

# Optimum Power Allocation for Single-User MIMO and Multi-User MIMO-MAC with Partial CSI

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

**Abstract**—We consider both the single-user and the multi-user power allocation problems in MIMO systems, where the receiver side has the perfect channel state information (CSI), and the transmitter side has partial CSI, which is in the form of covariance feedback. In a single-user MIMO system, we consider an iterative algorithm that solves for the eigenvalues of the optimum transmit covariance matrix that maximizes the rate. The algorithm is based on enforcing the Karush-Kuhn-Tucker (KKT) optimality conditions of the optimization problem at each iteration. We prove that this algorithm converges to the unique global optimum power allocation when initiated at an arbitrary point. We, then, consider the multi-user generalization of the problem, which is to find the eigenvalues of the optimum transmit covariance matrices of all users that maximize the sum rate of the MIMO multiple access channel (MIMO-MAC). For this problem, we propose an algorithm that finds the unique optimum power allocation policies of all users. At a given iteration, the multi-user algorithm updates the power allocation of one user, given the power allocations of the rest of the users, and iterates over all users in a round-robin fashion. Finally, we make several suggestions that significantly improve the convergence rate of the proposed algorithms.

**Index Terms**—Multi-user MIMO, MIMO multiple access channel, partial CSI, covariance feedback, optimum power allocation.

## I. INTRODUCTION

IN GAUSSIAN MIMO multiple access systems, when the receiver side has the perfect CSI, the calculation of the information theoretic capacity boils down to finding the transmit covariance matrices of the users. Finding the transmit covariance matrices, in turn, involves two components: finding the optimum transmit directions and finding the optimum power allocation policies. In a single-user MIMO system, when both the receiver and the transmitter have the perfect CSI and the channel is fixed, [2] showed that the optimum transmit directions are the right singular vectors of the deterministic channel matrix, and the optimum power allocation policy is to water-fill over the singular values of the deterministic channel matrix. In a multi-user MIMO system, when both the receiver and the transmitters have the perfect CSI and the channel is fixed, [3] showed that the optimum transmit directions and the power allocation policies can be found using an iterative algorithm that updates the transmit directions and the power allocation policy of one user at a time. When the channel is changing over time due to fading, and perfect and instantaneous CSI is known both at the receiver and at

the transmitter side, these solutions extend to water-filling over both the antennas and the channel states in single-user [2], and multi-user [4] MIMO systems. However, in most of the wireless communication scenarios, especially in wireless MIMO communications, it is unrealistic to assume that the transmitter side has the perfect knowledge of the instantaneous CSI. In such scenarios, it might be more realistic to assume that only the receiver side can perfectly estimate the instantaneous CSI, while the transmitter side has only a statistical knowledge of the channel.

When the fading in the channel is assumed to be a Gaussian process, statistics of the channel reduce to the mean and covariance information of the channel. The problem in this setting as well is to find the optimum transmit covariance matrices, i.e., the optimum transmit directions and the optimum power allocation policies. However, in this case the transmit directions and the power allocations are not functions of the channel states, but they are functions of the statistics of the channel states, that are fed by the receiver back to the transmitters. The optimization criteria that we consider are the maximum rate in a single-user system, and the maximum sum rate in a multi-user system. For the covariance feedback case, it was shown in [5] for a multi-input single-output (MISO) system, and in [6], [7] for a MIMO system that the optimal transmit covariance matrix and the channel covariance matrix have the same eigenvectors, i.e., the optimal transmit directions are the eigenvectors of the channel covariance matrix. For the mean feedback case, the eigenvectors of the optimal transmit covariance matrix were shown to be the same as the right singular vectors of the channel mean matrix for a MISO system in [5] and for a MIMO system in [6]. In [8]–[10], we generalized these results, both in covariance and mean feedback cases, to MIMO-MAC systems. We showed that in a MIMO-MAC system with partial CSI at the transmitters, all users should transmit in the direction of the eigenvectors of their *own* channel parameter matrices. Consequently, we showed that, the transmit directions of the users in a MIMO-MAC with partial CSI at the transmitters 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.

On the other hand, in this aforementioned literature, the optimization of the eigenvalues of the transmit covariance matrices, i.e., the power allocation policies, are left as additional optimization problems. The optimum eigenvalues are known only for specific conditions, called *beamforming optimality conditions*. If the channel statistics satisfy these conditions, then unit rank transmit covariance matrices are optimum for all users, i.e., users allocate all of their powers to the direction of their strongest eigenvectors. References [5], [6], and [9]–

Manuscript received June 20, 2006; revised January 1, 2007. This work was supported by NSF Grants CCR 03-11311, CCF 04-47613 and CCF 05-14846, and presented in part at IEEE Asilomar Conference 2006 [1].

The authors are with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA (e-mail: {alkan,ulukus}@umd.edu).

Digital Object Identifier 10.1109/JSAC.2007.070913.

[10] derived beamforming optimality conditions in single-user MISO, single-user MIMO, and multi-user MIMO-MAC, respectively.

Although having beamforming optimality conditions is extremely helpful, beamforming is unconditionally optimal only when the number of users grows to infinity in a fading multi-user MIMO setting when partial CSI is available at the transmitters [10]. In a single-user MIMO or in a MIMO-MAC with finite number of users, the channel statistics might be so that beamforming may never be optimal. For such scenarios, efficient and globally convergent algorithms are needed in order to solve for the optimum eigenvalues of the transmit covariance matrices. References [11] and [12]–[13] proposed algorithms that solve this problem for a single-user MISO system, and for a single-user MIMO system, respectively. However, in both cases, the convergence proofs for these algorithms were not provided. In a MIMO-MAC scenario with partial CSI available at the transmitters, although the eigenvectors of the optimal transmit covariance matrices are known [8]–[10], no algorithm is available to find the optimum eigenvalues in a multi-user setting.

In this paper, first, we give an alternative derivation for the algorithm proposed in [12], [13] for a single-user MIMO system by enforcing the KKT optimality conditions at each iteration. We prove that the convergence point of this algorithm is unique and is equal to the optimum eigenvalue allocation. The proposed algorithm converges to this unique point starting from any point on the space of feasible eigenvalues. Next, we consider the multi-user version of the problem. In this case, the problem is to find the optimum eigenvalues of the transmit covariance matrices of all users that maximize the sum rate of the MIMO-MAC system. We apply the single-user algorithm iteratively to reach the global optimum point. At any given iteration, the multi-user algorithm updates the eigenvalues of one user, using the algorithm proposed for the single-user case, assuming that the eigenvalues of the remaining users are fixed. The algorithm iterates over all users in a round-robin fashion. We prove that, this algorithm converges to the unique global optimum power allocation for all users.

## 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<sup>1</sup>. The receiver has the perfect knowledge of the channel, while the transmitters have only the statistical model of the channel. Each transmitter sends a vector  $\mathbf{x}_k$ , and the received vector is

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

where  $K$  is the number of users,  $\mathbf{n}$  is a zero-mean, identity-covariance 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  $P_k$ ,  $\text{tr}(\mathbf{Q}_k) \leq P_k$ .

The statistical model that we consider in this paper 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. In this model, there exists correlation between the signals transmitted by or received at different antenna elements. For each user, the channel is modeled as [14],

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

where the receive antenna correlation matrix,  $\Phi_k$ , is the correlation between the signals transmitted by user  $k$ , and received at the  $n_R$  receive antennas of the receiver, and the transmit antenna correlation matrix,  $\Sigma_k$ , 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 [14] to every single transmitter-receiver link. In this paper, we will assume that the receiver does not have any physical restrictions and therefore, there is sufficient spacing between the antenna elements on the receiver such that the signals received at different antenna elements are uncorrelated<sup>2</sup>. As a result, the receive antenna correlation matrix becomes the identity matrix, i.e.,  $\Phi_k = \mathbf{I}$ . Now, the channel of user  $k$  is written as

$$\mathbf{H}_k = \mathbf{Z}_k \Sigma_k^{1/2} \quad (3)$$

where the entries of  $\mathbf{Z}_k$  are i.i.d., zero-mean, unit-variance complex Gaussian random variables. From this point on, we will refer to matrix  $\Sigma_k$  as the channel covariance feedback matrix of user  $k$ . Similar covariance feedback models have been used in [5], [6], [7], [11].

## III. POWER ALLOCATION FOR SINGLE-USER MIMO

In this section, we will assume that  $K = 1$ . In a single-user system with partial CSI in the form of the channel covariance matrix at the transmitter, the optimization problem is that of choosing a transmit covariance matrix  $\mathbf{Q}$ , which is subject to a trace constraint representing the average transmit power constraint,

$$C = \max_{\text{tr}(\mathbf{Q}) \leq P} E [\log |\mathbf{I}_{n_R} + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger|] \quad (4)$$

where  $E[\cdot]$  is the expectation operator with respect to the channel matrix  $\mathbf{H}$ , and  $|\cdot|$  is the determinant operator. We note that the cost function of the optimization problem in (4) is concave in  $\mathbf{Q}$  and the constraint set is convex.

The channel covariance matrix  $\Sigma$ , which is known at the transmitter, has the eigenvalue decomposition  $\Sigma = \mathbf{U}_\Sigma \Lambda_\Sigma \mathbf{U}_\Sigma^\dagger$  with unitary  $\mathbf{U}_\Sigma$  and diagonal  $\Lambda_\Sigma$  of ordered eigenvalues. The transmit covariance matrix  $\mathbf{Q}$  has the eigenvalue decomposition  $\mathbf{Q} = \mathbf{U}_Q \Lambda_Q \mathbf{U}_Q^\dagger$  with unitary  $\mathbf{U}_Q$  and diagonal  $\Lambda_Q$ . It has been shown that the eigenvectors of the optimum transmit covariance matrix must be equal to the eigenvectors of the

<sup>1</sup>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.

