Power Region for Fading Multiple-Access Channel with Multiple Antennas

Rui Zhang, Mehdi Mohseni, John M. Cioffi
Department of Electrical Engineering, Stanford University, Stanford, CA 94305
Email: {ee.ruizhang, mmohseni, cioffi}@stanford.edu

***Abstract*— We consider the fading multiple-access channel (MAC) with additive Gaussian noise and multiple transmit and receive antennas. It is assumed that the receiver has perfect channel state information (CSI) while the mobile transmitters have no such knowledge. We study the transmission scheme under which the receiver determines the transmit covariance matrices for all the transmitters, based on the long-term CSI statistics, and then feeds them back to each transmitter. We characterize the achievable multi-user *power region* under this scheme. The power region constitutes all transmit power-tuples under which a given set of "ergodic" rates is achievable for the transmitters. We show that all the boundary points of the power region can be obtained through solving a sequence of weighted sum-power minimization problems. We observe through numerical results that the transmit optimization based on the long-term CSI statistics can provide substantial power savings in the fading MAC.**

I. Introduction

Transmission through multiple transmit and receive antennas, or the *so-called* multiple-input multiple-output (MIMO) channel, is known as an efficient scheme in providing enormous information rates in rich-scattering mobile environment [1]. The characterization of MIMO channel capacity limits under various assumptions on transmitter and receiver channel state information (CSI) or channel distribution information (CDI), has motivated a great deal of valuable scholarly work (see [2] and references therein). For a single user case and when CSI is perfectly known at the receiver but only CDI (in the form of either channel mean information or channel covariance information) is known at the transmitter, solutions to transmit signal covariance matrix for ergodic capacity optimization as well as the conditions under which beamforming is optimal, have been recently established in several works (see [2] and references therein).

This paper considers the multi-user transmission through fading MIMO multiple-access channel (MIMO-MAC) with additive Gaussian noise at the receiver. Two commonly adopted means to measure the information-theoretic limits of the Gaussian MAC are the multi-user *capacity region* and *power region*. The capacity region is defined as the constitution of all *achievable*[1] rate-tuples for the users given their individual power constraints. On the other hand, the power region consists of all possible power-tuples for the users under which a given rate-tuple is achievable. The capacity region of a *deterministic* Gaussian MAC with single transmit and receive antenna (SISO-MAC) has the well-known *polymatroid* structure [4], which also holds for the fading MIMO-MAC. On the other hand, though the power region of a deterministic SISO-MAC assumes the *contra-polymatroid* structure [4], this structure is non-existent for the fading MAC or MIMO-MAC. As a result, the characterization of the power region for the fading MIMO-MAC is not yet fully understood.

In [5], the authors have explored the characterization of the power region for the fading MIMO-MAC when CSI is perfectly known at both the transmitters and the receiver. In this paper, we use similar techniques as in [5] to characterize the power region for the fading MIMO-MAC when CSI is perfectly known at the receiver but unknown at each transmitter. We consider the transmission scheme under which the receiver determines the transmit covariance matrices for all the transmitters, based on the long-term CSI statistics (or equivalently, the CDI) as well as the rate requirement and power budget of each transmitter. Under the assumption that there is a reliable link from the receiver to each transmitter, the receiver feeds back the transmit signal covariance matrices to all transmitters. The main goal of this paper is to characterize the achievable power region under this scheme. Similar system setup as in this paper has been considered in [6] in characterizing the capacity region for the fading MIMO-MAC.

Notations: We use boldface letters to denote matrices and vectors. $|\boldsymbol{S}|$ denotes the determinant, $\boldsymbol{S}^{-1}$ the inverse and $\mathtt{Tr}(\boldsymbol{S})$ the trace of a square matrix $\boldsymbol{S}$. For any general matrix $\boldsymbol{M}$, $\boldsymbol{M}^{\dagger}$ denote its conjugate transpose. $\boldsymbol{I}$ denotes the identity matrix. $\mathbb{E}[\cdot]$ denotes statistical expectation. $\mathbb{C}^{x\times y}$ denotes the space of $x\times y$ matrices with complex entries. $\mathbb{R}^{M}$ denotes the M-dimensional real Euclidean space and $\mathbb{R}_{+}^{M}$ its non-negative orthant. The distribution of a complex Gaussian vector with the mean vector $\boldsymbol{x}$ and the covariance matrix $\boldsymbol{\Sigma}$ is denoted by $\mathcal{CN}(\boldsymbol{x},\boldsymbol{\Sigma})$ and $\sim$ means "distributed as." The sign $\succeq$ denotes the generalized inequality [7] and for a square matrix $\boldsymbol{S}$, $\boldsymbol{S}\succeq 0$ means that $\boldsymbol{S}$ is positive semi-definite.

