger = \lambda_{l1}^\Sigma(-\mathbf{z}_{l1})(-\mathbf{z}_{l1})^\dagger$, the matrix $\mathbf{A}$ and the value of the probability density function are the same for $\mathbf{z}_{l1}$ and $-\mathbf{z}_{l1}$. Hence, we conclude that $E[\bar{\boldsymbol{\Lambda}}^{1/2} \bar{\mathbf{Z}}^\dagger \mathbf{A}^{-1} \bar{\mathbf{Z}} \bar{\boldsymbol{\Lambda}}^{1/2}]$

is diagonal, and therefore its diagonal elements are the same as its eigenvalues.

Now, we will show that the eigenvalues of $E[\bar{\boldsymbol{\Lambda}}^{1/2}\bar{\mathbf{Z}}^\dagger\mathbf{A}^{-1}\bar{\mathbf{Z}}\bar{\boldsymbol{\Lambda}}^{1/2}]$ increase in $\boldsymbol{\lambda}$. We will show this in two steps. First, we will show that the eigenvalues of $\bar{\mathbf{P}}\bar{\boldsymbol{\Lambda}}^{1/2}\bar{\mathbf{Z}}^\dagger\mathbf{A}^{-1}\bar{\mathbf{Z}}\bar{\boldsymbol{\Lambda}}^{1/2}$, for any given realization of $\bar{\mathbf{Z}}$, increase in $\boldsymbol{\lambda}$, and then we will show that the eigenvalues of $E[\bar{\mathbf{P}}\bar{\boldsymbol{\Lambda}}^{1/2}\bar{\mathbf{Z}}^\dagger\mathbf{A}^{-1}\bar{\mathbf{Z}}\bar{\boldsymbol{\Lambda}}^{1/2}]$ increase in $\boldsymbol{\lambda}$. This immediately implies that the eigenvalues of $E[\bar{\boldsymbol{\Lambda}}^{1/2}\bar{\mathbf{Z}}^\dagger\mathbf{A}^{-1}\bar{\mathbf{Z}}\bar{\boldsymbol{\Lambda}}^{1/2}]$ increase in $\boldsymbol{\lambda}$. First consider a fixed realization of the random matrix $\bar{\mathbf{Z}}$. Note that,

$$\bar{\mathbf{Z}}\bar{\mathbf{P}}\bar{\boldsymbol{\Lambda}}\bar{\mathbf{Z}}^\dagger = \sum_{k=1}^{K} P_k \lambda_{k1}^{\Sigma} \mathbf{z}_{k1}\mathbf{z}_{k1}^\dagger \qquad (26)$$

If we increase any one of λ_{k1}^{Σ}, $k = 1, \ldots, K$, to $(\lambda_{k1}^{\Sigma})'$, this can be seen as an addition of the positive semidefinite matrix $P_k\left((\lambda_{k1}^{\Sigma})' - \lambda_{k1}^{\Sigma}\right)\mathbf{z}_{k1}\mathbf{z}_{k1}^\dagger$ to the summation in (26). Using the corollary to Weyl's monotonicity theorem [26, page 181-182] which states that all eigenvalues of a Hermitian matrix increase if a positive semidefinite matrix is added to it, we can conclude that the eigenvalues of $\bar{\mathbf{Z}}\bar{\mathbf{P}}\bar{\boldsymbol{\Lambda}}\bar{\mathbf{Z}}^\dagger$ increase in $\boldsymbol{\lambda}$ for any fixed $\bar{\mathbf{Z}}$. Now, note that, if we denote the eigenvalues of $\bar{\mathbf{Z}}\bar{\mathbf{P}}\bar{\boldsymbol{\Lambda}}\bar{\mathbf{Z}}^\dagger$ as α_i, then the eigenvalues of $\mathbf{A}^{-1}\bar{\mathbf{Z}}\bar{\mathbf{P}}\bar{\boldsymbol{\Lambda}}\bar{\mathbf{Z}}^\dagger$ are given by $\frac{\alpha_i}{1+\alpha_i}$. Further, the eigenvalues of $\bar{\mathbf{P}}\bar{\boldsymbol{\Lambda}}^{1/2}\bar{\mathbf{Z}}^\dagger\mathbf{A}^{-1}\bar{\mathbf{Z}}\bar{\boldsymbol{\Lambda}}^{1/2}$ are either $\frac{\alpha_i}{1+\alpha_i}$ or 0, depending on the dimensions of $\bar{\mathbf{Z}}$. Therefore, we conclude that when $\boldsymbol{\lambda}$ increases, all α_i increase as shown above, and therefore all $\frac{\alpha_i}{1+\alpha_i}$ increase as well.

Until now, we have shown that the eigenvalues of the random matrix $\bar{\mathbf{P}}\bar{\boldsymbol{\Lambda}}^{1/2}\bar{\mathbf{Z}}^\dagger\mathbf{A}^{-1}\bar{\mathbf{Z}}\bar{\boldsymbol{\Lambda}}^{1/2}$ increase in $\boldsymbol{\lambda}$. Next, we will show that the eigenvalues of the expectation $E[\bar{\mathbf{P}}\bar{\boldsymbol{\Lambda}}^{1/2}\bar{\mathbf{Z}}^\dagger\mathbf{A}^{-1}\bar{\mathbf{Z}}\bar{\boldsymbol{\Lambda}}^{1/2}]$ increase in $\boldsymbol{\lambda}$ as well. We note that this expectation can be written as a positive weighted sum of positive semidefinite Hermitian matrices $\bar{\mathbf{P}}\bar{\boldsymbol{\Lambda}}^{1/2}\bar{\mathbf{Z}}^\dagger\mathbf{A}^{-1}\bar{\mathbf{Z}}\bar{\boldsymbol{\Lambda}}^{1/2}$ for all realizations of the random matrix $\bar{\mathbf{Z}}$. An increase in $\boldsymbol{\lambda}$, can again be seen as an addition of a positive semidefinite matrix to the expectation. Therefore, invoking the corollary to Weyl's monotonicity theorem [26, page 181-182] once again, we conclude that the eigenvalues of $E[\bar{\mathbf{P}}\bar{\boldsymbol{\Lambda}}^{1/2}\bar{\mathbf{Z}}^\dagger\mathbf{A}^{-1}\bar{\mathbf{Z}}\bar{\boldsymbol{\Lambda}}^{1/2}]$, and consequently the eigenvalues of $E[\bar{\boldsymbol{\Lambda}}^{1/2}\bar{\mathbf{Z}}^\dagger\mathbf{A}^{-1}\bar{\mathbf{Z}}\bar{\boldsymbol{\Lambda}}^{1/2}]$ increase in $\boldsymbol{\lambda}$. We also note that $E\left[\text{tr}(\mathbf{A}^{-1})\right]$ decreases in $\boldsymbol{\lambda}$, since the eigenvalues of $\mathbf{A}^{-1}$, i.e., $\frac{1}{1+\alpha_i}$, decrease as $\boldsymbol{\lambda}$ increases. Therefore, the ratios on the right hand side of (25), and therefore, $\mathbf{f}(\boldsymbol{\lambda})$, increase in $\boldsymbol{\lambda}$. $\square$

Theorem 3 shows that with the addition of more and more users into the system, beamforming becomes optimal for more and more channel covariance matrices. Whether the growth in the region where beamforming is optimal is bounded, or whether beamforming is unconditionally optimal for very large numbers of users in a fading environment will be addressed in Section IV.