<sup>2</sup>We refer the reader to Section VII, for a discussion on extending 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.

channel covariance matrix, i.e.,  $\mathbf{U}_Q = \mathbf{U}_\Sigma$  [6]. By inserting this into (4), and using the fact that the random matrices  $\mathbf{Z}\mathbf{U}_\Sigma$  and  $\mathbf{Z}$  have the same probability distribution for zero-mean identity-covariance Gaussian  $\mathbf{Z}$  and unitary  $\mathbf{U}_\Sigma$  [2], we get

$$C = \max_{\text{tr}(\mathbf{\Lambda}_Q) \leq P} E \left[ \log \left| \mathbf{I}_{n_R} + \mathbf{Z}\mathbf{U}_\Sigma \mathbf{\Lambda}_\Sigma^{1/2} \mathbf{U}_\Sigma^\dagger \mathbf{U}_Q \mathbf{\Lambda}_Q \mathbf{U}_Q^\dagger \right. \right. \\ \left. \left. \mathbf{U}_\Sigma \mathbf{\Lambda}_\Sigma^{1/2} \mathbf{U}_\Sigma^\dagger \mathbf{Z}^\dagger \right| \right] \quad (5)$$

$$= \max_{\text{tr}(\mathbf{\Lambda}_Q) \leq P} E \left[ \log \left| \mathbf{I}_{n_R} + \mathbf{Z}\mathbf{U}_\Sigma \mathbf{\Lambda}_Q \mathbf{\Lambda}_\Sigma \mathbf{U}_\Sigma^\dagger \mathbf{Z}^\dagger \right| \right] \quad (6)$$

$$= \max_{\text{tr}(\mathbf{\Lambda}_Q) \leq P} E \left[ \log \left| \mathbf{I}_{n_R} + \mathbf{Z}\mathbf{\Lambda}_Q \mathbf{\Lambda}_\Sigma \mathbf{Z}^\dagger \right| \right] \quad (7)$$

$$= \max_{\sum_{i=1}^{n_T} \lambda_i^Q \leq P} E \left[ \log \left| \mathbf{I}_{n_R} + \sum_{i=1}^{n_T} \lambda_i^Q \lambda_i^\Sigma \mathbf{z}_i \mathbf{z}_i^\dagger \right| \right] \quad (8)$$

where  $\mathbf{z}_i$  is the  $i^{\text{th}}$  column of  $\mathbf{Z}$ , i.e.,  $\{\mathbf{z}_i, i = 1, \dots, n_T\}$  is a set of  $n_R \times 1$  dimensional i.i.d., zero-mean, identity-covariance Gaussian random vectors. The Lagrangian for the above optimization problem is,

$$L = E \left[ \log \left| \mathbf{I}_{n_R} + \sum_{i=1}^{n_T} \lambda_i^Q \lambda_i^\Sigma \mathbf{z}_i \mathbf{z}_i^\dagger \right| \right] - \mu \left( \sum_{i=1}^{n_T} \lambda_i^Q - P \right) \quad (9)$$

where  $\mu$  is the Lagrange multiplier. In order to derive the KKT conditions, we need the following identity which is proved in [6],

$$\frac{\partial}{\partial x} \log |\mathbf{A} + x\mathbf{B}| = \text{tr} [(\mathbf{A} + x\mathbf{B})^{-1} \mathbf{B}] \quad (10)$$

Using this identity, the KKT conditions can be written for  $i = 1, \dots, n_T$  as

$$\lambda_i^\Sigma E \left[ \mathbf{z}_i^\dagger \left( \mathbf{I}_{n_R} + \sum_{j=1}^{n_T} \lambda_j^Q \lambda_j^\Sigma \mathbf{z}_j \mathbf{z}_j^\dagger \right)^{-1} \mathbf{z}_i \right] \leq \mu \quad (11)$$

Defining  $\mathbf{A} = \mathbf{I}_{n_R} + \sum_{j=1}^{n_T} \lambda_j^Q \lambda_j^\Sigma \mathbf{z}_j \mathbf{z}_j^\dagger$ , and  $\mathbf{A}_i = \mathbf{A} - \lambda_i^Q \lambda_i^\Sigma \mathbf{z}_i \mathbf{z}_i^\dagger$ , and using the matrix inversion lemma [15, page 19], we get for  $i = 1, \dots, n_T$

$$E_i(\boldsymbol{\lambda}^Q) \triangleq E \left[ \frac{\lambda_i^\Sigma \mathbf{z}_i^\dagger \mathbf{A}_i^{-1} \mathbf{z}_i}{1 + \lambda_i^Q \lambda_i^\Sigma \mathbf{z}_i^\dagger \mathbf{A}_i^{-1} \mathbf{z}_i} \right] \leq \mu, \quad (12)$$

where we defined the left hand side of (12) as  $E_i(\boldsymbol{\lambda}^Q)$ . The  $i^{\text{th}}$  inequality in (12) is satisfied with equality whenever the optimum  $\lambda_i^Q$  is non-zero, and with strict inequality whenever the optimum  $\lambda_i^Q$  is zero. We note that in classical water-filling solutions, since the channel is either fixed or known instantaneously at the transmitter, the corresponding KKT conditions do not involve an expectation, and therefore, non-zero  $\lambda_i^Q$ 's can be solved for in terms of the Lagrange multiplier and the eigenvalues of the fixed/instantaneous channel matrix. However, in our case, we cannot directly solve for  $\lambda_i^Q$  in (12). Instead, we multiply both sides of (12) by  $\lambda_i^Q$ ,

$$\lambda_i^Q E_i(\boldsymbol{\lambda}^Q) = \mu \lambda_i^Q, \quad i = 1, \dots, n_T \quad (13)$$

We note that when  $\lambda_i^Q = 0$ , both sides of (13) are equal to zero. Therefore, unlike (12), (13) is always satisfied with equality for optimum eigenvalues. By summing both sides

over all antennas, we find  $\mu$ , and by substituting this  $\mu$  into (13), we find the fixed point equations which have to be satisfied by the optimum eigenvalues for  $i = 1, \dots, n_T$ ,

$$\lambda_i^Q = \frac{\lambda_i^Q E_i(\boldsymbol{\lambda}^Q)}{\sum_{j=1}^{n_T} \lambda_j^Q E_j(\boldsymbol{\lambda}^Q)} P = \frac{P}{\sum_j \frac{\lambda_j^Q E_j(\boldsymbol{\lambda}^Q)}{\lambda_j^Q E_i(\boldsymbol{\lambda}^Q)}} \triangleq f_i(\boldsymbol{\lambda}^Q) \quad (14)$$

where  $\boldsymbol{\lambda}^Q = [\lambda_1^Q, \dots, \lambda_{n_T}^Q]$ , and we defined the right hand side of (14) which depends on all of the eigenvalues as  $f_i(\boldsymbol{\lambda}^Q)$ . It is important to emphasize that the optimum solution of the KKT conditions always satisfies the fixed point in (14), even if the optimum solution has some zero components.

We propose to use the following fixed point algorithm

$$\boldsymbol{\lambda}^Q(n+1) = \mathbf{f}(\boldsymbol{\lambda}^Q(n)) \quad (15)$$

where  $\mathbf{f} = [f_1, \dots, f_{n_T}]$ . In order to solve for the optimum eigenvalues, (15) updates the eigenvalues at step  $n+1$  as a function of the eigenvalues at step  $n$ . We claim that this algorithm converges and that the unique stable fixed point of the algorithm is equal to the optimum eigenvalues. Although this algorithm is the same as the one proposed in [12], [13], here, we also provide a convergence proof.

#### IV. CONVERGENCE PROOF

As stated in (8), the constraint set of the optimization problem is  $\sum_{i=1}^{n_T} \lambda_i^Q \leq P$ . We know that the optimum value is obtained when the summation is equal to  $P$ . If the summation was strictly less than  $P$ , we could increase the value of the objective function by increasing any one of the  $\lambda_i^Q$ 's, while keeping the rest fixed. Therefore, the constraint set becomes  $\sum_{i=1}^{n_T} \lambda_i^Q = P$ . This equality defines a simplex in the  $n_T$ -dimensional space (see Fig. 1), and all feasible eigenvalue vectors are located on this simplex. Note that if the algorithm is initiated at an exact corner point of the simplex, then the updates stay at the same point indefinitely. The reason for this is that while we obtain (13) from (12), we create some artificial fixed points. That is, although some non-optimum  $\lambda_i^Q = 0$  does not satisfy (12) with equality, the same non-optimum  $\lambda_i^Q = 0$  always satisfies (13) with equality.

As a result, in addition to the point that is the solution of the KKT conditions, the solution set of the fixed point equation in (14) includes some artificial fixed points. Since our optimization problem is concave and the constraint set is convex, the solution of the KKT conditions is the unique optimum point of the optimization problem. On the other hand, artificial fixed points are the solutions to some reduced optimization problems, which are obtained by forcing some of the components of the power allocation vector to be zero. When we force a choice of  $n_T - 1$  components to be zero, we can find one optimum solution to the corresponding reduced optimization problem for each choice. Since there are  $\binom{n_T}{n_T-1}$  ways of choosing zero components, this adds  $n_T$  artificial fixed points, which are the corner points of the simplex, to the solution set of the fixed point equation. Similarly, when we force a choice of  $n_T - 2$  components to be zero, we can find one optimum solution to the corresponding reduced optimization problem for each choice. This adds  $\binom{n_T}{n_T-2}$  artificial fixed points to the solution set of the fixed point

equation. By counting all possibilities, we find that there are a total of  $2^{n_T} - 2$  artificial fixed points. However, it is important to note that one of these counted points might be the optimum solution of the KKT conditions, if there are some zero components in the optimum eigenvalue vector. If the optimum eigenvalues are all non-zero, then the solution of the KKT conditions is different than these artificial fixed points. Therefore, we call a point an artificial fixed point only if it is not the optimum solution.

In this section, we will first prove that our algorithm converges. Then, we will prove that the algorithm cannot converge to an artificial fixed point, and therefore, the only point that the algorithm can converge to is the unique solution of the KKT conditions. The main ingredient of our convergence proof is the following lemma.