II. System Model

We consider a fading MIMO-MAC with r receive antennas at the base station and K mobile users equipped with $t_1,\cdots,t_K$ transmit antennas, respectively. The transmission from each user to the base station is assumed to be synchronously. We assume the space of fading states is continuous

[1] A rate-tuple is achievable if there exist a sequence of codes for each of the users such that the probability of decoding error for all the users approaches zero as the code length goes to infinity [3].

and infinite and the fading process is stationary and ergodic. At each state ν, the fading MAC is given by

$$\boldsymbol{y}(\nu) = [\boldsymbol{H}_1(\nu) \cdots \boldsymbol{H}_K(\nu)] \begin{bmatrix} \boldsymbol{x}_1(\nu) \\ \vdots \\ \boldsymbol{x}_K(\nu) \end{bmatrix} + \boldsymbol{z}(\nu), \quad (1)$$

where $\boldsymbol{y} \in \mathbb{C}^{r\times 1}$ denotes the received signal vector. $\boldsymbol{x}_k \in \mathbb{C}^{t_k\times 1}$ and $\boldsymbol{H}_k \in \mathbb{C}^{r\times t_k}$ denote, respectively, the transmitted signal vector and the channel matrix of mobile user k, $k = 1, \ldots, K$. $\boldsymbol{z} \in \mathbb{C}^{r\times 1}$ denotes the AWGN at the receiver where $\boldsymbol{z} \sim \mathcal{CN}(0, \boldsymbol{I})$.

We consider the case where the CSI is perfectly known at the receiver but unknown at each transmitter. With availability of CSI, the receiver can acquire the long-term CSI statistics of each mobile user and hence determine the transmit signal covariance matrices for all mobile users based on the individual rate requirements and power budgets. Let the covariance matrix of the transmitted signal of user k be $\boldsymbol{S}_k \triangleq \mathbb{E}[\boldsymbol{x}_k\boldsymbol{x}_k^\dagger]$, where the expectation is taken over the code-book and $\boldsymbol{S}_k \succeq 0$. Recall that since CSI is not available at the transmitter, $\boldsymbol{S}_k$ is fixed for all the fading states. Any code-book for user k should satisfy $\mathtt{Tr}(\boldsymbol{S}_k) = p_k$, where $\boldsymbol{p} = (p_1, \ldots, p_K) \in \mathbb{R}_+^K$ is the vector of average transmit powers for all the users. For a fixed set of $\{\boldsymbol{S}_k\}$, $k = 1, \ldots, K$, it was shown in [8] that all the rate-tuples in the set $\mathcal{C}_f(\{\boldsymbol{H}_k(\nu)\}, \{\boldsymbol{S}_k\})$ defined as below are achievable.

$$\mathcal{C}_f(\{\boldsymbol{H}_k(\nu)\}, \{\boldsymbol{S}_k\}) = \Bigg\{\boldsymbol{r} \in \mathbb{R}_+^K : \sum_{k\in J} r_k \le \mathbb{E}\left[\frac{1}{2}\log\left|\sum_{k\in J}\boldsymbol{H}_k(\nu)\boldsymbol{S}_k\boldsymbol{H}_k^\dagger(\nu) + \boldsymbol{I}\right|\right] \;\; \forall J \subseteq \{1, \ldots, K\}\Bigg\},$$

where the expectation is taken over the distribution of $\boldsymbol{\nu}$.

The *power region* for the fading MAC is then defined as

$$\mathcal{P}_{\mathrm{MAC}}(\boldsymbol{R}) \triangleq \{\boldsymbol{p} : \; \boldsymbol{r} \succeq \boldsymbol{R}, \; \boldsymbol{r} \in \mathcal{C}_f(\{\boldsymbol{H}_k(\nu)\}, \{\boldsymbol{S}_k\})\}, \quad (2)$$

where $\boldsymbol{R} = (R_1, \cdots, R_K) \in \mathbb{R}_+^K$ denotes the average-rate constraint vector. It is not hard to show that the power region defined in (2) is a convex set.