B. Mean Feedback at the Transmitters

1) Transmit Directions: As in the case of covariance feedback, for a single-user system with partial CSI in the form of the channel mean matrix at the transmitter, the capacity is no longer achieved by an identity transmit covariance matrix. The optimization problem in this case is the same as (7), with the

difference that, in this setting, the channel covariance matrix is identity, i.e., $\boldsymbol{\Sigma} = \mathbf{I}$, and the channel mean matrix $\mathbf{H}_\mu$ is fedback to the transmitter. With the assumption that $\mathbf{H}_\mu$ is unit-rank, [5], [6] showed that the optimal transmit covariance matrix $\mathbf{Q}$ that solves (7) can be written as

$$\mathbf{Q} = \mathbf{U}_\mu \boldsymbol{\Lambda}_Q \mathbf{U}_\mu^\dagger \qquad (27)$$

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. In the multi-user setting, let the singular value decomposition of the channel mean matrix of user k be

$$\mathbf{H}_{\mu_k} = \mathbf{U}_{\mu_k} \boldsymbol{\Lambda}_{\mu_k} \mathbf{V}_{\mu_k}^\dagger \qquad (28)$$

Since $\mathbf{H}_{\mu_k}$ is a unit-rank matrix as in (4), 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. Unlike $\mathbf{U}_{\mu_k}$, $\mathbf{V}_{\mu_k}$ is different for different users. The diagonal matrix $\boldsymbol{\Lambda}_{\mu_k}$ has only one non-zero element, which is $|\mathbf{a}_R||\mathbf{a}_{T_k}|$.

The following theorem identifies the optimum transmit directions for all users. The single-user version of this theorem was proved in [21], [22].

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

The proof of this theorem can be found in [25]. It follows [21] closely, and is omitted here.

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

$$C_{sum}(\{\mathbf{H}_{\mu_k}\}_{k=1}^{K}) = \max_{\substack{\text{tr}(\mathbf{Q}_k)\leq P_k, \mathbf{Q}_k \succeq \mathbf{0} \\ k=1,\ldots,K}} E\left[\log\left|\mathbf{I}_{n_R} + \sum_{k=1}^{K}(\mathbf{H}_{\mu_k} + \mathbf{Z}_k)\right.\right.$$
$$\left.\left.\mathbf{Q}_k(\mathbf{H}_{\mu_k} + \mathbf{Z}_k)^\dagger\right|\right] \qquad (29)$$

Using Theorem 4, we can write the optimization problem in (29) as

$$C_{sum} = \max_{\substack{\text{tr}(\boldsymbol{\Lambda}_k)\leq P_k, \boldsymbol{\Lambda}_k \succeq \mathbf{0} \\ k=1,\ldots,K}} E\left[\log\left|\mathbf{I}_{n_R} + \sum_{k=1}^{K}\hat{\mathbf{Z}}_k\boldsymbol{\Lambda}_k\hat{\mathbf{Z}}_k^\dagger\right|\right] \qquad (30)$$

$$= \max_{\substack{\sum_{i=1}^{n_T}\lambda_{ki}^Q \leq P_k, \lambda_{ki}^Q \geq 0 \\ k=1,\ldots,K}} E\left[\log\left|\mathbf{I}_{n_R} + \sum_{k=1}^{K}\sum_{i=1}^{n_T}\lambda_{ki}^Q\hat{\mathbf{z}}_{ki}\hat{\mathbf{z}}_{ki}^\dagger\right|\right] \qquad (31)$$

where $\hat{\mathbf{Z}}_k = \boldsymbol{\Lambda}_{\mu_k} + \mathbf{Z}_k$. Note that while the first column of this matrix is a non-zero mean Gaussian vector, all of the remaining columns are zero-mean Gaussian vectors.

Similar to the covariance feedback case, in a MIMO system, a transmit strategy is a combination of a transmit direction strategy and a transmit power allocation strategy. A result of Theorem 4 is that the optimal multi-user transmit direction strategies are decoupled into single-user transmit direction strategies. However, in general, this is not true for the optimal transmit power allocation strategies. On the other hand, we know that when beamforming is optimal, the optimal transmit power allocation strategy for each user is to allocate all of its power to its strongest eigen-direction. Therefore, for the range of parameters where beamforming is optimal, both the optimal transmit direction and the optimal transmit power allocation strategies are decoupled among users.

2) Conditions for the Optimality of Beamforming: In this section, we determine 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 [6] identified these conditions for a single-user system. For a single-user system, let λ^μ denote the non-zero eigenvalue of the channel mean matrix $\mathbf{H}_\mu$. Then, the necessary and sufficient condition for the optimality of beamforming is [6]:

$$P < \frac{1 - E\left[\frac{1}{1+P\hat{\mathbf{z}}^\dagger \hat{\mathbf{z}}}\right]}{n_R - 1 + E\left[\frac{1}{1+P\hat{\mathbf{z}}^\dagger \hat{\mathbf{z}}}\right]} \qquad (32)$$

where $\hat{\mathbf{z}}$ is an $n_R \times 1$ dimensional Gaussian random vector with identity-covariance. The first entry of $\hat{\mathbf{z}}$ has a mean of λ_μ, while all other entries have zero-mean.

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 reduce them to (32). 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. We have the following theorem.

Theorem 5: 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 (31) have unit-rank (i.e., beamforming is optimal for all users) if and only if

$$P_k < \frac{1 - E\left[\frac{1}{1+P_k \hat{\mathbf{z}}_{k1}^\dagger \mathbf{B}_k^{-1} \hat{\mathbf{z}}_{k1}}\right]}{n_R - K + \sum_{l=1}^K E\left[\frac{1}{1+P_l \hat{\mathbf{z}}_{l1}^\dagger \mathbf{B}_l^{-1} \hat{\mathbf{z}}_{l1}}\right]}, \quad k = 1, \ldots, K \tag{33}$$

where $\mathbf{B} = \mathbf{I}_{n_R} + \sum_{l=1}^K P_l \hat{\mathbf{z}}_{l1} \hat{\mathbf{z}}_{l1}^\dagger$, $\mathbf{B}_k = \mathbf{B} - P_k \hat{\mathbf{z}}_{k1} \hat{\mathbf{z}}_{k1}^\dagger$, and $\hat{\mathbf{z}}_{k1} = \lambda_k^\mu \mathbf{e}_1 + \mathbf{z}_{k1}$ is the first column of the matrix $\hat{\mathbf{Z}}_k$.