*Lemma 1:* Let us have two feasible vectors on the simplex,  $\lambda^Q$  and  $\bar{\lambda}^Q$ , such that  $\lambda_i^Q > \bar{\lambda}_i^Q$ , then  $f_i(\lambda^Q) > f_i(\bar{\lambda}^Q)$ .

*Proof:* Note that  $\lambda_i^Q > \bar{\lambda}_i^Q$  implies  $\sum_{j \neq i} \lambda_j^Q < \sum_{j \neq i} \bar{\lambda}_j^Q$ , since all  $\lambda_i^Q$  sum up to  $P$ . Therefore, the lemma can be proved equivalently by proving that  $f_i(\lambda^Q)$  is increasing in  $\lambda_i^Q$  when the rest of the  $\lambda_j^Q$ ,  $j \neq i$  are fixed, and  $f_i(\lambda^Q)$  is decreasing in  $\sum_{j \neq i} \lambda_j^Q$ , when  $\lambda_i^Q$  is fixed. The first part of the claim is easy to show. Consider (14), it can be shown that the partial derivative of  $\lambda_i^Q E_i(\lambda^Q)$  with respect to  $\lambda_i^Q$  is positive, and the partial derivatives of  $\lambda_j^Q E_j(\lambda^Q)$ , for  $j \neq i$ , with respect to  $\lambda_i^Q$  are all negative. Therefore,  $\frac{\lambda_j^Q E_j(\lambda^Q)}{\lambda_i^Q E_i(\lambda^Q)}$  is decreasing (for all  $j$ ), and  $f_i(\lambda^Q)$  is increasing, in  $\lambda_i^Q$  when the rest of the  $\lambda_j^Q$ ,  $j \neq i$  are fixed. The second part of the claim is a little bit involved. In order to show that  $f_i(\lambda^Q)$  is decreasing in  $\sum_{j \neq i} \lambda_j^Q$ , we need to show that  $\sum_{j \neq i} \frac{\partial f_i(\lambda^Q)}{\partial \lambda_j^Q} < 0$ . It is sufficient to show  $\frac{\partial f_i(\lambda^Q)}{\partial \lambda_j^Q} < 0$  for all  $j \neq i$ . In order to show this, consider (14), it is easy to show that the partial derivative of  $\lambda_i^Q E_i(\lambda^Q)$  with respect to  $\lambda_j$  is negative. We, then, need to show that  $\frac{\partial(\sum_{k=1}^{n_T} \lambda_k^Q E_k(\lambda^Q))}{\partial \lambda_j^Q} > 0$ . We will give the proof of this in the Appendix.  $\square$

In Lemma 1, we showed the monotonicity property of the algorithm. By using this property, in the next lemma, we will show that the algorithm converges.

*Lemma 2:* The algorithm in (15) converges to one of the points in the solution set of the fixed point equation in (14) when it is initiated at any arbitrary feasible point,  $\lambda^Q(0)$ , that is not on the boundary of the simplex.

*Proof:* After the first iteration of the algorithm, we have one of the following three cases for each  $\lambda_i^Q$ . The first case is that  $\lambda_i^Q(1) = f_i(\lambda^Q(0)) = \lambda_i^Q(0)$ . This means that we have started the algorithm at the optimum point that solves the KKT conditions. Since all of the artificial fixed points are on the boundary of the simplex, this point cannot be an artificial fixed point.

The second case is that  $\lambda_i^Q(1) = f_i(\lambda^Q(0)) > \lambda_i^Q(0)$ . In this case, by applying Lemma 1 repeatedly, we get  $\lambda_i^Q(n) > \lambda_i^Q(n-1) > \dots > \lambda_i^Q(1) > \lambda_i^Q(0)$ . Since  $\lambda_i^Q(n)$  is a monotonically increasing sequence and it is upper bounded, it is guaranteed to converge.

The third case is that  $\lambda_i^Q(1) = f_i(\lambda^Q(0)) < \lambda_i^Q(0)$ . In this case, by applying Lemma 1 repeatedly, we get  $\lambda_i^Q(n) <$

$\lambda_i^Q(n-1) < \dots < \lambda_i^Q(1) < \lambda_i^Q(0)$ . Since  $\lambda_i^Q(n)$  is a monotonically decreasing sequence and it is lower bounded, it is guaranteed to converge.

Finally, since each component of  $\lambda^Q$  converges, the vector itself also converges to a point inside the solution set of the fixed point equation.  $\square$

Although we proved that the algorithm converges when it is initiated at any arbitrary feasible point that is not on the boundary of the simplex, there is a possibility that it converges to an artificial fixed point instead of the optimum solution of the KKT conditions. In the following lemma, we will show that this is never the case.

*Lemma 3:* The artificial fixed points are unstable. For a very small and fixed  $\epsilon$ , if we are  $\epsilon$  away from an artificial fixed point, with one iteration of the algorithm, we will move further away from that artificial fixed point.

*Proof:* The main idea of the proof is the following. We will start from an artificial fixed point that is not the optimum solution of the KKT conditions of the original optimization problem, and show that by perturbing this artificial fixed point by an  $\epsilon$  amount, we move further away from that artificial fixed point. We give the proof of the most general scenario with  $n_T$  antennas and starting from any arbitrary artificial fixed point in the Appendix. Here, we give the outline and the basic methodology of the general proof by considering a simple case where  $n_T = 3$ . In this case, we have  $2^{n_T} - 2 = 6$  artificial fixed points. Three of them are the corner points of the 3-dimensional simplex. The other three of them lie on the boundary of the simplex, each point corresponding to a solution of the reduced optimization problem where one of the components is forced to be zero. These points can be seen in Fig. 1.

Here, in this outline of the general proof, we will also assume that the artificial fixed point we focus on has only one zero component. In particular, we assume that we are at the artificial fixed point  $\mathbf{p}_4 = (a, b, 0)$ , see Fig. 1. Since this is a fixed point, the following equalities hold from (14),

$$a = \frac{aE_1(\mathbf{p}_4)}{aE_1(\mathbf{p}_4) + bE_2(\mathbf{p}_4)}P, \quad b = \frac{bE_2(\mathbf{p}_4)}{aE_1(\mathbf{p}_4) + bE_2(\mathbf{p}_4)}P \quad (16)$$

From above, we find that  $aE_1(\mathbf{p}_4) + bE_2(\mathbf{p}_4) = PE_1(\mathbf{p}_4) = PE_2(\mathbf{p}_4)$ . This is equivalent to saying that the KKT conditions of the reduced optimization problem corresponding to the first and second components are satisfied with equality, that is,  $E_1(\mathbf{p}_4) = E_2(\mathbf{p}_4) = \mu'$ . We call this Lagrange multiplier  $\mu'$ , because this is possibly different than the Lagrange multiplier of the original optimization problem. For  $E_3(\mathbf{p}_4)$ , we have three possibilities.  $E_3(\mathbf{p}_4) = \mu'$  cannot hold, because that would mean that the third KKT condition is also satisfied with equality and this can only happen when optimal  $\lambda_3^Q$  is non-zero.  $E_3(\mathbf{p}_4) < \mu'$  cannot hold, because that would mean that we satisfy all three KKT conditions of the original optimization problem with  $\mu' = \mu$ , and this fixed point is optimum. This contradicts with our assumption that we are at an artificial fixed point that is not the optimum solution of the original optimization problem. Therefore, the only possibility at an artificial fixed point is that  $E_3(\mathbf{p}_4) > \mu'$ .

Now, we will show that by perturbing this artificial fixed

point by an  $\epsilon$  amount, we move further away from this fixed point. We run the algorithm for  $\mathbf{p}'_4 = (a - \epsilon, b, \epsilon)$ . We first calculate  $E_1(\mathbf{p}'_4)$ ,

$$E_1(\mathbf{p}'_4) = E \left[ \frac{h(\epsilon)}{1 + (a - \epsilon)h(\epsilon)} \right] \quad (17)$$

where  $h(\epsilon)$  is the value of the function  $h(x) = \lambda_1^\Sigma \mathbf{z}_1^\dagger \left( \mathbf{A}_{p_4} - x(\lambda_1^\Sigma \mathbf{z}_1 \mathbf{z}_1^\dagger) \right)^{-1} \mathbf{z}_1$  evaluated at  $x = \epsilon$ , where  $\mathbf{A}_{p_4} = \mathbf{I}_{n_R} + a\lambda_1^\Sigma \mathbf{z}_1 \mathbf{z}_1^\dagger + b\lambda_2^\Sigma \mathbf{z}_2 \mathbf{z}_2^\dagger$ . Using the matrix inversion lemma [15, page 19], we get  $h(x) = \frac{\lambda_1^\Sigma \mathbf{z}_1^\dagger \mathbf{A}_{p_4} \mathbf{z}_1}{1 - x\lambda_1^\Sigma \mathbf{z}_1^\dagger \mathbf{A}_{p_4} \mathbf{z}_1}$ . Using the Taylor series expansion formula around  $x = 0$ , and denoting  $w_1 = \lambda_1^\Sigma \mathbf{z}_1^\dagger \mathbf{A}_{p_4} \mathbf{z}_1$ , we obtain

$$h(\epsilon) = w_1 + \epsilon w_1^2 + \epsilon^2 w_1^3 + \dots \quad (18)$$

$$= w_1 + O(\epsilon) \quad (19)$$

where  $O(\epsilon)$  is an asymptotic upper bound for the magnitude of the residual in terms of  $\epsilon$ . Mathematically, a function,  $\bar{h}(\epsilon)$  is order  $O(\epsilon)$  as  $\epsilon \rightarrow 0$  if and only if  $0 < \limsup_{\epsilon \rightarrow 0} \frac{\bar{h}(\epsilon)}{\epsilon} < \infty$  [16]. Now, when we insert this into (17), we obtain

$$E_1(\mathbf{p}'_4) = E \left[ \frac{w_1 + O(\epsilon)}{1 + aw_1 + O(\epsilon)} \right] \quad (20)$$

$$= \frac{1}{a} \left( 1 - E \left[ \frac{1}{1 + aw_1 + O(\epsilon)} \right] \right) \quad (21)$$