III. Characterization of Power-Region

The power region defined in (2) is illustrated in Fig. 1 for a two-user MAC. The solid line in Fig. 1 represents all the boundary points of the power region. Each boundary point, *e.g.*, point A as indicated in Fig. 1, can be characterized by two means. Firstly, due to the convexity of the power region, each of its boundary points can be expresses as the solution to a weighted sum-power minimization problem for some non-negative weights denoted by λ_ks,

$$\begin{aligned} \mathtt{Minimize} \quad & \sum_{k=1}^K \lambda_k \mathtt{Tr}(\boldsymbol{S}_k) && (3) \\ \mathtt{s.t.} \quad & r_k \ge R_k \;\; \forall k, && (4) \\ & \boldsymbol{r} \in \mathcal{C}_f(\{\boldsymbol{H}_k(\nu)\}, \{\boldsymbol{S}_k\}), && (5) \\ & \boldsymbol{S}_k \succeq 0 \quad \forall k. && (6) \end{aligned}$$

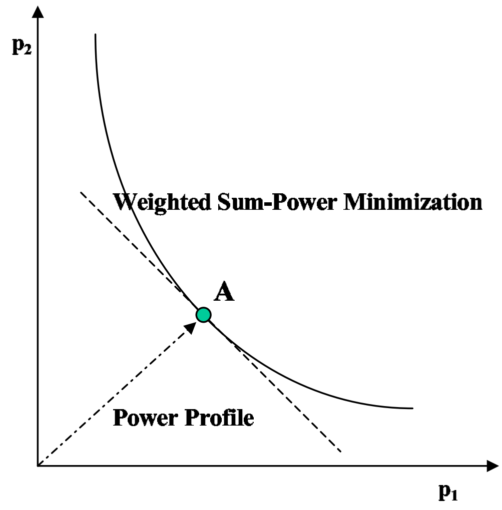

Fig. 1. A two-user MAC power region.

Alternatively, each boundary point of the power region can be considered geometrically as the intersection of a line passing through the origin (*i.e.*, the power-tuple with all zeros) and the power region (see Fig. 1). Equivalently, it can be cast as the solution to the following sum-power minimization problem with a sequence of power-ratio constraints:

$$\begin{aligned} \mathtt{Minimize} \quad & \sum_{k=1}^K \mathtt{Tr}(\boldsymbol{S}_k) && (7) \\ \mathtt{s.t.} \quad & r_k \ge R_k \;\; \forall k, && (8) \\ & \boldsymbol{r} \in \mathcal{C}_f(\{\boldsymbol{H}_k(\nu)\}, \{\boldsymbol{S}_k\}), && (9) \\ & \boldsymbol{S}_k \succeq 0 \quad \forall k, && (10) \\ & \mathtt{Tr}(\boldsymbol{S}_k) = \alpha_k \sum_{j=1}^K \mathtt{Tr}(\boldsymbol{S}_j) \;\; \forall k. && (11) \end{aligned}$$

In the above, $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_K) \in \mathbb{R}_+^K$ is the prescribed *power-profile* vector and $\sum_{k=1}^K \alpha_k = 1$. It is not hard to show that the above two problems are both convex. Therefore, they can be solved by standard convex optimization techniques.

A. Weighted Sum-Power Minimization

We first consider the weighted sum-power minimization problem in (3). The Lagrangian of this primal problem with the vector of dual variables $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_K) \in \mathbb{R}_+^K$, associated with the inequality constraints in (4) is defined as in [7] and is given by

$$\mathcal{L}(\{\boldsymbol{S}_k\}, \{r_k\}, \boldsymbol{\mu}) = \sum_{k=1}^K \lambda_k \mathtt{Tr}(\boldsymbol{S}_k) - \sum_{k=1}^K \mu_k(r_k - R_k). \quad (12)$$

Note that the variables $\{\boldsymbol{S}_k\}$ and $\{r_k\}$ belong to the region $\mathcal{D}$ which is specified by the remaining constrains in (5) and (6). Then the Lagrange dual function [7] defined as

$$g(\boldsymbol{\mu}) = \min_{\{\boldsymbol{S}_k\}, \{r_k\} \in \mathcal{D}} \mathcal{L}(\{\boldsymbol{S}_k\}, \{r_k\}, \boldsymbol{\mu}), \quad (13)$$

serves as a lower bound on the optimal value of the primal problem, denoted by p^*. For a convex problem, Slater's condition states that the duality gap is zero if the feasible set has non-empty interior [7]. By using large enough powers, the sets $\mathcal{C}_f(\{\boldsymbol{H}_k(\nu)\}, \{\boldsymbol{S}_k\})$ can be made arbitrary large to contain any rate $\boldsymbol{R}$ as an interior point. Thus Slater's condition holds and the duality gap is zero for our problem, *i.e.*,

$$p^* = \max_{\boldsymbol{\mu} \succeq 0} g(\boldsymbol{\mu}) \triangleq g(\boldsymbol{\mu}^*), \tag{14}$$

where $\boldsymbol{\mu}^*$ denote an optimal solution in maximizing the dual function. The above equality suggests that the optimal solution to the primal problem can be obtained by first minimizing the Lagrangian, $\mathcal{L}$, to obtain the Lagrange dual function $g(\boldsymbol{\mu})$ for a given $\boldsymbol{\mu}$, and then maximizing $g(\boldsymbol{\mu})$ over all possible values of μ_ks. In the following, we provide a numerical routine to solve this optimization problem.