Proof (Theorem 5): The Lagrangian for the optimization problem in (31), with ν_k as the Lagrange multiplier of user k corresponding to its power constraint, is

$$E\left[\log\left|\mathbf{I}_{n_R} + \sum_{k=1}^K \sum_{i=1}^{n_T} \lambda_{ki}^Q \hat{\mathbf{z}}_{ki} \hat{\mathbf{z}}_{ki}^\dagger\right|\right] - \sum_{k=1}^K \nu_k \left(\sum_{i=1}^{n_T} \lambda_{ki}^Q - P_k\right) \tag{34}$$

This Lagrangian for the mean feedback case is similar to the Lagrangian for the covariance feedback case in (15) with the difference that there are no second largest eigenvalues of

the channel mean matrices. The following KKT conditions for user k can be derived using (16),

$$E\left[\hat{\mathbf{z}}_{ki}^\dagger \left(\mathbf{I} + \sum_{k=1}^K \sum_{i=1}^{n_T} \lambda_{ki}^Q \hat{\mathbf{z}}_{ki} \hat{\mathbf{z}}_{ki}^\dagger\right)^{-1} \hat{\mathbf{z}}_{ki}\right] \leq \nu_k \tag{35}$$

for $i = 1, \ldots, n_T$. Similar to the covariance feedback case, in order for beamforming to be optimal, we should have $\lambda_{k1}^Q = P_k$, and all other eigenvalues of the transmit covariance matrices to be zero. We have the following for user k,

$$E_{k1} = E\left[\hat{\mathbf{z}}_{k1}^\dagger \mathbf{B}^{-1} \hat{\mathbf{z}}_{k1}\right] = \nu_k \tag{36}$$

$$E_{ki} = E\left[\hat{\mathbf{z}}_{ki}^\dagger \mathbf{B}^{-1} \hat{\mathbf{z}}_{ki}\right] < \nu_k, \quad \forall i \neq 1 \tag{37}$$

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

$$\frac{E_{k1}}{E_{ki}} > 1, \quad \forall i \neq 1, \quad k = 1, \ldots, K \tag{38}$$

Finally, using Lemma 1, we have (33). □

Note that inserting $K = 1$ in (33), we obtain the condition in (32), which is derived in [6]. In our case, the condition in (33) depends only on the sole non-zero channel mean eigenvalues of all users. Therefore, if the powers and the eigenvalues of the feedback mean matrices of all users are such that they satisfy the inequalities in (33), then beamforming is optimal for all users.

Contrary to the covariance feedback case, in the mean feedback case, the region where beamforming is optimal does not necessarily grow with the addition of new users into the system. The reason that the proof of Theorem 3 does not follow in the mean feedback case is the following. Note that, in the covariance feedback case, the off-diagonal entries of the matrix in the numerator of (25) were zero. However, in the mean feedback case, the off-diagonal entries of the corresponding matrix are not zero. Therefore, proving that the eigenvalues of that matrix increase, does not prove that the diagonal entries of the same matrix increase as well. However, for relatively large numbers of users, we see through simulations that it is harder to violate the beamforming condition. We discuss this issue in more detail in Section V.

We have proved for the covariance feedback case and observed through simulations for the mean feedback case with relatively large numbers of users that the region where beamforming is optimal for all users grows, as new users are added to the system. These results and the asymptotic results of [15] with deterministic channel assumption motivate us to investigate whether the growth of the region where beamforming is optimal is bounded, or whether beamforming is unconditionally optimal for very large numbers of users in a fading environment. We address this issue in the next section.

IV. ASYMPTOTIC ANALYSIS

It is not immediate from the previous section that the region where beamforming is optimal covers the entire channel parameter space for all users when the number of users grows to infinity. In this section, we show that for very large numbers of users, even with the assumption that the transmitters have no knowledge of the channel, beamforming achieves a sum rate

which approaches the sum capacity. For asymptotic analysis, we need the following lemma.

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

$$E\left[\log\left|\mathbf{I}_M + \sum_{i=1}^N \alpha_i \mathbf{x}_i \mathbf{x}_i^\dagger\right|\right] \doteq M \log\left(1 + \sum_{i=1}^N \alpha_i\right) \quad (39)$$

where the symbol $\doteq$ denotes "equal for asymptotically large N".

This is a version of the Strong Law of Large Numbers (SLLN), which states that the sum of independent, non-identically distributed random variables, converges to the sum of the means of the random variables. In particular, this version of SLLN is applied to independent random vectors $\sqrt{\alpha_i} \mathbf{x}_i$ in (39) which are non-identically distributed. A formal proof of Lemma 2 is given in Appendix A.

Lemma 2 will be used to state a form of channel "hardening" in the next three sub-sections. We will use Lemma 2 to say that, when the number of users grows to infinity, a form of channel hardening will occur, i.e., roughly speaking $\sum_{k=1}^K \sum_{i=1}^{n_T} \mathbf{h}_{ki} \mathbf{h}_{ki}^\dagger$ in (6), $\sum_{k=1}^K \sum_{i=1}^{n_T} \lambda_{ki}^Q \lambda_{ki}^\Sigma \mathbf{z}_{ki} \mathbf{z}_{ki}^\dagger$ in (10), and $\sum_{k=1}^K \sum_{i=1}^{n_T} \lambda_{ki}^Q \hat{\mathbf{z}}_{ki} \hat{\mathbf{z}}_{ki}^\dagger$ in (31) will converge to deterministic quantities almost surely, and that those deterministic quantities can be approached if simple beamforming is used. When beamforming is used, the sum $\sum_{i=1}^{n_T}$ drops in all three sums, however the sum over k, i.e., $\sum_{k=1}^K$ suffices to create the effect of SLLN.

The concept of channel hardening has been observed in [16] also, where instead of a SLLN approach, a Central Limit Theorem (CLT) approach is used to conclude that the mutual information converges to a Gaussian random variable whose variance vanishes. In [16], the number(s) of transmit and receive antennas grow(s) large for a single-user system, while here, the number of transmit and receive antennas are fixed, but the number of users goes to infinity in a MAC. Nevertheless, we observe similar mathematical phenomena as in [16].

A. No CSI at the Transmitters

When there is no CSI at the transmitters, the optimal transmit strategy is to use an identity transmit covariance matrix for all users [3]. In this section, we show that when there is no CSI at the transmitters, for an asymptotically large system, an arbitrary antenna selection scheme is sufficient to achieve the sum capacity. This is stated in the following theorem.

Theorem 6: 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 transmit covariance matrices approaches the sum capacity. In particular, this unit-rank transmission scheme takes the form of a simple antenna selection.

Proof (Theorem 6): The sum capacity in this case is given in (5) with $\mathbf{Q}_k = \frac{P_k}{n_T}\mathbf{I}$, for all k. We define C_{sum}^{aas} as the achievable sum rate by performing arbitrary antenna selection (aas) at all transmitters:

$$C_{sum}^{aas} = E\left[\log\left|\mathbf{I}_{n_R} + \sum_{k=1}^K P_k \mathbf{h}_{ka_k} \mathbf{h}_{ka_k}^\dagger\right|\right] \quad (40)$$

where $\mathbf{h}_{ka_k}$ is the a_k^{th} column of the channel matrix of user k, and a_k is the antenna chosen by user k, $1 \le a_k \le n_T$. The choice of the columns does not affect our result. All users may select their first antenna, i.e., $a_k = 1$, for all k, or they may select an antenna arbitrarily. Since Strong Law of Large Numbers (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 noting that P_k are a series of bounded numbers, we apply Lemma 2 to (5) by inserting $\mathbf{Q}_k = \frac{P_k}{n_T}\mathbf{I}$, for all k, and to (40). We have

$$C_{sum} \doteq C_{sum}^{aas} \doteq n_R \log\left(1 + \sum_{k=1}^K P_k\right) \quad (41)$$

Therefore, we see that the sum rates achievable by the optimal power allocation and the arbitrary antenna selection scheme converge to the same quantity asymptotically. □

We note that this result does not contradict with the result of [3] which is stated in Section III. For a multi-user system, full-rank transmit covariance matrices are optimum in the sense of maximizing the sum rate [3]. Theorem 6 states that arbitrary antenna selection scheme is also sufficient to achieve the optimum when the number of users grows to infinity. In other words, the performance of the arbitrary antenna selection scheme converges to the optimum when the number of users goes to infinity.