We again use the Taylor series expansion formula, this time with  $h(x) = 1/x$ , around  $x = 1 + aw_1$ ,

$$h(1 + aw_1 + O(\epsilon)) = \frac{1}{1 + aw_1 + O(\epsilon)} \quad (22)$$

$$= \frac{1}{1 + aw_1} - \frac{O(\epsilon)}{(1 + aw_1)^2} + \dots \quad (23)$$

$$= \frac{1}{1 + aw_1} + O(\epsilon) \quad (24)$$

Finally, (17) becomes

$$E_1(\mathbf{p}'_4) = E \left[ \frac{w_1}{1 + aw_1} \right] + O(\epsilon) \quad (25)$$

$$= E_1(\mathbf{p}_4) + O(\epsilon) \quad (26)$$

By using similar arguments, we can conclude that  $E_i(\mathbf{p}'_4) = E_i(\mathbf{p}_4) + O(\epsilon)$ , for  $i = 1, 2, 3$ . If we insert these into  $f_3(\mathbf{p}'_4)$ , we obtain

$$f_3(\mathbf{p}'_4) = \frac{\epsilon E_3(\mathbf{p}_4)}{aE_1(\mathbf{p}_4) + bE_2(\mathbf{p}_4) + O(\epsilon)} P + O(\epsilon^2), \quad (27)$$

since the summation of terms that are in the order of  $O(\epsilon)$  and smaller will be in the order of  $O(\epsilon)$ . Finally, by applying Taylor series expansion one more time with  $h(x) = 1/x$ , we get

$$f_3(\mathbf{p}'_4) = \frac{\epsilon E_3(\mathbf{p}_4)}{aE_1(\mathbf{p}_4) + bE_2(\mathbf{p}_4)} P + O(\epsilon^2) \quad (28)$$

We know from (16) that  $aE_1(\mathbf{p}_4) + bE_2(\mathbf{p}_4) = PE_1(\mathbf{p}_4) = PE_2(\mathbf{p}_4)$ . Inserting this into the above equation, we have

$$f_3(\mathbf{p}'_4) = \epsilon \frac{E_3(\mathbf{p}_4)}{E_1(\mathbf{p}_4)} + O(\epsilon^2) \quad (29)$$

$$> \epsilon \quad (30)$$

where the last inequality follows from the fact that  $E_3(\mathbf{p}_4) > \mu' = E_1(\mathbf{p}_4)$ . This result tells us that starting from  $\epsilon$  away from an artificial fixed point, the third component of the updated vector, and therefore the updated vector itself moves further away from that artificial fixed point. Finally, by using Lemma 1, we note that the algorithm will move away from the artificial fixed point at each iteration. Therefore, this artificial fixed point is unstable.  $\square$

As a result of Lemma 3, the algorithm never converges to an artificial fixed point, if it is not initiated at the boundary of the simplex. Therefore, the point that the algorithm converges to, always satisfies the KKT conditions of the original optimization problem. Since this point is unique, when the algorithm converges, it does so to the unique optimum power allocation policy.

#### A. Comparison to Water-filling

In this section, we will compare our results to the classical water-filling solution. We note that in [2], the channel matrix  $\mathbf{H}$  is known to both the receiver and the transmitter. The singular value decomposition of  $\mathbf{H}$  can be written as  $\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^\dagger$ , where  $n_R \times n_R$  dimensional  $\mathbf{U}$ , and  $n_T \times n_T$  dimensional  $\mathbf{V}$  are unitary, and  $n_R \times n_T$  dimensional  $\mathbf{D}$  is non-negative and diagonal. Let the diagonal elements of  $\mathbf{D}$  be denoted by  $d_i$ , for  $i = 1, \dots, \min(n_R, n_T)$ . The solution of the KKT conditions for this case yields,

$$\lambda_i^Q = \left( \frac{1}{\mu} - \frac{1}{d_i} \right)^+, \quad i = 1, \dots, \min(n_R, n_T) \quad (31)$$

where  $(x)^+ = \max\{0, x\}$ . Although  $\lambda_i^Q$  is given explicitly, the Lagrange multiplier  $\mu$  still has to be solved. On the other hand, note that the algorithm proposed in this paper calculates the eigenvalues directly, without the need for calculating the Lagrange multiplier of the KKT conditions. Considering this fact, we can propose the following new algorithm for the water-filling solution in [2], using the idea in this paper. For  $i = 1, \dots, \min(n_R, n_T)$ , we have

$$\lambda_i^Q(n+1) = \frac{\frac{\lambda_i^Q(n)d_i}{1+\lambda_i^Q(n)d_i}}{\sum_j \frac{\lambda_j^Q(n)d_j}{1+\lambda_j^Q(n)d_j}} P \quad (32)$$

Note that this algorithm has the same properties as (15), and finds the optimum eigenvalues without the need for calculating the Lagrange multiplier  $\mu$ .

#### V. POWER ALLOCATION FOR MULTI-USER MIMO

The sum capacity of a MIMO-MAC is given as [17],

$$C_{sum} = \max_{\substack{u(\mathbf{Q}_k) \leq P_k \\ k=1, \dots, K}} E \left[ \log \left| \mathbf{I}_{n_R} + \sum_{k=1}^K \mathbf{H}_k \mathbf{Q}_k \mathbf{H}_k^\dagger \right| \right] \quad (33)$$

Let  $\Sigma_k = \mathbf{U}_{\Sigma_k} \Lambda_{\Sigma_k} \mathbf{U}_{\Sigma_k}^\dagger$  be the spectral decomposition of the channel covariance matrix of user  $k$ . Then, the optimum transmit covariance matrix  $\mathbf{Q}_k$  of user  $k$  has the form  $\mathbf{Q}_k = \mathbf{U}_{\Sigma_k} \Lambda_{Q_k} \mathbf{U}_{\Sigma_k}^\dagger$ , for all users [10]. This means that each user transmits along the directions of its own channel covariance matrix. While proving this in [10], we used the fact that the

random matrices  $\{\mathbf{Z}_k \mathbf{U}_{\Sigma_k}\}_{k=1}^K$  and  $\{\mathbf{Z}_k\}_{k=1}^K$  have the same joint distribution for zero-mean identity-covariance Gaussian  $\{\mathbf{Z}_k\}_{k=1}^K$  and unitary  $\{\mathbf{U}_{\Sigma_k}\}_{k=1}^K$ . Since the structure of the sum capacity expression is similar to the single-user capacity expression except for the summation inside the determinant, single-user solution easily generalizes to the multi-user case. By inserting this into (33), we get

$$C_{sum} = \max_{\substack{\text{tr}(\mathbf{A}_{Q_k}) \leq P_k \\ k=1, \dots, K}} E \left[ \log \left| \mathbf{I}_{n_R} + \sum_{k=1}^K \mathbf{Z}_k \mathbf{A}_{Q_k} \mathbf{A}_{\Sigma_k} \mathbf{Z}_k^\dagger \right| \right] \quad (34)$$

$$= \max_{\substack{\sum_{i=1}^{n_T} \lambda_{ki}^Q \leq P_k \\ k=1, \dots, K}} E \left[ \log \left| \mathbf{I}_{n_R} + \sum_{k=1}^K \sum_{i=1}^{n_T} \lambda_{ki}^Q \lambda_{ki}^{\Sigma} \mathbf{z}_{ki} \mathbf{z}_{ki}^\dagger \right| \right] \quad (35)$$

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

A result of [10] 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. If the eigenvalues of the channel covariance matrices satisfy the conditions given in [10], then beamforming becomes optimal, and 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. However, if the eigenvalues of the channel covariance matrices do not satisfy these conditions, finding the optimum eigenvalues becomes a harder task. In this section, we will give an iterative algorithm that finds the optimum eigenvalues for all users. We will follow a similar direction as in the single-user case. By writing the Lagrangian for (35) and using the identity in (10), we obtain the  $i^{\text{th}}$  KKT condition for user  $k$  as

$$E_{ki}(\boldsymbol{\lambda}^Q) \triangleq E \left[ \frac{\lambda_{ki}^{\Sigma} \mathbf{z}_{ki}^\dagger \mathbf{A}_{ki}^{-1} \mathbf{z}_{ki}}{1 + \lambda_{ki}^Q \lambda_{ki}^{\Sigma} \mathbf{z}_{ki}^\dagger \mathbf{A}_{ki}^{-1} \mathbf{z}_{ki}} \right] \leq \mu_k \quad (36)$$

where  $\boldsymbol{\lambda}^Q = [\lambda_1^Q, \dots, \lambda_K^Q]$ ,  $\boldsymbol{\lambda}_k^Q = [\lambda_{k1}^Q, \dots, \lambda_{kn_T}^Q]$  is the eigenvalue vector of user  $k$ , and  $\mu_k$  is the Lagrange multiplier corresponding to user  $k$ ,  $\mathbf{A}_{ki} = \mathbf{A} - \lambda_{ki}^Q \lambda_{ki}^{\Sigma} \mathbf{z}_{ki} \mathbf{z}_{ki}^\dagger$ , and  $\mathbf{A} = \mathbf{I}_{n_R} + \sum_{k=1}^K \sum_{j=1}^{n_T} \lambda_{kj}^Q \lambda_{kj}^{\Sigma} \mathbf{z}_{kj} \mathbf{z}_{kj}^\dagger$ . The inequalities in (36) are satisfied with equality whenever the optimum  $\lambda_{ki}^Q$  is non-zero, and with strict inequality whenever the optimum  $\lambda_{ki}^Q$  is zero. Similar to the single-user case,  $\lambda_{ki}^Q$  cannot be solved directly from (36) because of the expectation operator. Again, we will multiply both sides of (36) by  $\lambda_{ki}^Q$ ,

$$\lambda_{ki}^Q E_{ki}(\boldsymbol{\lambda}^Q) = \lambda_{ki}^Q \mu_k \quad (37)$$

Note that, similar to the single-user case, (37) is satisfied with equality for all  $\lambda_{ki}^Q$ , and we have created some artificial fixed points while obtaining (37) from (36). For any  $k$ , we can find  $\mu_k$  by summing over all antennas, and by inserting this  $\mu_k$  into (37), we can find the fixed point equations that have to

be satisfied by the optimum power values of user  $k$ ,

$$\lambda_{ki}^Q = \frac{\lambda_{ki}^Q E_{ki}(\boldsymbol{\lambda}^Q)}{\sum_j \lambda_{kj}^Q E_{kj}(\boldsymbol{\lambda}^Q)} P_k \triangleq g_{ki}(\boldsymbol{\lambda}^Q), \quad i = 1, \dots, n_T \quad (38)$$

where we defined the right hand side of (38) which depends on all of the eigenvalues as  $g_{ki}(\boldsymbol{\lambda}^Q)$ .