Numerical Routine: We first look further into the minimization of the Lagrangian, $\mathcal{L}$, to attain the dual function $g(\boldsymbol{\mu})$ in (13). In this case, $\boldsymbol{\mu}$ is fixed and the variables are $\{\boldsymbol{S}_k\}$ and $\{r_k\}$. From (12), it is easy to see that the minimization of $\mathcal{L}$ can be rewritten as the following equivalent problem:

$$\begin{aligned} \texttt{Maximize} \quad & \sum_{k=1}^{K} \mu_k r_k - \sum_{k=1}^{K} \lambda_k \texttt{Tr}(\boldsymbol{S}_k) && (15) \\ \texttt{s.t.} \quad & \boldsymbol{r} \in \mathcal{C}_f(\{\boldsymbol{H}_k(\nu)\}, \{\boldsymbol{S}_k\}), && (16) \\ & \boldsymbol{S}_k \succeq 0 \quad \forall k. && (17) \end{aligned}$$

To simplify the above problem further, the following lemma borrowed from [4] is utilized to remove the constraint in (16).

Lemma 1: For any permutation π over $\{1, 2, \ldots, K\}$, $\boldsymbol{r}^{(\pi)}$ defined as

$$r_{\pi(k)}^{(\pi)} = \mathbb{E}\left[\frac{1}{2}\log\frac{\left|\sum_{i=1}^{k}\boldsymbol{H}_{\pi(i)}(\nu)\boldsymbol{S}_{\pi(i)}^*\boldsymbol{H}_{\pi(i)}^\dagger(\nu) + \boldsymbol{I}\right|}{\left|\sum_{i=1}^{k-1}\boldsymbol{H}_{\pi(i)}(\nu)\boldsymbol{S}_{\pi(i)}^*\boldsymbol{H}_{\pi(i)}^\dagger(\nu) + \boldsymbol{I}\right|}\right], \tag{18}$$

is a *vertex* of the polymatroid $\mathcal{C}_f(\{\boldsymbol{H}_k(\nu)\}, \{\boldsymbol{S}_k\})$ in $\mathbb{R}_+^K$. Furthermore, for any $\boldsymbol{\mu} \succeq 0$, the solution to the problem

$$\texttt{Maximize} \quad \sum_{k=1}^{K} \mu_k r_k \quad \texttt{s.t.} \quad \boldsymbol{r} \in \mathcal{C}_f(\{\boldsymbol{H}_k(\nu)\}, \{\boldsymbol{S}_k\})$$

is attained by a vertex $\boldsymbol{r}^{(\pi^*)}$, where π^* is such that $\mu_{\pi^*(1)} \geq \mu_{\pi^*(2)} \geq \ldots \geq \mu_{\pi^*(K)}$.

With the above lemma, the problem in (15) can be further simplified as the maximization of

$$\sum_{k=1}^{K}\left(\mu_{\pi(k)} - \mu_{\pi(k+1)}\right)\mathbb{E}\left[\frac{1}{2}\log\left|\sum_{i=1}^{k}\Big(\boldsymbol{H}_{\pi(i)}(\nu)\right.\right.$$
$$\left.\left.\boldsymbol{S}_{\pi(i)}\boldsymbol{H}_{\pi(i)}^\dagger(\nu)\Big) + \boldsymbol{I}\right|\right] - \sum_{k=1}^{K}\lambda_k \texttt{Tr}(\boldsymbol{S}_k), \tag{19}$$

where $\boldsymbol{S}_k \succeq 0$ for $k = 1, \ldots, K$ and π is a permutation such that $\mu_{\pi(1)} \geq \mu_{\pi(2)} \geq \ldots \geq \mu_{\pi(K)} \geq \mu_{\pi(K+1)} = 0$. Since the above maximization is concave with twice differentiable objective function and only positive semi-definite constraints, it can be solved numerically, *e.g.*, by interior-point method [7].

Next, we consider the maximization of the dual function, $g(\boldsymbol{\mu})$ over all possible values of μ_ks. Although $g(\boldsymbol{\mu})$ is concave, it is not necessarily differentiable and therefore optimization algorithms that exploit the function's differentials such as Newton method cannot be employed directly. Nevertheless, performing a "gradient" based search to find the optimal values of μ_ks is still possible. The search direction is based on what is known as a "sub-gradient". For the concave function $g(\boldsymbol{\mu})$, $\boldsymbol{\theta} \in \mathbb{R}^K$ is called a sub-gradient at $\boldsymbol{\mu}$ if $g(\boldsymbol{\phi}) \leq g(\boldsymbol{\mu}) + (\boldsymbol{\phi} - \boldsymbol{\mu})^T\boldsymbol{\theta}$ for all $\boldsymbol{\phi} \in \mathbb{R}_+^K$. Thus the optimal values of $\boldsymbol{\mu}^*$, cannot lie in the half-space $(\boldsymbol{\phi} - \boldsymbol{\mu})^T\boldsymbol{\theta} < 0$. The following lemma provides us with a sub-gradient for $g(\boldsymbol{\mu})$.