B. Covariance Feedback at the Transmitters

When the transmitters have the partial CSI in the form of covariance feedback, Theorem 1 shows that for any number of users, the transmit directions of a user are the eigenvectors of its own channel covariance feedback matrix. For sufficiently large numbers of users, the asymptotic optimality of beamforming in achieving the sum capacity is proved in the following theorem.

Theorem 7: 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 transmit covariance matrices (i.e., beamforming) approaches the 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 7): Note that λ_{ki}^Q is bounded for all (k, i), since power constraints for all users are finite, and λ_{ki}^Σ is bounded for all (k, i), since the covariance matrix, $\mathbf{\Sigma}_k$, of the channel has finite trace. Now, we can apply Lemma 2 to (10) with $\alpha_{ki} = \lambda_{ki}^Q \lambda_{ki}^\Sigma$. We have

$$C_{sum} \doteq \max_{\substack{\sum_{i=1}^{n_T} \lambda_{ki}^Q \le P_k \\ k=1\ldots K}} n_R \log\left(1 + \sum_{k=1}^K \sum_{i=1}^{n_T} \lambda_{ki}^Q \lambda_{ki}^\Sigma\right) \quad (42)$$

In order to solve the above optimization problem, we form the Lagrangian with μ_k's as the Lagrange multipliers,

$$n_R \log \left(1 + \sum_{k=1}^{K} \sum_{i=1}^{n_T} \lambda_{ki}^Q \lambda_{ki}^\Sigma \right) - \sum_{k=1}^{K} \mu_k \left(\sum_{i=1}^{n_T} \lambda_{ki}^Q - P_k \right)$$
(43)

The KKT optimality conditions for $i = 1, \ldots, n_T$, and $k = 1, \ldots, K$ are,

$$\frac{n_R \lambda_{ki}^\Sigma}{1 + \sum_{k=1}^{K} \sum_{i=1}^{n_T} \lambda_{ki}^Q \lambda_{ki}^\Sigma} \leq \mu_k$$
(44)

where (44) 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 $\lambda_{kn_T}^\Sigma < \cdots < \lambda_{k1}^\Sigma$ 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 λ_{kj}^Q, $j = 1, \cdots, n_T$ can be non-zero. From the objective function in (42), we observe that the non-zero λ_{kj}^Q must correspond to the largest eigenvalue of the channel covariance feedback matrix. Hence, the only non-zero power component in Λ_{Q_k} is the first diagonal element. Finally, from the trace constraint, we have $\lambda_{k1}^Q = P_k$, for all k. The asymptotic sum capacity becomes

$$C_{sum} \doteq n_R \log \left(1 + \sum_{k=1}^{K} P_k \lambda_{k1}^\Sigma \right)$$
(45)

□

C. Mean Feedback at the Transmitters

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 own channel mean feedback matrix. For sufficiently large numbers of users, the asymptotic optimality of beamforming in achieving the sum capacity is proved in the following theorem.

Theorem 8: In a system where there is partial CSI at the transmitters in the form of mean feedback, if the number of users grows to infinity, then the sum rate achieved by unit-rank transmit covariance matrices (i.e., beamforming) approaches the 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 8): First, note that λ_{ki}^μ for all (k, i) is bounded, since the channel has finite mean information. Applying Lemma 2 to (31), while noting that λ_{ki}^μ is non-zero for only $i = 1$, and therefore using different α_{ki} (all of which are bounded) for diagonal and off-diagonal entries, we get

$$C_{sum} \doteq \max_{\substack{\sum_{i=1}^{n_T} \lambda_{ki}^Q \leq P_k \\ k=1\ldots K}} (n_R - 1) \log \left(1 + \sum_{k=1}^{K} P_k \right)$$

$$+ \log \left(1 + \sum_{k=1}^{K} \left((\lambda_{k1}^\mu)^2 \lambda_{k1}^Q + P_k \right) \right)$$
(46)

Since only one eigenvalue from the transmit covariance matrix of each user appears in (46), the optimum choice is to allocate all of the power of each user to the eigenvector of its own channel mean matrix corresponding to the only non-zero eigenvalue, i.e., $\lambda_{k1}^Q = P_k$, for all k. The resulting sum capacity,

$$C_{sum} \doteq (n_R - 1) \log \left(1 + \sum_{k=1}^{K} P_k \right)$$

$$+ \log \left(1 + \sum_{k=1}^{K} \left((\lambda_{k1}^\mu)^2 P_k + P_k \right) \right)$$
(47)

□

V. NUMERICAL RESULTS

The region where beamforming is optimal is multi-dimensional. In order to illustrate the effects of having more than two users, we plot two dimensional slices from the region where beamforming is optimal for all users. We first consider the covariance feedback case, and plot these slices for $K = 1, 2, 3, 5, 10$ users in Figure 1. These slices give the maximum possible λ_{12}^Σ for a range of λ_{11}^Σ. The largest eigenvalues of the remaining users are kept constant. The number of transmit and receive antennas is $n_T = n_R = 2$. We see that the region where beamforming is optimal gets larger with increasing number of users. Note that these curves have to lie below the $\lambda_{12}^\Sigma = \lambda_{11}^\Sigma$ line, because λ_{11}^Σ is the largest eigenvalue. The top-most line in Figure 1 is the $\lambda_{12}^\Sigma = \lambda_{11}^\Sigma$ line. 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 into the system, a larger range of $(\lambda_{11}^\Sigma, \lambda_{12}^\Sigma)$ pairs becomes optimal.

For the mean feedback model, we will demonstrate two different cases. In the first case, the region where beamforming is optimal gets larger by addition of new users into the system. In Figure 2, we plot one dimensional slices from the region corresponding to $K = 1, 2, 3, 5, 10$. These lines give λ_1^μ values for beamforming to be optimal for a given power constraint, $P_k = 1$ for all k.

The largest eigenvalues of all other users, which are kept constant, are comparable in value to each other. The number of transmit and receive antennas is $n_T = n_R = 2$. The curves in Figure 2 correspond to the left hand side of (38) for $k = 1$. Beamforming is optimal for the range of λ_1^μ, where the curves stay above the horizontal line at $P_k = 1$. For example, in the single user case, beamforming is optimal for λ_1^μ values to the left of point A and to the right of point B, while beamforming is not optimal for all λ_1^μ values between points A and B. In the second case, the region where beamforming is optimal first gets smaller by the addition of new users into the system, however it then starts to get larger as the number of users is increased further. In Figure 3, we plot one dimensional slices from the region corresponding to $K = 1, 2, 5, 10, 20, 30$. These curves give λ_1^μ values for beamforming to be optimal for a given power constraint, $P_k = 1$ for all k. The largest eigenvalues of all other users are kept constant, and each new user that is added to the system has a larger mean channel value than those of the users that are already in the system.
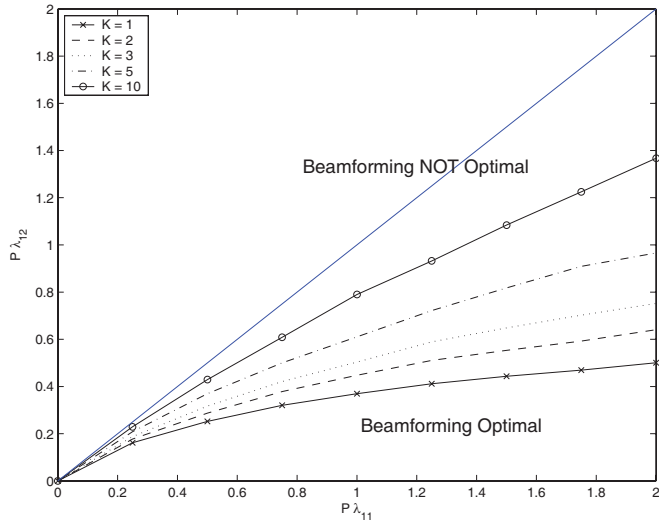
Fig. 1. The region where beamforming is optimal for various numbers of users in the covariance feedback model.
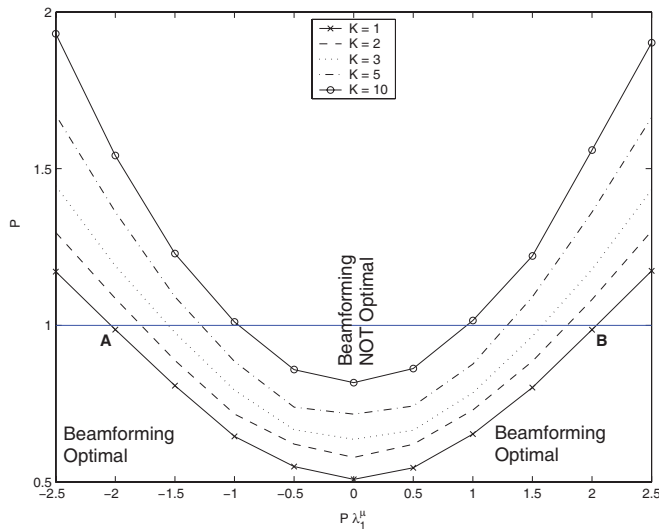