We propose the following algorithm, that enforces (38),

$$\boldsymbol{\lambda}_k^Q(n+1) = \mathbf{g}_k \left( \lambda_1^Q, \dots, \lambda_{k-1}^Q, \lambda_k^Q(n), \lambda_{k+1}^Q, \dots, \lambda_K^Q \right) \quad (39)$$

where  $\mathbf{g}_k = [g_{k1}, \dots, g_{kn_T}]$  is the vector valued update function of user  $k$ . This algorithm finds the optimum eigenvalues of a given user by assuming that the eigenvalues of the rest of the users are fixed. The algorithm moves to another user, after (39) converges. A complete update corresponding to user  $k$  only, i.e., running the algorithm in (39) for user  $k$  until it converges while the eigenvalues of the other users are fixed, is equivalent to the single-user algorithm proposed in (15). Therefore, we know from the previous section that the algorithm in (39) converges to the unique optimum point, when the eigenvalues of the rest of the users are fixed. The optimization problem that is solved by (39) is,

$$C_k = \max_{\sum_{i=1}^{n_T} \lambda_{ki}^Q \leq P_k} E \left[ \log \left| \mathbf{B}_k + \sum_{i=1}^{n_T} \lambda_{ki}^Q \lambda_{ki}^{\Sigma} \mathbf{z}_{ki} \mathbf{z}_{ki}^\dagger \right| \right] \quad (40)$$

where  $\mathbf{B}_k = \mathbf{I}_{n_R} + \sum_{l \neq k} \sum_{i=1}^{n_T} \lambda_{li}^Q \lambda_{li}^{\Sigma} \mathbf{z}_{li} \mathbf{z}_{li}^\dagger$  depends on the fixed eigenvalues of all other users. Such an algorithm is guaranteed to converge to the global optimum [18, page 219], since  $C_{sum}$  is a concave function of  $\lambda_{ki}$  for all  $k$  and  $i$ ,  $C_k$  is a strictly concave function of  $\lambda_{ki}$  for all  $i$ , and the constraint set is convex and has a Cartesian product structure among the users. Note that in [3], this kind of an algorithm is used in order to find the iterative water-filling solution. However, in that setting, where both the receiver and the transmitters know the perfect CSI, an iteration corresponding to user  $k$  does not include another algorithm, but it is just a single-user water-filling solution.

In order to improve the convergence rate, we also propose the following multi-user algorithm,

$$\boldsymbol{\lambda}_{k'}^Q(n+1) = \mathbf{g}_{k'} \left( \lambda_1^Q(n+1), \dots, \lambda_{k'-1}^Q(n+1), \lambda_{k'}^Q(n), \lambda_{k'+1}^Q(n), \dots, \lambda_K^Q(n) \right) \quad (41)$$

where  $k' = (n+1) \bmod K$ . At a given time  $n+1$ , this algorithm updates the eigenvalues of user  $k'$ . In the next iteration, it moves to another user. Since at a given iteration corresponding to user  $k$ , this algorithm does not solve (40) completely, we cannot conclude its convergence using [18, page 219]. However, we have observed the convergence of this algorithm experimentally through many simulations. One potential method to prove the convergence of this algorithm could be through proving that each iteration of the single-user algorithm in (15) increases the objective function of the optimization problem, i.e., the rate. Even though we proved that each iteration of this algorithm either monotonically increases or monotonically decreases each eigenvalue, and therefore, monotonically decreases the distance between the iterated eigenvalue vector and the optimum eigenvalue vector, we have

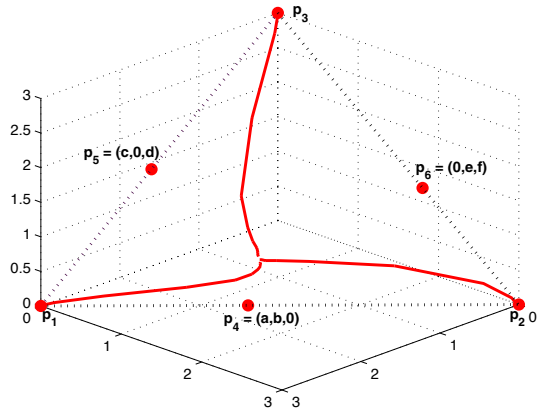


Fig. 1. The trajectories of the single-user algorithm when it is started from the corner points of the simplex for the case where the optimal eigenvalues are all non-zero.

not been able to prove mathematically that each iteration monotonically increases the objective function. Yet, we have observed this monotonicity through extensive simulations. Given that the objective function is a strictly concave function of the eigenvalue vector, we conjecture that the algorithm in (15) increases the objective function monotonically. On the other hand, we have observed experimentally that the algorithm in (41) converges much faster than the algorithm in (39). This could be due to the fact that, while the algorithm in (39) runs many iterations of the same user before it moves to another user, the algorithm in (41) runs only one iteration for each user before it moves to the next user.

## VI. NUMERICAL RESULTS

In this section, we will provide numerical examples for the performances of the proposed algorithms. In Fig. 1 and Fig. 2, we plot the trajectories of the iterations of the proposed single-user algorithm for a MIMO system with  $n_R = n_T = P = 3$ . We run the algorithm three times for each figure with different initial points, which are  $\epsilon$  away from the three corner points of the 3-dimensional simplex. In Fig. 1, all of the optimum eigenvalues are non-zero, and in Fig. 2, one of the optimum eigenvalues is zero. We observe, from the two figures, that the algorithm converges to the unique optimum point.

In Fig. 3 and Fig. 4, we plot the eigenvalues as a function of the iteration index. We observe that the eigenvalues converge to the same unique convergence point starting from various initial points. In addition to the points that are  $\epsilon$  away from the corner points, the other initial points are: the all-one vector, and the point corresponding to the channel covariance matrix eigenvalues, which is normalized to satisfy the power constraint. In Fig. 3, all of the optimum eigenvalues are non-zero, and in Fig. 4, one of the optimum eigenvalues is zero. As we see from Fig. 3, the algorithm needs much less time to converge to the optimum point when it is started from the normalized channel covariance eigenvalue point compared to the cases when it is started from any other points on the simplex. This is true mainly because of an argument similar to the water-filling argument, where we allocate more power to the strongest channel. As a result, the unique optimum transmit

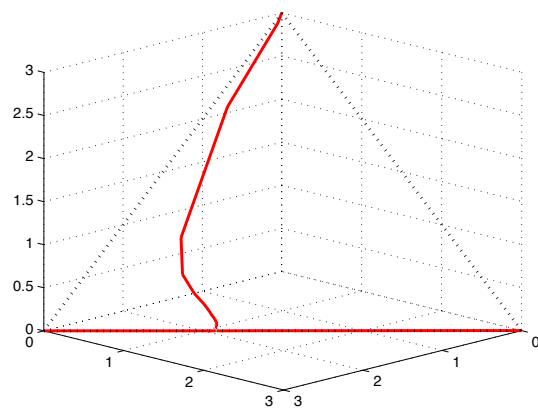


Fig. 2. The trajectories of the single-user algorithm when it is started from the corner points of the simplex for the case where one of the optimal eigenvalues is zero.

covariance eigenvalue vector is located close to the normalized channel covariance eigenvalue vector. Since they are located close by, it takes less time for the algorithm to converge to the optimum. Therefore, we may prefer to start the algorithm from the normalized channel covariance matrix eigenvalues, in order to improve the convergence rate of the algorithm. We note however that the algorithm converges to the optimum point from any arbitrary initial point.

We also note that, even when we start the algorithm from the normalized channel covariance matrix eigenvalues, we observe from Fig. 4 that it may still take some time for the algorithm to converge. In this case, this occurs mainly because one of the optimum eigenvalues is equal to zero. In order to improve the convergence rate, we can check if any one of the optimum eigenvalues will be zero, before we start the algorithm. We can use the beamforming optimality conditions [6], [10] in order to check if the second component of the eigenvalue vector is zero. For the rest of the components, similar conditions can easily be derived by using the ideas in [6], [10]. If there are any eigenvalues that will be zero at the optimum, we can drop them from the optimization problem, and solve a reduced problem with fewer dimensions. In Fig. 5, we have selected the eigenvalues of the channel covariance matrix so that the third eigenvalue of the optimum transmit covariance matrix happens to be zero. We considered two different initial points: the normalized channel covariance eigenvalue vector, and a vector obtained by setting the third component of the channel covariance eigenvalue vector to zero, before the normalization. We observe that the algorithm converges much faster if we identify the components that will be zero at the convergence point and remove them from the iterations.