Lemma 2: If $\{\boldsymbol{S}_k^*\}$ and $\{r_k^*\}$ minimize $\mathcal{L}(\{\boldsymbol{S}_k\}, \{r_k\}, \boldsymbol{\mu})$ over $\mathcal{D}$, *i.e.* $\mathcal{L}(\{\boldsymbol{S}_k^*\}, \{r_k^*\}, \boldsymbol{\mu}) = g(\boldsymbol{\mu})$, then the vector $\boldsymbol{\theta}$ defined as $\theta_k = R_k - r_k^*$ is a sub-gradient at $\boldsymbol{\mu}$.

Proof : For any $\boldsymbol{\phi} \succeq 0$, $g(\boldsymbol{\phi}) \leq \mathcal{L}(\{\boldsymbol{S}_k^*\}, \{r_k^*\}, \boldsymbol{\phi}) = g(\boldsymbol{\mu}) + \sum_{k=1}^{K}(\phi_k - \mu_k)(R_k - r_k^*)$.

With the sub-gradient at hand, the optimal dual variable $\boldsymbol{\mu}^*$ can be found by a gradient-based search method like sub-gradient method or ellipsoid method [7]. We omit here their details that can be found in standard optimization texts like [7]. We only point out that from programming implementation, the ellipsoid method is more suitable than the sub-gradient method in solving the problem due to its superior convergence behavior and guaranteed convergence. It is also noted that the sub-gradient method may exhibit oscillation when some of the μ_k^*s are indeed equal. It is also remarked that the algorithm proposed in [4, Algorithm 5.3] can be modified to search for the optimal value of $\boldsymbol{\mu}^*$, which, like sub-gradient method, may also exhibit oscillation when some of μ_k^*s are equal.

So far we have developed numerical routines to determine the optimal value for the problem in (3), $p^* = \sum_{k=1}^{K}\texttt{Tr}(\boldsymbol{S}_k^*)$, and associated primal and dual optimal solutions, $\{\boldsymbol{S}_k^*, r_k^*\}$ and $\boldsymbol{\mu}^*$, respectively. In the following, we address the issue on the uniqueness of the solutions. While uniqueness of the dual optimal solution $\boldsymbol{\mu}^*$ is not an issue in convergence of the proposed algorithms, uniqueness of the primal variables plays an important role in their convergence behavior. Since we are using a primal-dual approach to solve the optimization problem in (3), the primal variables $\{\boldsymbol{S}_k^*, r_k^*\}$ that optimize the Lagrangian for $\boldsymbol{\mu}^*$ might not satisfy the rate constraints in (4). However if we prove that these solutions are unique, from the Karush-Kuhn-Tucker (KKT) optimality conditions, they satisfy the rate constraints automatically.

Uniqueness of the Solutions: First we study the effect of $\boldsymbol{\mu}^*$ on the convergence of our algorithms. Consider a fading SIMO (single transmit antenna and multiple receive antennas) MAC with two users. The channel of the two users, $\{\boldsymbol{H}_1(\nu)\}$ and $\{\boldsymbol{H}_2(\nu)\}$, are independent and each has independent entries distributed as $\mathcal{CN}(0, 1)$. Fig. 2 shows the capacity region of this fading SIMO-MAC under a sum-power constraint, *i.e.*, $p_1 + p_2 \leq 10$. It is seen that the capacity region is indeed symmetric. The dashed line and the dotted line show how two vertices of the constituting polymatroids sweep the capacity region as p_1 and p_2 vary while keeping their sum equal to 10. The dashed line corresponds to the case when user 2's message is decoded before user 1's while the dotted line does for the reversed decoding order. Any point, *e.g.*, point A shown in Fig. 2, on the dashed or dotted line corresponds to a vertex of some polymatroid. Also note that there is a region on the boundary that does not consist of the vertices. We refer to this part as the "time-sharing" region which corresponds to -45 degree straight lines of constituting polymatroids (point B is in this region). Hence, any point in the time-sharing region is not achievable by successive decoding and time-sharing of transmission rates and powers as well as decoding orders between two users is required to achieve this point if successive decoding is used at the receiver.