Fig. 2. The region where beamforming is optimal for various numbers of users in the mean feedback model. This is an example where the region gets larger as more users are added into the system.
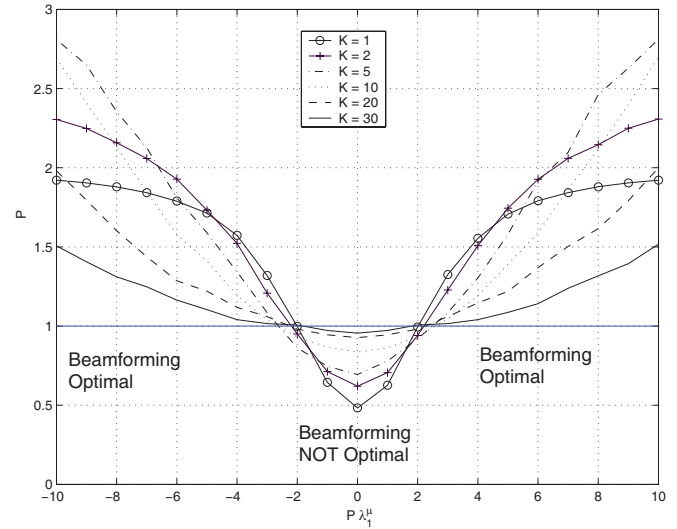

Fig. 3. The region where beamforming is optimal for various numbers of users in the mean feedback model. This is an example where the region does not get larger as more users are added into the system.
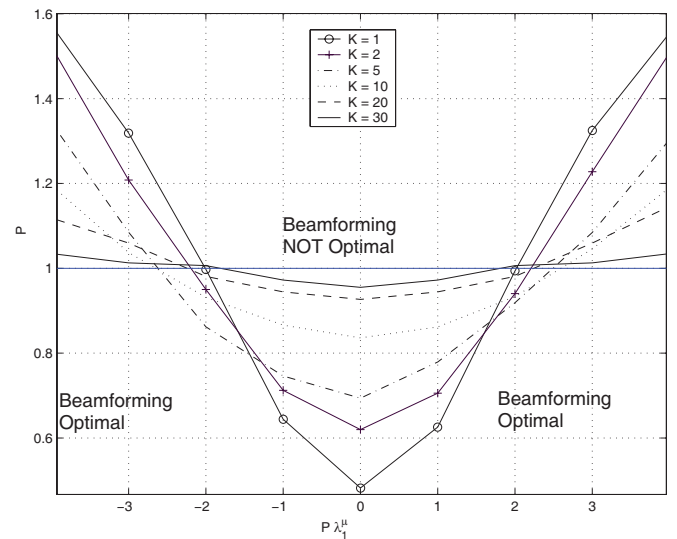

Fig. 4. This is the same as Figure 3 where we zoom into the center of the figure to show details.

The number of transmit and receive antennas is $n_T = n_R = 2$. In Figure 4, we zoom into the center of Figure 3 in order to show the details.

In light of these two examples, we conclude that, for the mean feedback model, the region where beamforming is optimal does not necessarily get larger with increasing number of users. However, we observe that as the number of users grows, the region starts to get larger regardless of the mean channel values of the users. The situation in Figures 3 and 4 is a worst case scenario. Even in this worst case, only the region corresponding to the first user does not get larger, while the corresponding regions for all other users in the system get larger. This possibly follows from the fact that the first user has the lowest mean channel value.

In Figure 5, we illustrate the change in the region where beamforming is optimal with the number of receive antennas for the covariance feedback model, while the number of transmit antennas is kept at $n_T = 2$. We observe that the region

gets smaller as the number of receive antennas is increased. However, for a fixed number of receive antennas, the region grows with the number of users, and eventually equals the entire parameter region asymptotically as the number of users goes to infinity.

For the asymptotic analysis, in Figure 6, we show three simple examples for different numbers of receive and transmit antennas. We plot the sum rates resulting from optimal power allocation and arbitrary antenna selection schemes for the no CSI model. We observe that, for this instance, even for a small number of users, arbitrary antenna selection performs very close to the optimum power allocation scheme.

VI. CONCLUSION

We determined the optimal transmit directions and the region where beamforming is optimal for all users under covariance and mean feedback CSI models for a multi-user
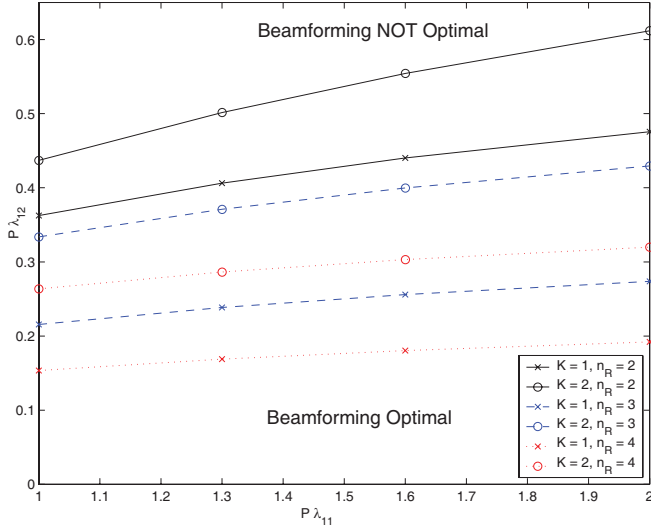
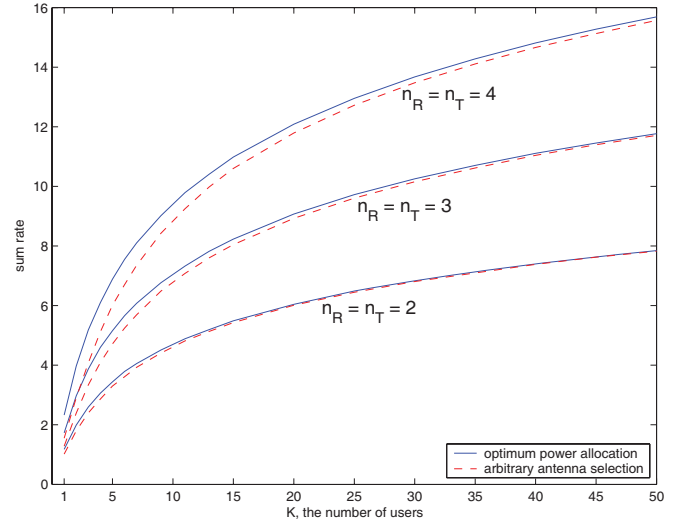Fig. 5. The region where beamforming is optimal for various numbers of receive antennas in the covariance feedback model.