Finally, we consider a multi-user MIMO-MAC scenario. Note that, for a given user, the multi-user algorithm given in (39) demonstrates the same convergence behavior as in Fig. 3 and Fig. 4, when the eigenvalues of the other users are kept constant. Therefore, we plot Fig. 6 by running the multi-user algorithm proposed in (41). In this figure, we consider 3 users with different channel covariance matrices. The algorithm is started at the normalized channel covariance eigenvalue vectors of the users. Each iteration in the figure corresponds

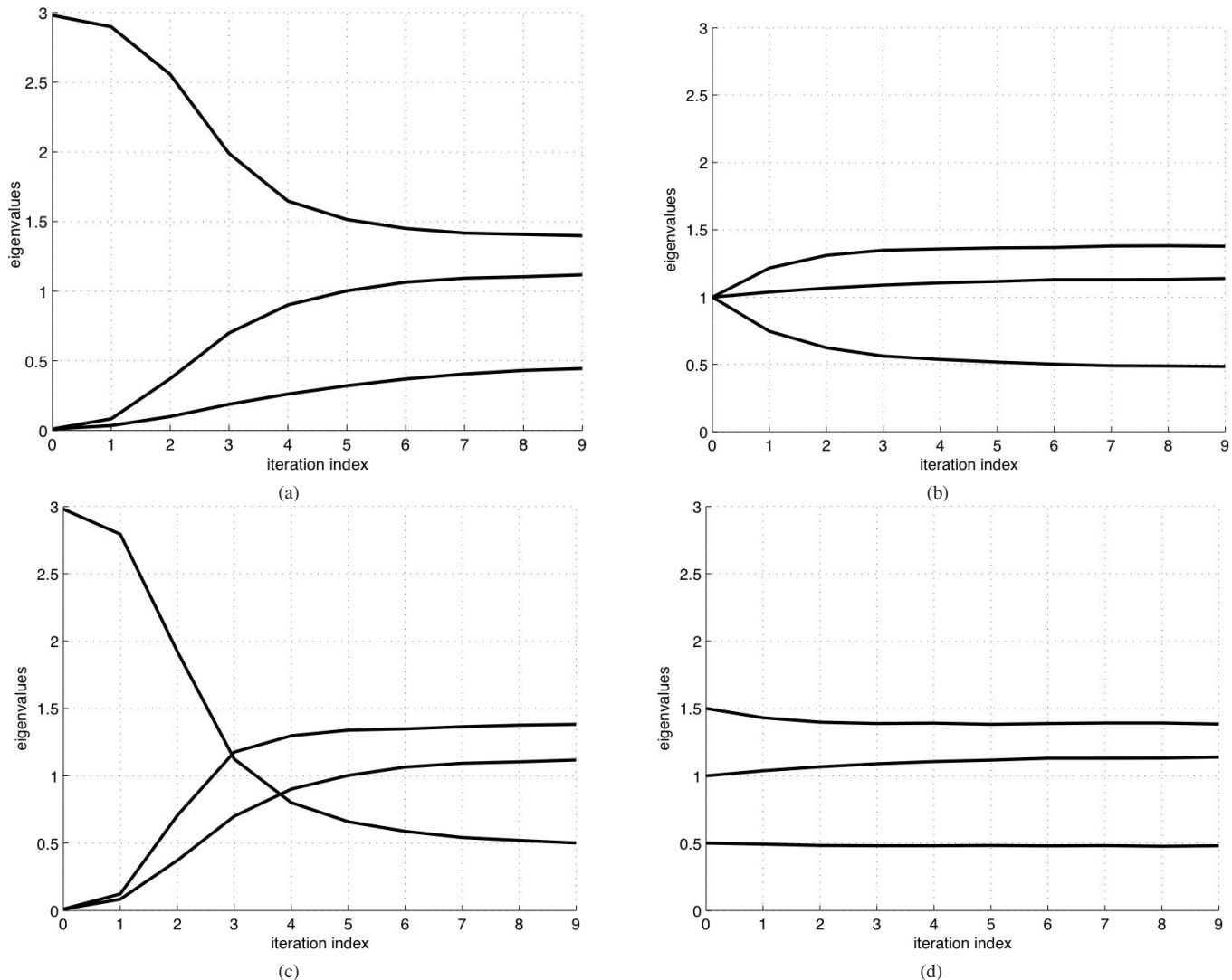


Fig. 3. The convergence of the single-user algorithm starting from various points, when all of the optimal eigenvalues are non-zero: (a) convergence of all three eigenvalues from  $(P - 2\epsilon, \epsilon, \epsilon)$ ; (b) convergence of all three eigenvalues from  $(\frac{P}{n_T}, \frac{P}{n_T}, \frac{P}{n_T})$ ; (c) convergence of all three eigenvalues from  $(\epsilon, \epsilon, P - 2\epsilon)$ ; (d) convergence of all three eigenvalues from the normalized channel eigenvalue vector.

to an update of the eigenvalues of the transmit covariance matrices of all users. At the end of the first iteration, all users have run the algorithm in (41) once. We can see in Fig. 6 that the multi-user algorithm converges quite quickly, and at the end of the fourth iteration, all users are almost at their optimum eigenvalue points.

## VII. CONCLUSIONS AND DISCUSSIONS

We proposed globally convergent algorithms for finding the optimum power allocation policies for both single-user MIMO and MIMO-MAC systems. Combining this with our previous results on the optimum transmit directions and the asymptotic behavior of MIMO-MAC systems [10], the sum capacity maximization problem is completely solved for a finite or infinite sized MIMO-MAC with the full CSI at the receiver and the partial CSI at the transmitters in the form of channel covariance information. In this paper, for a single-user case, we proved the convergence and the uniqueness of the convergence point of a pre-existing algorithm. This proof handles the complications arising from the existence of the artificial fixed points, and it gives some insights to the classical

water-filling solution. For the multi-user case, we derived and proved the convergence of a multi-user algorithm, which finds the optimum power allocations of all users.

Due to the nature of our optimization problem, our algorithms include calculation of some expectations at each iteration. Direct calculation of these expectations is sometimes difficult. However, by exploiting the ergodicity of the system and using sample averages, we can get very fast results. Although the number of expectations that has to be calculated increases as the number of users increases, fortunately, we can eliminate most of the components inside the expectations using the results of [10], which state that beamforming becomes optimal as the number of users in the system increases. As it can be seen in [10], even for a fairly low number of users, beamforming is almost optimal. Therefore, by combining beamforming optimality conditions with the proposed algorithms, we can find the optimum power allocations of the users much faster. Fig. 5 shows that the number of iterations is significantly less when beamforming optimality conditions are utilized. Although it cannot be seen in the figure, each iteration takes less time as well, i.e., the expectations are computed



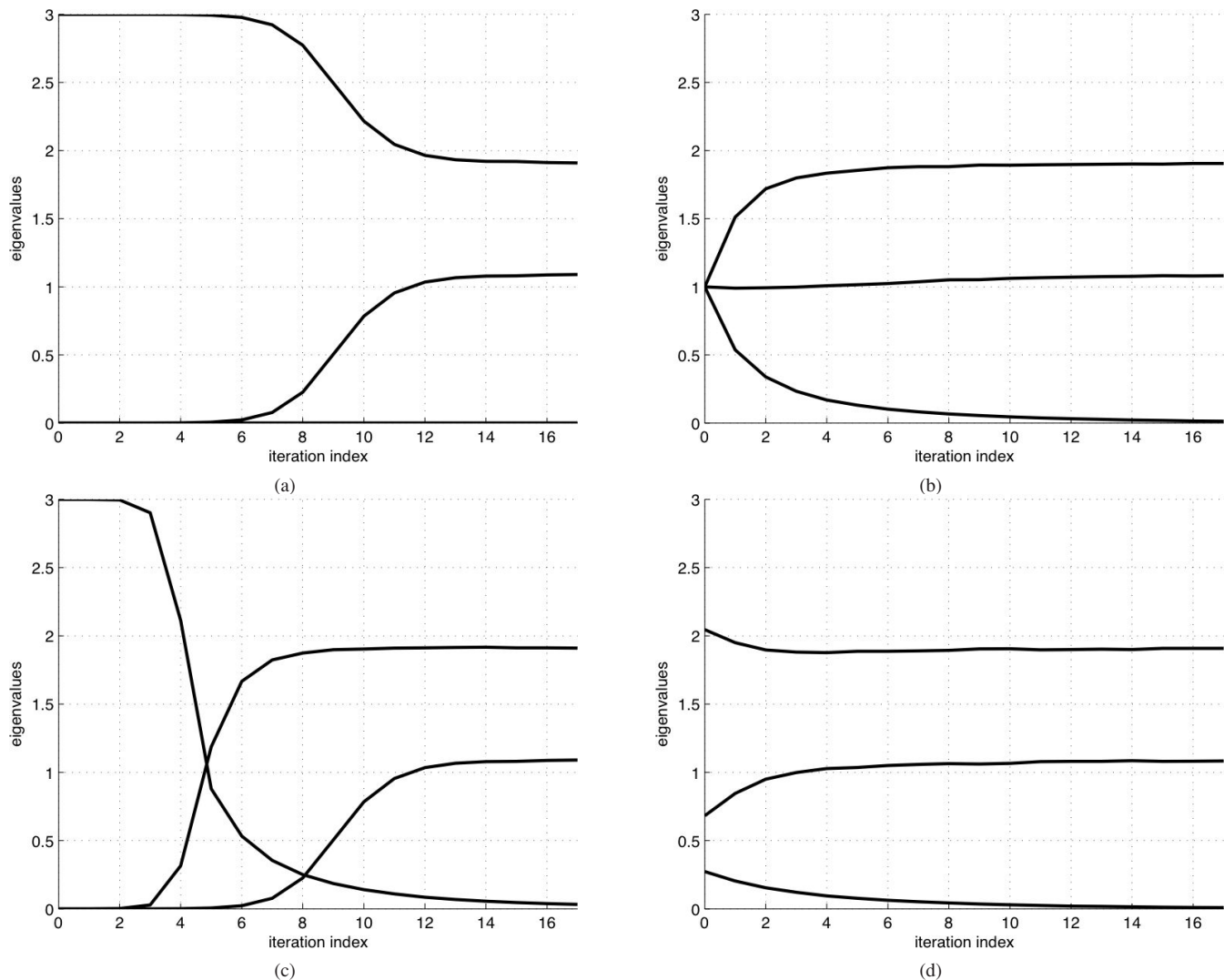


Fig. 4. The convergence of the single-user algorithm starting from various points, when one of the optimal eigenvalues is zero: (a) convergence of all three eigenvalues from  $(P - 2\epsilon, \epsilon, \epsilon)$ ; (b) convergence of all three eigenvalues from  $(\frac{P}{n_T}, \frac{P}{n_T}, \frac{P}{n_T})$ ; (c) convergence of all three eigenvalues from  $(\epsilon, \epsilon, P - 2\epsilon)$ ; (d) convergence of all three eigenvalues from the normalized channel eigenvalue vector.

faster, since there is less randomness in the system as a result of setting some eigenvalues to zero.