Suppose now the rate-pair indicated by point A is the actual rate constraint for the problem in (3) and the problem minimizes the weighted sum-power of the two users with equal weights, *i.e.*, $\lambda_1 = \lambda_2 = 1$. In this case, the minimum sum-power required is thus 10. Because point A is a vertex, Lemma 1 implies that $\mu_1^* \leq \mu_2^*$ in the associated optimal dual variables, $\boldsymbol{\mu}^*$. On the other hand, if the rate-pair indicated by point B is the rate constraint for a sum-power minimization problem in (3), $\boldsymbol{\mu}^*$ must have equal entries, *i.e.*, $\mu_1^* = \mu_2^*$. Since there is no "sharp" vertex with multiple tangent lines on the boundary of this capacity region under the sum-power constraint, we can conclude that $\boldsymbol{\mu}^*$ is unique for any rate pair on the boundary.

Furthermore, in the Appendix, it is shown that the set of optimal transmit covariance matrices, $\{\boldsymbol{S}_k^*\}$ is indeed unique. The only question remains to be answered is whether the set of optimal rates, r_k^* is unique. The answer is yes for the case where all μ_k^*s are different and positive. Recall that $\boldsymbol{r}^*$ maximizes $\sum_k \mu_k^* r_k$ over $\mathcal{C}_f(\{\boldsymbol{H}_k(\nu)\}, \{\boldsymbol{S}_k^*\})$ and since $\{\boldsymbol{S}_k^*\}$ is unique and all weights are different and positive , $\boldsymbol{r}^*$ is unique. In this case, from KKT conditions, r_k^*s automatically satisfy the rate constraints in (4). Note that for the case where some of the μ_k^*s are equal, the optimal rates, r_k^*s may not be unique and consequently they may not satisfy the rate constraints. As an example, assume the target rate is point B in Fig. 2. In this case, $\mu_1^* = \mu_2^*$ and any point on the straight line maximizes $\sum_k \mu_k^* r_k$ over $\mathcal{C}_f(\{\boldsymbol{H}_k(\nu)\}, \{\boldsymbol{S}_k^*\})$. Moreover, our proposed numerical routine obtains either point C or D as the optimal $\boldsymbol{r}^*$ which clearly does not satisfy the rate constraints. This is because the simplified optimization in (19) always tends to use a vertex of the capacity region as the solution for $\boldsymbol{r}^*$. However, this is not an issue in the convergence of our algorithm. As far as we know $\{\boldsymbol{S}_k^*\}$ is unique, we can obtain the target rate as a convex combination of the vertices of the unique polymatroid, $\mathcal{C}_f(\{\boldsymbol{H}_k(\nu)\}, \{\boldsymbol{S}_k^*\}\}$.

B. Sum-Power Minimization with Power-Profile

In this subsection, we propose a solution to the problem in (7), for which a new set of equality constraints is imposed in (11) to regulate the transmitted power of each user according to a given power profile vector, $\boldsymbol{\alpha}$. Similarly as done in Section III-A, we consider the Lagrangian of the primal problem in (7). Let the vector of dual variables $\boldsymbol{\mu} = (\mu_1, \dots, \mu_K) \in \mathbb{R}_+^K$ be associated with the inequality constraints in (8) and $\boldsymbol{\delta} = (\delta_1, \dots, \delta_K) \in \mathbb{R}^K$ associated with the equality constraints in (11) and $\mathcal{D}$ denotes the set specified by the remaining constraints in (9) and (10). We have

$$\mathcal{L}(\{\boldsymbol{S}_k\}, \{r_k\}, \boldsymbol{\mu}, \boldsymbol{\delta}) = \sum_{k=1}^{K} \lambda_k \mathtt{Tr}(\boldsymbol{S}_k) - \sum_{k=1}^{K} \mu_k (r_k - R_k)$$
$$- \sum_{k=1}^{K} \delta_k \left(\mathtt{Tr}\left(\boldsymbol{S}_k\right) - \alpha_k \sum_{k'=1}^{K} \mathtt{Tr}(\boldsymbol{S}_{k'}) \right). \tag{20}$$

Define the Lagrange dual function as

$$g(\boldsymbol{\mu}, \boldsymbol{\delta}) = \min_{\{\boldsymbol{S}_k\}, \{r_k\} \in \mathcal{D}} \mathcal{L}(\{\boldsymbol{S}_k\}, \{r_k\}, \boldsymbol{\mu}, \boldsymbol{\delta}). \tag{21}$$

And the optimal value of the primal problem can be then attained as

$$p^* = \max_{\boldsymbol{\mu} \succeq 0, \boldsymbol{\delta}} g(\boldsymbol{\mu}, \boldsymbol{\delta}). \tag{22}$$

Similar numerical routine as in Section III-A can be developed for solving the above problem and the details are omitted here for brevity. It is worth pointing out that the dual function, $g(\boldsymbol{\mu}, \boldsymbol{\delta})$, has also a sub-gradient ρ_k defined as $\rho_k = \alpha \sum_{k'=1}^{K} \mathtt{Tr}(\boldsymbol{S}_{k'}^*) - \mathtt{Tr}(\boldsymbol{S}_k^*)$, with respect to the dual variable δ_k, for $k = 1, \dots, K$, where $\mathcal{L}(\{\boldsymbol{S}_k^*\}, \{r_k^*\}, \boldsymbol{\mu}, \boldsymbol{\delta}) = g(\boldsymbol{\mu}, \boldsymbol{\delta})$.