Fig. 6. Sum rates for the optimal and arbitrary antenna selection schemes, as a function of the number of users in the no CSI model.

MIMO-MAC. We proved that the region where beamforming is optimal gets larger by the addition of new users into the system in the covariance feedback case. In the mean feedback case, we observed through simulations that the region where beamforming is optimal gets larger for relatively large numbers of users. We showed that in an asymptotically large system, beamforming is always optimal for all users not only for the covariance and the mean feedback cases, but also for the no CSI case as well. Combining our results with those of [15], we conclude that in a large multi-user MIMO-MAC system, beamforming is optimal under full, partial (covariance and mean), and no CSI cases.

VII. APPENDIX

A. Proofs of the Lemmas

Proof (Lemma 1): Using the matrix inversion lemma [26, page 19]

$$\mathbf{A}^{-1} = \mathbf{A}_k^{-1} - \mathbf{A}_k^{-1} \mathbf{z}_{k1} \left(\frac{1}{P_k \lambda_{k1}^{\Sigma}} + \mathbf{z}_{k1}^{\dagger} \mathbf{A}_k^{-1} \mathbf{z}_{k1} \right)^{-1} \mathbf{z}_{k1}^{\dagger} \mathbf{A}_k^{-1}$$

(A-1)

Multiplying this with $\mathbf{z}_{k1}^{\dagger}$ from left, and $\mathbf{z}_{k1}$ from right yields

$$\mathbf{z}_{k1}^{\dagger} \mathbf{A}^{-1} \mathbf{z}_{k1} = \frac{1}{P_k \lambda_{k1}^{\Sigma}} \left(1 - \frac{1}{1 + P_k \lambda_{k1}^{\Sigma} \mathbf{z}_{k1}^{\dagger} \mathbf{A}_k^{-1} \mathbf{z}_{k1}} \right)$$

(A-2)

By taking the expectation of both sides, (12) follows. In order to derive the identity in (13), note that

$$\lambda_{ki}^{\Sigma} E \left[\mathbf{z}_{ki}^{\dagger} \mathbf{A}^{-1} \mathbf{z}_{ki} \right] = \lambda_{ki}^{\Sigma} E \left[\mathrm{tr}(\mathbf{A}^{-1}) \right] \quad \text{(A-3)}$$

since $\mathbf{z}_{k2}$ is independent of $\mathbf{A}$, and has identity-covariance. Applying the matrix inversion lemma [26, page 19] to $\mathbf{A} = \mathbf{I}_{n_R} + \bar{\mathbf{Z}} \tilde{\mathbf{\Lambda}} \bar{\mathbf{Z}}^{\dagger}$, with $\tilde{\mathbf{\Lambda}} = \bar{\mathbf{P}} \bar{\mathbf{\Lambda}}$

$$\mathbf{A}^{-1} = \mathbf{I}_{n_R} - \bar{\mathbf{Z}} \tilde{\mathbf{\Lambda}}^{1/2} (\mathbf{I}_K + \tilde{\mathbf{\Lambda}}^{1/2} \bar{\mathbf{Z}}^{\dagger} \bar{\mathbf{Z}} \tilde{\mathbf{\Lambda}}^{1/2})^{-1} \tilde{\mathbf{\Lambda}}^{1/2} \bar{\mathbf{Z}}^{\dagger}$$

(A-4)

Calculating the traces of both sides, we obtain

$$\mathrm{tr}(\mathbf{A}^{-1}) = n_R - \mathrm{tr}\left((\mathbf{I}_K + \tilde{\mathbf{\Lambda}}^{1/2} \bar{\mathbf{Z}}^{\dagger} \bar{\mathbf{Z}} \tilde{\mathbf{\Lambda}}^{1/2})^{-1} \right.$$
$$\left. (\mathbf{I}_K + \tilde{\mathbf{\Lambda}}^{1/2} \bar{\mathbf{Z}}^{\dagger} \bar{\mathbf{Z}} \tilde{\mathbf{\Lambda}}^{1/2} - \mathbf{I}_K) \right)$$

(A-5)

$$= n_R - K + \mathrm{tr}\left((\mathbf{I}_K + \tilde{\mathbf{\Lambda}}^{1/2} \bar{\mathbf{Z}}^{\dagger} \bar{\mathbf{Z}} \tilde{\mathbf{\Lambda}}^{1/2})^{-1} \right) \quad \text{(A-6)}$$

$$= n_R - K + \frac{\sum_{k=1}^{K} |\mathbf{A}_k|}{|\mathbf{A}|}$$

(A-7)

where in the last equation, we used the definition of an inverse of a matrix [26, page 21] and $|\mathbf{I} + \mathbf{AB}| = |\mathbf{I} + \mathbf{BA}|$. Noting that $\frac{|\mathbf{A}_k|}{|\mathbf{A}|} = \frac{1}{1 + P_k \lambda_{k1}^{\Sigma} \mathbf{z}_{k1}^{\dagger} \mathbf{A}_k^{-1} \mathbf{z}_{k1}}$, and taking the expectations of both sides, we have (13). $\square$

Proof (Lemma 2): We will apply a version of the SLLN from [27, page 27, Theorem D]. In this version of the SLLN, the sum of a sequence of independent random variables with different means and variances converges to the sum of the sequence of means of the random variables, subject to the condition that $\sum_i^N \frac{\sigma_i^2}{i^2}$ converges, where σ_i^2 are the variances of the random variables. We will apply this theorem to every element of the matrix at hand, that is, to $\sum_{i=1}^N \alpha_i x_{ik} x_{ij}^*$, for all (k, j). In order to invoke the theorem, we let $\alpha_i x_{ik} x_{ij}^*$, for all i, be the sequence of independent but not identically distributed random variables. Note that the expectations of the diagonal elements are α_i, and the expectations of the off-diagonal elements are zero. Since α_i are assumed to be bounded and $\mathbf{x}_i$ have zero-mean and identity covariance, $\sum_i^N \frac{\alpha_i^2 E[x_{ik}^2 x_{ij}^2]}{i^2}$ converges, for all (k, j). As a result, we have $\sum_{i=1}^N \alpha_i \mathbf{x}_i \mathbf{x}_i^{\dagger} \rightarrow \sum_{i=1}^N \alpha_i \mathbf{I}_M$.