Another issue that we want to discuss here is the possibility of having a channel with double-sided correlation. In our model, 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 away from each other, the receiver side correlation matrix becomes the identity matrix. In a different 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 [19]. Beamforming optimality condition for this case is also found previously [19]. For the power allocation problem, an approach similar to the one in our paper can be applied and a similar but more cumbersome algorithm can be found. This algorithm includes extra terms that are similar to the terms in beamforming optimality conditions that are given in [19]. 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. This might be motivated by assuming that the receiver side channel correlation is only a result of the physical structure of the receiver and the environment around the receiver, therefore it is the same for all users. In this case, it is possible to find similar but again more cumbersome algorithms in order to solve the optimum power allocation policies of all users.

## VIII. APPENDIX

### A. Proof of Lemma 1

Without loss of generality, let us take  $j = 1$ . We will show that  $\frac{\partial(\sum_{k=1}^{n_T} \lambda_k^Q E_k(\lambda^Q))}{\partial \lambda_1^Q} > 0$ . It can be shown that

$$\frac{\partial \lambda_1^Q E_1(\lambda^Q)}{\partial \lambda_1^Q} = E \left[ \frac{\lambda_1^\Sigma \mathbf{z}_1^\dagger \mathbf{A}_1^{-1} \mathbf{z}_1}{\left(1 + \lambda_1^Q \lambda_1^\Sigma \mathbf{z}_1^\dagger \mathbf{A}_1^{-1} \mathbf{z}_1\right)^2} \right] \quad (42)$$

Now, for  $k = 2, \dots, n_T$ , let us consider  $\lambda_k^Q E_k(\lambda^Q) = \lambda_k^Q \lambda_k^\Sigma E[\mathbf{z}_k \mathbf{A}^{-1} \mathbf{z}_k]$ . Applying the matrix inversion lemma [15, page 19] to  $\mathbf{A} = \mathbf{A}_1 + \lambda_1^Q \lambda_1^\Sigma \mathbf{z}_1 \mathbf{z}_1^\dagger$ , we get

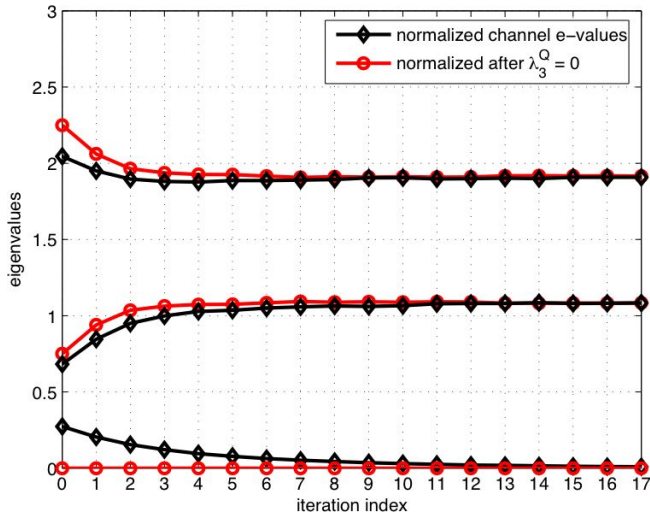


Fig. 5. The convergence of the single-user algorithm when one of the optimum eigenvalues is zero.

$$E_k(\lambda^Q) = \lambda_k^\Sigma \left( E[\mathbf{z}_k \mathbf{A}_1^{-1} \mathbf{z}_k] - E \left[ \frac{\lambda_1^Q \lambda_1^\Sigma (\mathbf{z}_k \mathbf{A}_1^{-1} \mathbf{z}_1)^2}{1 + \lambda_1^Q \lambda_1^\Sigma \mathbf{z}_1^\dagger \mathbf{A}_1^{-1} \mathbf{z}_1} \right] \right) \quad (43)$$

By taking the derivative of (43) with respect to  $\lambda_1^Q$ , we get

$$\frac{\partial E_k(\lambda^Q)}{\partial \lambda_1^Q} = -\lambda_k^\Sigma E \left[ \frac{\lambda_1^\Sigma (\mathbf{z}_k^\dagger \mathbf{A}_1^{-1} \mathbf{z}_1)^2}{(1 + \lambda_1^Q \lambda_1^\Sigma \mathbf{z}_1^\dagger \mathbf{A}_1^{-1} \mathbf{z}_1)^2} \right] \quad (44)$$

Combining (42) and (44) with  $\mathbf{s}_k = (\lambda_k^\Sigma \lambda_k^Q)^{1/2} \mathbf{z}_k$ , we have

$$\frac{\partial \left( \sum_{k=1}^{n_T} \lambda_k^Q E_k(\lambda^Q) \right)}{\partial \lambda_1^Q} = E \left[ \frac{\mathbf{s}_1^\dagger \mathbf{A}_1^{-1} \mathbf{s}_1 - \sum_{k=2}^{n_T} (\mathbf{s}_k^\dagger \mathbf{A}_1^{-1} \mathbf{s}_1)^2}{\lambda_1^Q (1 + \mathbf{s}_1^\dagger \mathbf{A}_1^{-1} \mathbf{s}_1)^2} \right] \quad (45)$$

We note that  $\mathbf{A}_1 = \mathbf{I} + \mathbf{S}_1 \mathbf{S}_1^\dagger$ , where  $\mathbf{S}_1 = [\mathbf{s}_2, \dots, \mathbf{s}_{n_T}]$ . Then, by using the matrix inversion lemma, we have  $\mathbf{A}_1^{-1} = \mathbf{I} - \mathbf{S}_1 (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} \mathbf{S}_1^\dagger$ . Finally, note that  $\sum_{k=2}^{n_T} (\mathbf{s}_k^\dagger \mathbf{A}_1^{-1} \mathbf{s}_1)^2 = \mathbf{s}_1^\dagger \mathbf{A}_1^{-1} \mathbf{S}_1 \mathbf{S}_1^\dagger \mathbf{A}_1^{-1} \mathbf{s}_1$ . Now, we will find equivalent expressions for the numerator of (45). Let us first look at  $\mathbf{s}_1^\dagger \mathbf{A}_1^{-1} \mathbf{s}_1$ ,

$$\mathbf{s}_1^\dagger \mathbf{A}_1^{-1} \mathbf{s}_1 = \mathbf{s}_1^\dagger \mathbf{s}_1 - \mathbf{s}_1^\dagger \mathbf{S}_1 (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} \mathbf{S}_1^\dagger \mathbf{s}_1 \quad (46)$$

Now, let us look at  $\mathbf{s}_1^\dagger \mathbf{A}_1^{-1} \mathbf{S}_1$ ,

$$\mathbf{s}_1^\dagger \mathbf{A}_1^{-1} \mathbf{S}_1 = \mathbf{s}_1^\dagger \mathbf{S}_1 - \mathbf{s}_1^\dagger \mathbf{S}_1 (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} \mathbf{S}_1^\dagger \mathbf{S}_1 \quad (47)$$

$$= \mathbf{s}_1^\dagger \mathbf{S}_1 - \mathbf{s}_1^\dagger \mathbf{S}_1 (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} (\mathbf{S}_1^\dagger \mathbf{S}_1 + \mathbf{I} - \mathbf{I}) \quad (48)$$

$$= \mathbf{s}_1^\dagger \mathbf{S}_1 (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} \quad (49)$$

Inserting (46) and (49) into (45), it is sufficient to show that the expression

$$\mathbf{s}_1^\dagger \mathbf{s}_1 - \mathbf{s}_1^\dagger \mathbf{S}_1 (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} \mathbf{S}_1^\dagger \mathbf{s}_1 - \mathbf{s}_1^\dagger \mathbf{S}_1 (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-2} \mathbf{S}_1^\dagger \mathbf{s}_1 \quad (50)$$

is positive. In order to proceed, we note that  $\mathbf{s}_1^\dagger \mathbf{s}_1 \geq \mathbf{s}_1^\dagger \mathbf{S}_1 (\mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} \mathbf{S}_1^\dagger \mathbf{s}_1$  holds. This can be seen by noting that the matrix  $\mathbf{S}_1 (\mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} \mathbf{S}_1^\dagger$  is idempotent, and therefore its eigenvalues are either zero or one. Hence,  $\mathbf{I} - \mathbf{S}_1 (\mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} \mathbf{S}_1^\dagger$  is

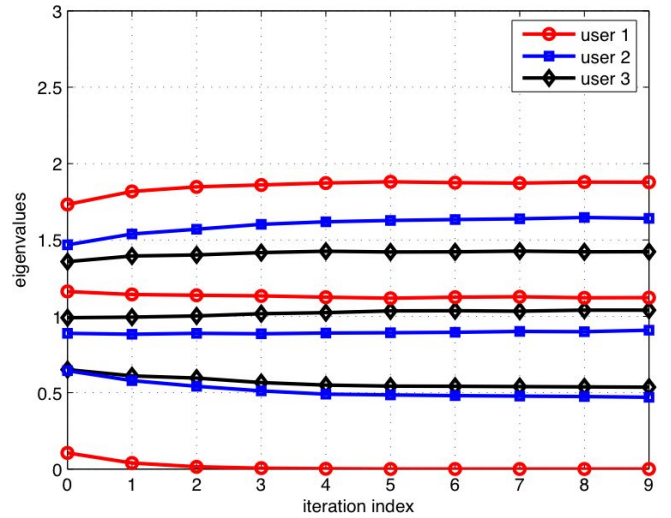


Fig. 6. The convergence of the multi-user algorithm where each iteration corresponds to a single update of all users.