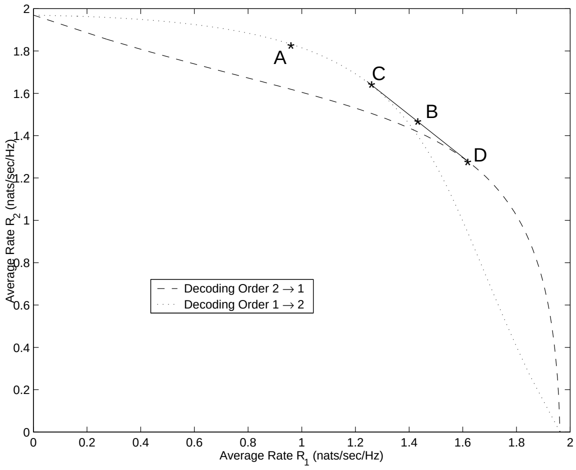

Fig. 2. Capacity region for two-user fading MAC with $t_1 = t_2 = 1$, $r = 2$ and $p_1 + p_2 \leq 10$.

IV. Numerical Results

In this section we characterize the power region for a Rayleigh fading MIMO-MAC with $r = 2$ antennas at the base station and $K = 2$ mobile users each equipped with $t = 2$ antennas. We consider the case where the receive antennas are separated enough that they experience independent fading. However, due to the size limitations, this might not be the case for the mobile users and the fading levels might be correlated across transmit antennas. Given that, we employ the channel model that describes the channel for user k as $\boldsymbol{H}_k = \boldsymbol{H}_w \boldsymbol{R}_{tk}^{1/2}$ for $k = 1, 2$, where $\boldsymbol{R}_{tk}$ is the covariance matrix for the transmit antenna fading levels of user k, and $\boldsymbol{H}_w$ is a r by t matrix, assumed to be independent across two users and each having independent entries distributed as $\mathcal{CN}(0, 1)$.

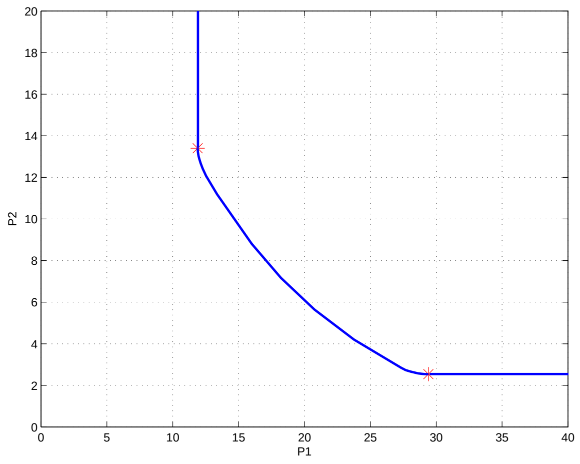

Fig. 3. Power region for two-user Rayleigh fading MAC with $t_1 = t_2 = 2$, $r = 2$.

We assume the receiver has complete CSI knowledge and can exctract $\boldsymbol{R}_{tk}$ from it. Based on $\boldsymbol{R}_{tk}$ it computes the transmit covariance matrix of user k that minimizes the weighted-sum power. Similar to the proof given in [6], it can be shown that the expression in (19) is maximized for $\boldsymbol{S}_k$ that has the same set of eigenvector as $\boldsymbol{R}_{tk}$ for $k = 1, 2$. Equivalently, if $\boldsymbol{Q}_k \boldsymbol{\Lambda}_k \boldsymbol{Q}_k^\dagger$ be the eigenvalue decomposition of $\boldsymbol{R}_{tk}$, then the optimal $\boldsymbol{S}_k$ would be in the form of $\boldsymbol{Q}_k \boldsymbol{\Sigma}_k \boldsymbol{Q}_k^\dagger$ for some diagonal matrix $\boldsymbol{\Sigma}_k$. This observation reduces the number of variables from Kt^2 to Kt that in turn reduces the complexity of our proposed algorithm.