Due to [27, page 24], if a random variable converges to a deterministic number, a, then a function, f, of that random variable converges to $f(a)$. Therefore,

$$\log \left| \mathbf{I}_M + \sum_{i=1}^N \alpha_i \mathbf{x}_i \mathbf{x}_i^{\dagger} \right| \rightarrow M \log \left(1 + \sum_{i=1}^N \alpha_i \right), \text{ a.s.} \quad \text{(A-8)}$$

If a random variable converges to a number almost surely, then the expectation of that random variable will be equal to

the same number (for large N). That is,

$$E\left[\log\left|\mathbf{I}_M + \sum_{i=1}^N \alpha_i \mathbf{x}_i \mathbf{x}_i^\dagger\right|\right] \doteq M \log\left(1 + \sum_{i=1}^N \alpha_i\right). \quad \text{(A-9)}$$

$\square$

B. General Receive Antenna Correlation Matrix

In the model that we considered in the main part of this paper, the receiver side correlation matrix is the identity matrix as a result of the assumption that the receiver (e.g., a base station) is not physically limited and one can place the antenna elements sufficiently far away from each other. In a different physical model with receiver side correlation present in the system, similar results can be found. For the single-user scenario, it is already known that the transmit directions are still the eigenvectors of the transmitter side channel correlation matrix, even when there is receiver side channel correlation in the system [8]. Beamforming optimality condition for this case is also found previously [8]. For the multi-user scenario, our approach generalizes to the case where there is receiver side channel correlation in the system, when the receiver side channel correlation matrices of all users are the same. In this case, the channel is modeled as,

$$\mathbf{H}_k = \boldsymbol{\Phi}^{1/2} \mathbf{Z}_k \boldsymbol{\Sigma}_k^{1/2} \quad \text{(B-1)}$$

where the receive antenna correlation matrix, $\boldsymbol{\Phi}$, is the correlation between the signals received at the n_R receive antennas of the receiver. This correlation matrix does not depend on the specific transmit antenna from which the signal is transmitted [19]. In a MAC, since we have a single receiver and the correlation matrix does not depend on the transmitters, we have the same $\boldsymbol{\Phi}$ for all users. The transmit antenna correlation matrix, $\boldsymbol{\Sigma}_k$, is the correlation between the signals transmitted from the n_T transmit antennas of user k.

The result of Theorem 1, which states that the optimal transmit directions are the eigenvectors of the transmit antenna correlation matrix, remains exactly the same. The result of Theorem 2, which states the conditions under which beamforming is optimal for all users, changes slightly. The refined conditions involve the eigenvalues of the receive correlation matrix. The region formed by these refined conditions still grows with the addition of a new user, and therefore, Theorem 3 also remains exactly the same. Finally, the asymptotic sum capacity expression in Theorem 7 changes slightly and involves the eigenvalues of the receive correlation matrix. Below, we outline the reasons that Theorems 1 and 3 remain the same, and state the refined conditions for the optimality of beamforming and the new asymptotic sum capacity expression.

Proof (Theorem 1): Let $\boldsymbol{\Phi} = \mathbf{U}_\Phi \boldsymbol{\Lambda}_\Phi^{1/2} \mathbf{U}_\Phi^\dagger$ be the spectral decomposition of the receive antenna correlation matrix. Inserting this into (B-1), we have

$$\mathbf{H}_k = \mathbf{U}_\Phi \boldsymbol{\Lambda}_\Phi^{1/2} \mathbf{U}_\Phi^\dagger \mathbf{Z}_k \mathbf{U}_{\Sigma_k} \boldsymbol{\Lambda}_{\Sigma_k}^{1/2} \mathbf{U}_{\Sigma_k}^\dagger \quad \text{(B-2)}$$

Then, inserting (B-2) into (5) and following similar lines to [8], we obtain

$$C_{sum} = \max_{\substack{\text{tr}(\mathbf{Q}_k) \le P_k, \mathbf{Q}_k \succeq \mathbf{0} \\ k=1,\dots,K}} E\left[\log\left|\mathbf{I}_{n_R} + \boldsymbol{\Lambda}_\Phi \sum_{k=1}^K \mathbf{Z}_k \boldsymbol{\Lambda}_{Q_k} \boldsymbol{\Lambda}_{\Sigma_k} \mathbf{Z}_k^\dagger\right|\right]$$
$$\text{(B-3)}$$

The only difference between the proofs of Theorem 1 in uncorrelated and correlated receiver structures is that, here, we have the matrix $\boldsymbol{\Lambda}_\Phi$ in front of the summation term inside the logarithm compared to (9), which does not affect the derivations. Therefore, we observe that, even when the receive antenna correlation matrix is not equal to identity, the transmit directions of all users continue to depend only on their own transmit antenna correlation matrices. However, the resulting sum capacity is different, and the optimal power allocation will depend on the eigenvalues of the receive antenna correlation matrix. $\square$

The sum capacity expression in this case can be written, similar to (10), as

$$C_{sum} = \max_{\substack{\sum_{i=1}^{n_T} \lambda_{ki}^Q \le P_k, \lambda_{ki}^Q \ge 0 \\ k=1,\dots,K}} E\left[\log\left|\mathbf{I}_{n_R} + \sum_{k=1}^K \sum_{i=1}^{n_T} \lambda_{ki}^Q \lambda_{ki}^\Sigma \tilde{\mathbf{z}}_{ki} \tilde{\mathbf{z}}_{ki}^\dagger\right|\right]$$
$$\text{(B-4)}$$

where $\{\tilde{\mathbf{z}}_{ki} = \boldsymbol{\Lambda}_\Phi^{1/2} \mathbf{z}_{ki}, k = 1,\dots,K, i = 1,\dots,n_T\}$ is a set of i.i.d. Gaussian random vectors with zero-mean and covariance matrix $\boldsymbol{\Lambda}_\Phi$.

Proof (Theorem 2): After taking the derivative of the Lagrangian for the optimization problem in (B-4), the conditions for the optimality of beamforming for $k = 1,\dots,K$ become

$$\frac{\tilde{E}_{k1}}{\tilde{E}_{ki}} = \frac{\lambda_{k1}^\Sigma E\left[\tilde{\mathbf{z}}_{k1}^\dagger \tilde{\mathbf{A}}^{-1} \tilde{\mathbf{z}}_{k1}\right]}{\lambda_{ki}^\Sigma E\left[\tilde{\mathbf{z}}_{ki}^\dagger \tilde{\mathbf{A}}^{-1} \tilde{\mathbf{z}}_{ki}\right]} > 1, \quad \forall i \ne 1 \quad \text{(B-5)}$$

where $\tilde{\mathbf{A}} = \mathbf{I}_{n_R} + \sum_{l=1}^K P_l \lambda_{l1}^\Sigma \tilde{\mathbf{z}}_{l1} \tilde{\mathbf{z}}_{l1}^\dagger$. The identities in Lemma 1 change slightly for the general $\boldsymbol{\Phi}$ case. The details of the derivations only require matrix algebra and are omitted here due to space limitations. Inserting the new identities from Lemma 1 into (B-5), we have $k = 1,\dots,K$

$$P\lambda_{k2}^\Sigma < \frac{1 - E\left[\frac{1}{1 + P_k \lambda_{k1}^\Sigma \tilde{\mathbf{z}}_{k1}^T \tilde{\mathbf{A}}_k^{-1} \tilde{\mathbf{z}}_{k1}}\right]}{\sum_{i=1}^{n_R} \lambda_i^\Phi - \sum_{l=1}^K E\left[\frac{P_l \lambda_{l1}^\Sigma \tilde{\mathbf{z}}_{l1}^\dagger \boldsymbol{\Lambda}_\Phi^{1/2} \tilde{\mathbf{A}}_l^{-1} \boldsymbol{\Lambda}_\Phi^{1/2} \tilde{\mathbf{z}}_{l1}}{1 + P_l \lambda_{l1}^\Sigma \tilde{\mathbf{z}}_{l1}^T \tilde{\mathbf{A}}_l^{-1} \tilde{\mathbf{z}}_{l1}}\right]}$$
$$\text{(B-6)}$$

where $\tilde{\mathbf{A}}_k = \tilde{\mathbf{A}} - P_k \lambda_{k1}^\Sigma \tilde{\mathbf{z}}_{k1} \tilde{\mathbf{z}}_{k1}^\dagger$, for all k. $\square$

Inserting $\boldsymbol{\Phi} = \mathbf{I}$, and adding and subtracting "1" from the numerator of the expectation term in the denominator of (B-6), we get (14). And, inserting $K = 1$, and $\tilde{\mathbf{A}}_k = \mathbf{I}$, for all k, in (B-6), we get the single-user condition derived in [8].