positive definite. Using this inequality, the condition becomes,

$$\mathbf{s}_1^\dagger \mathbf{S}_1 \left[ (\mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} - (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} - (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-2} \right] \mathbf{S}_1^\dagger \mathbf{s}_1 \geq 0 \quad (51)$$

Now, let us look at the term between the square brackets,

$$(\mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} - (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} - (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-2} = \quad (52)$$

$$= \left[ (\mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1) - \mathbf{I} \right] (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} - (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-2} \quad (53)$$

$$= (\mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} - (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-2} \quad (54)$$

$$= \left( (\mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} - (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} \right) (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} \quad (55)$$

$$= (\mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} \quad (56)$$

$$= (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} (\mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} (\mathbf{I} + \mathbf{S}_1^\dagger \mathbf{S}_1)^{-1} \quad (57)$$

Now, let the singular value decomposition of  $\mathbf{S}_1$  be  $\mathbf{S}_1 = \mathbf{U} \mathbf{D} \mathbf{V}^\dagger$ , then  $\mathbf{S}_1^\dagger \mathbf{S}_1 = \mathbf{V} \mathbf{D}^2 \mathbf{V}^\dagger$ . Inserting this into (57), and (57) into (51), we get

$$\mathbf{s}_1^\dagger \mathbf{U} \mathbf{D} (\mathbf{I} + \mathbf{D}^2)^{-1} \mathbf{D}^{-2} (\mathbf{I} + \mathbf{D}^2)^{-1} \mathbf{D} \mathbf{U}^\dagger \mathbf{s}_1 \geq 0 \quad (58)$$

$$\mathbf{s}_1^\dagger \mathbf{U} (\mathbf{I} + \mathbf{D}^2)^{-1} (\mathbf{I} + \mathbf{D}^2)^{-1} \mathbf{U}^\dagger \mathbf{s}_1 \geq 0 \quad (59)$$

Finally, since  $(\mathbf{I} + \mathbf{D}^2)^{-2}$  is positive definite, (59) holds and (45) is greater than zero.

### B. Proof of Lemma 3

For arbitrary number of antennas, we will assume that we are at some artificial fixed point, which is not a solution of the original optimization problem, with possibly more than one zero components. Let this artificial fixed point be  $\mathbf{p} = (a_1, a_2, \dots, a_{n_T})$ , and let  $S$  be the index set of the zero components so that  $a_j = 0$  for all  $j \in S$ . Since  $\mathbf{p}$  is a fixed point, the following equalities hold for  $i \notin S$

$$a_i = \frac{a_i E_i(\mathbf{p})}{\sum_{j \notin S} a_j E_j(\mathbf{p})} P \quad (60)$$

From above, we find that  $\sum_{j \notin S} a_j E_j(\mathbf{p}) = P E_i(\mathbf{p}) = P \mu'$ , for all  $i \notin S$ . This is equivalent to saying that the KKT conditions of the reduced optimization problem corresponding to components,  $i \notin S$ , are satisfied with equality, where  $\mu'$  is

possibly different than  $\mu$ . We will show that some conditions on  $E_j(\mathbf{p})$ ,  $j \in S$  cannot hold. The case where  $E_j(\mathbf{p}) = \mu'$  for all  $j \in S$  cannot hold, because this would mean that the KKT conditions of the original optimization problem are all satisfied with equality with  $\mu' = \mu$ , and this can only happen when optimal  $\lambda_i^Q$ 's for all  $i$  are non-zero. Now, let  $k$  be the smallest index in  $S$ , then because of the ordering of the eigenvalues of the channel covariance matrix,  $E_k(\mathbf{p})$  is greater than all  $E_j(\mathbf{p})$ , for all  $j \neq k$ , and  $j \in S$ . The case where  $E_k(\mathbf{p}) = \mu'$  and  $E_j(\mathbf{p}) \leq \mu'$ , for all  $j \neq k$ , and  $j \in S$  cannot hold, because that would mean that the KKT conditions of the reduced optimization problem is violated. The case where  $E_k(\mathbf{p}) < \mu'$  and  $E_j(\mathbf{p}) < E_k(\mathbf{p}) < \mu'$ , for all  $j \neq k$ , and  $j \in S$  cannot hold, because that would mean that we satisfy all KKT conditions of the original optimization problem with  $\mu' = \mu$ . This contradicts with our assumption that we are at an artificial fixed point that is not the solution of the original optimization problem. Therefore, in all other possibilities, we have at least  $E_k(\mathbf{p}) > \mu'$ , where  $k$  is the smallest index in  $S$ . Now, we will show that by perturbing the artificial fixed point by an  $\epsilon$  amount, we move further away from that artificial fixed point. We run the algorithm for  $\mathbf{p}'$  which is different from  $\mathbf{p}$  in two components: the  $k^{\text{th}}$  component is  $\epsilon$ , and any  $i^{\text{th}}$  component, for  $i \notin S$ , is  $a_i - \epsilon$ . By using the same Taylor series arguments, we can say that  $E_i(\mathbf{p}') = E_i(\mathbf{p}) + O(\epsilon)$ , for  $i = 1, \dots, n_T$ . If we insert these into  $f_k(\mathbf{p}')$ , we have

$$f_k(\mathbf{p}') = \frac{\epsilon E_k(\mathbf{p})}{\sum_{i \notin S} a_i E_i(\mathbf{p})} P + O(\epsilon^2) \quad (61)$$

We know from (60) that  $\sum_{i \notin S} a_i E_i(\mathbf{p}) = P\mu'$ . Inserting this into the above equation, we have

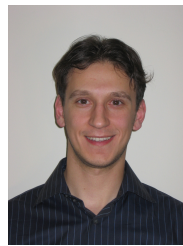
$$f_k(\mathbf{p}') = \epsilon \frac{E_k(\mathbf{p})}{\mu'} + O(\epsilon^2) \quad (62)$$

$$> \epsilon \quad (63)$$

where the last inequality follows from the fact that  $E_k(\mathbf{p}) > \mu'$ . This result tells us that starting from  $\epsilon$  away from an artificial fixed point, the  $k^{\text{th}}$  component of the updated vector, and therefore the updated vector itself moves further away from the artificial fixed point. By using Lemma 1, the algorithm will move away from the artificial fixed point at each iteration. Therefore, this artificial fixed point is unstable.

## REFERENCES

- [1] A. Soysal and S. Ulukus. Optimum power allocation in fading MIMO multiple access channels with partial CSI at the transmitters. In *Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, October 2006.
- [2] İ. E. Telatar. Capacity of multi-antenna Gaussian channels. *European Transactions on Telecommunication*, 10(6):585–596, November 1999.
- [3] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi. Iterative water-filling for Gaussian vector multiple access channels. *IEEE Transactions on Information Theory*, 50(1):145–151, January 2004.
- [4] W. Yu, W. Rhee, and J.M. Cioffi. Optimal power control in multiple access fading channels with multiple antennas. In *IEEE International Conference on Communications*, pages 575–579, 2001.
- [5] E. Visotsky and U. Madhow. Space-time transmit precoding with imperfect feedback. *IEEE Transactions on Information Theory*, 47(6):2632–2639, September 2001.
- [6] S. A. Jafar and A. Goldsmith. Transmitter optimization and optimality of beamforming for multiple antenna systems. *IEEE Transactions on Wireless Communications*, 3(4):1165–1175, July 2004.
- [7] 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.
- [8] A. Soysal and S. Ulukus. Asymptotic optimality of beamforming in multi-user MIMO-MAC with no or partial CSI at the transmitters. In *Vehicular Technology Conference*, Stockholm, Sweden, May 2005.
- [9] A. Soysal and S. Ulukus. Transmit directions and optimality of beamforming in MIMO-MAC with partial CSI at the transmitters. In *Conference on Information Sciences and Systems*, Baltimore, MD, March 2005.
- [10] A. Soysal and S. Ulukus. Optimality of beamforming in fading MIMO multiple access channel. Submitted for journal publication.
- [11] E. Jorswieck and H. Boche. Optimal transmission strategies and impact of correlation in multi-antenna systems with different types of channel state information. *IEEE Transactions on Signal Processing*, 52(12):3440–3453, December 2004.
- [12] A. M. Tulino, A. Lozano, and S. Verdú. Power allocation in multi-antenna communication with statistical channel information at the transmitter. In *IEEE Int. Symp. on PIMRC*, Barcelona, Spain, September 2004.
- [13] A. M. Tulino, A. Lozano, and S. Verdú. Capacity-achieving input covariance for single-user multi-antenna channels. *IEEE Transactions on Wireless Communications*, 5(3):662–671, March 2006.
- [14] C. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela. Capacity scaling in MIMO wireless systems under correlated fading. *IEEE Transactions on Information Theory*, 48(3):637–650, March 2002.
- [15] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.
- [16] S. C. Malik and S. Arora. *Mathematical Analysis*. Wiley, 1992.
- [17] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath. Capacity limits of MIMO channels. *IEEE Journal on Selected Areas in Communications*, 21(5), June 2003.
- [18] D. P. Bertsekas and J. N. Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods*. Athena Scientific, 1997.
- [19] E. Jorswieck and H. Boche. Channel capacity and capacity-range of beamforming in MIMO wireless systems under correlated fading with covariance feedback. *IEEE Transactions on Wireless Communications*, 3(5):1543–1553, September 2004.



**Alkan Soysal** (S'03) received the B.S. degree in electrical and electronics engineering from Middle East Technical University, Ankara, Turkey in 2002, and the M.S. degree in electrical and computer engineering from University of Maryland, College Park, MD in 2006. He is currently working toward the Ph.D. degree in the Department of Electrical and Computer Engineering, University of Maryland, College Park. His research interests include wireless communication theory and information theory with particular focus on MIMO networks.



**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 appointment 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 as an Associate Editor for the *IEEE Transactions on Communications* since 2003, as a Guest Editor for the *IEEE Journal on Selected Areas in Communications* in 2007, as the co-chair of the Communication Theory Symposium at the 2007 IEEE Global Telecommunications Conference and as the Secretary of the IEEE Communication Theory Technical Committee (CTTC).