Proof (Theorem 3): The proof exactly follows the original proof. We only use the fact that $\{\tilde{\mathbf{z}}_{ki}\}$ are independent and zero-mean random vectors. $\square$

Proof (Theorem 7): Applying Lemma 2 to (B-4), the objective function in (42) changes to a summation of $\log$ functions, instead of $n_R \log(\cdot)$. Using the Lagrangian method, (45)

becomes

$$C_{sum} \doteq \sum_{i=1}^{n_R} \log\left(1 + \lambda_i^\Phi \sum_{k=1}^{K} P_k \lambda_{k1}^\Sigma\right) \qquad (B-7)$$

□

REFERENCES

[1] A. Soysal and S. Ulukus, "Transmit directions and optimality of beamforming in MIMO-MAC with partial CSI at the transmitters," in *Proc. Conference on Information Sciences and Systems*, Baltimore, MD, Mar. 2005.

[2] A. Soysal and S. Ulukus, "Asymptotic optimality of beamforming in multi-user MIMO-MAC with no or partial CSI at the transmitters," in *Proc. Vehicular Technology Conference*, Stockholm, Sweden, May 2005.

[3] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, no. 6, pp. 585–596, Nov. 1999.

[4] G. J. Foschini and M. J Gans, "On limits of wireless communication in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, pp. 311–335, 1998.

[5] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2632–2639, Sept. 2001.

[6] S. A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1165–1175, July 2004.

[7] H. Boche and E. Jorswieck, "On the optimality range of beamforming for MIMO systems with covariance feedback," *IEICE Trans. Commun.*, vol. E85-A, no. 11, pp. 2521–2528, Nov. 2002.

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

[9] E. Jorswieck and H. Boche, "Optimal transmission with imperfect channel state information at the transmit antenna array," *Wireless Personal Commun.*, vol. 27, no. 1, pp. 33–56, 2003.

[10] A. M. Tulino, A. Lozano, and S. Verdú, "Power allocation in multi-antenna communication with statistical channel information at the transmitter," in *Proc. IEEE Int. Symp. on PIMRC*, Barcelona, Spain, Sept. 2004.

[11] A. M. Tulino, A. Lozano, and S. Verdú, "Capacity-achieving input covariance for single-user multi-antenna channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 662–671, Mar. 2006.

[12] A. Soysal and S. Ulukus, "Optimum power allocation in fading MIMO multiple access channels with partial CSI at the transmitters," in *Proc. Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 2006.

[13] A. Soysal and S. Ulukus, "Optimum power allocation for single-user MIMO and multi-user MIMO-MAC with partial CSI," *IEEE J. Select. Areas Commun.*, vol. 25, no. 7, pp. 1402–1412, Sept. 2007.

[14] S. A. Jafar, S. Vishwanath, and A. Goldsmith, "Vector MAC capacity region with covariance feedback," in *Proc. ISIT*, p. 54, June 2001.

[15] W. Rhee, W. Yu, and J. M. Cioffi, "The optimality of beamforming in uplink multiuser wireless systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 86–96, Jan. 2004.

[16] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 1893–1909, Sept. 2004.

[17] S. V. Hanly and D. N. C. Tse, "Resource pooling and effective bandwidths in CDMA networks with multiuser receivers and spatial diversity," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1328–1351, May 2001.

[18] A. Mantravadi, V. V. Veeravalli, and H. Viswanathan, "Spectral efficiency of MIMO multiaccess systems with single-user decoding," *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 382–394, Apr. 2003.

[19] C. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 637–650, Mar. 2002.

[20] W. C. Jakes, ed., *Microwave Mobile Communications*. Wiley, 1974.

[21] D. Hösli and A. Lapidoth, "The capacity of a MIMO Ricean channel is monotonic in the singular values of the mean," in *Proc. 5th International ITG Conference on Source and Channel Coding*, Jan. 2004.

[22] D. Hoesli, Y. Kim, and A. Lapidoth, "Monotonicity results for coherent MIMO Ricean channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4334–4339, Dec. 2005.

[23] A. Lozano, F. R. Farrokhi, and R. A Valenzuela, "Asymptotically optimal open-loop space-time architecture adaptive to scattering conditions," in *Proc. Vehicular Technology Conference*, Spring 2001.

[24] W. Rhee and J. M. Cioffi, "On the capacity of multiuser wireless channels with multiple antennas," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2580–2595, Oct. 2003.

[25] A. Soysal, "Optimum transmit strategies for Gaussian multi-user MIMO networks with partial CSI," Ph.D. thesis, University of Maryland, 2008.

[26] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[27] R. J. Serfling, *Approximation Theorems of Mathematical Statistics*. Wiley, 1980.

Alkan Soysal (S'03, M'08) received the B.S. degree in electrical and electronics engineering from Middle East Technical University, Ankara, Turkey in 2002, and the M.S. and Ph.D. degrees in electrical and computer engineering from University of Maryland, College Park, MD in 2006 and 2008 respectively. He joined Bahçeşehir University, Istanbul, Turkey in February 2008 as an Assistant Professor in the Department of Electrical and Electronics Engineering. His research interests include wireless communication theory and information theory with particular focus on MIMO networks with various assumptions on the availability of the channel state information.

Sennur Ulukus (S'90, M'98) received the B.S. and M.S. degrees in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 1991 and 1993, respectively, and the Ph.D. degree in electrical and computer engineering from Rutgers University, NJ, in 1998. During her Ph.D. studies, she was with the Wireless Information Network Laboratory (WINLAB), Rutgers University. From 1998 to 2001, she was a Senior Technical Staff Member at AT&T Labs-Research in NJ. In 2001, she joined the University of Maryland at College Park, where she is currently an Associate Professor in the Department of Electrical and Computer Engineering, with a joint appoinment at the Institute for Systems Research (ISR). Her research interests are in wireless communication theory and networking, network information theory for wireless networks, signal processing for wireless communications, and security for multi-user wireless communications.

Sennur Ulukus is a recipient of the 2005 NSF CAREER Award, and a co-recipient of the 2003 IEEE Marconi Prize Paper Award in Wireless Communications. She serves/served as an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY since 2007, as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS between 2003-2007, as a Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS in 2006-2008, as the co-chair of the Communication Theory Symposium at the 2007 IEEE Global Telecommunications Conference, as the co-chair of the Medium Access Control (MAC) Track at the 2008 IEEE Wireless Communications and Networking Conference, and as the Secretary of the IEEE Communication Theory Technical Committee (CTTC) since 2007.