Monte-Carlo simulation with $N = 5000$ samples is used to approximate the expectation terms. Fig. 3 plots the power region of the described MAC for $\mathbf{R} = [2 \quad 1]^T$ nuts per transmission and

$$\boldsymbol{R}_{t1} = \begin{bmatrix} 1 & 0.4 \\ 0.4 & 1 \end{bmatrix} \qquad \boldsymbol{R}_{t2} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}. \quad (23)$$

To illustrate the effectiveness of our proposed algorithm, we compare the weighted-sum power obtained from our algorithm for some weights to the one obtained from another sub-optimal algorithm. In this sub-optimal algorithm, first the receiver picks an arbitrary decoding order for the users. Starting from the last decoded user, it minimizes the power required to maintain the target rate for that user, while considering the users that have not been decoded yet as interference. For example, for decoding order 2 then 1, the receiver first computes $\boldsymbol{S_1}$ with minimum trace that satisfy $\mathbb{E}[\frac{1}{2}\log|\boldsymbol{H}_1\boldsymbol{S}_1\boldsymbol{H}_1^\dagger + \boldsymbol{I}|] \geq 2$, then it fixes $\boldsymbol{S_1}$ and computes the minimum trace $\boldsymbol{S_2}$ that satisfy $\mathbb{E}[\frac{1}{2}\log|\boldsymbol{H}_1\boldsymbol{S}_1\boldsymbol{H}_1^\dagger + \boldsymbol{H}_2\boldsymbol{S}_2\boldsymbol{H}_2^\dagger + \boldsymbol{I}|] \geq 3$. Each of these optimizations are convex. The powers obtained by this sub-optimal algorithm are shown by (*) in Fig. 3 for two different decoding orders. The minimum value of $0.4p_1 + 0.6p_2$ to support the target rate $[2 \quad 1]^T$ is 11.5 by our algorithm. However, the receiver running this sub-optimal algorithm will require 13.3 units of power if user 1 is decoded first and 12.8 units of power if user 2 is decoded first.

V. Conclusion

Power region for a fading MIMO-MAC is considered and efficient numerical algorithms for characterizing this region is proposed.

Appendix

In the Appendix, we prove that the optimal values of the transmit signal covariance matrices, $\{\boldsymbol{S}_k^*\}$, are unique. Without loss of generality assume $\mu_1^* \geq \mu_2^* \geq \ldots \geq \mu_K^* > \mu_{K+1}^* = 0$. Let $\{\boldsymbol{S}_k^{(1)}\}$ and $\{\boldsymbol{S}_k^{(2)}\}$ be two optimal solutions of the problem in (3). From (12) and using Lemma 1, it can be shown

$$p^* = \sum_{k=1}^{K} \lambda_k \mathtt{Tr}(\boldsymbol{S}_k^{(j)}) + \sum_{k=1}^{K} \mu_k^* R_k - \sum_{k=1}^{K} \Bigg((\mu_k^* - \mu_{k+1}^*) \\ \mathbb{E}\left[\frac{1}{2}\log\left|\sum_{i=1}^{k} \boldsymbol{H}_i(\nu)\boldsymbol{S}_i^{(j)}\boldsymbol{H}_i(\nu)^\dagger + \boldsymbol{I}\right|\right]\Bigg),$$

for $j = 1, 2$. Since the problem is convex, for any $\beta \in [0, 1]$, $\boldsymbol{S}_k = \beta\boldsymbol{S}_k^{(1)} + \bar{\beta}\boldsymbol{S}_k^{(2)}$ is also an optimal solution where $\bar{\beta} = 1 - \beta$. This fact together with the concavity of $\log|\cdot|$ function implies that

$$\mathbb{E}\left[\log\left|\beta\boldsymbol{A}^{(1)}(\nu) + \bar{\beta}\boldsymbol{A}^{(2)}(\nu)\right| - \\ \beta\log\left|\boldsymbol{A}^{(1)}(\nu)\right| - \bar{\beta}\log\left|\boldsymbol{A}^{(2)}(\nu)\right|\right] = 0,$$

where $\boldsymbol{A}^{(j)}(\nu) \triangleq \sum_{k=1}^{K} \boldsymbol{H}_k(\nu)\boldsymbol{S}_k^{(j)}\boldsymbol{H}_k(\nu)^\dagger + \boldsymbol{I}$. Let $f(\beta)$ denote the function on the LHS of the above equation, then $f(\beta) = 0$, for all $0 \leq \beta \leq 1$. Because $f(\beta)$ is twice continuously differentiable function, both its first and second derivatives must vanish, *i.e.*, $\frac{d^2}{d\beta^2} f(\beta) =$

$$-\mathbb{E}\left[\mathtt{Tr}\left(\left((\boldsymbol{A}^{(1)}(\nu) - \boldsymbol{A}^{(2)}(\nu))(\beta\boldsymbol{A}^{(1)}(\nu) + \bar{\beta}\boldsymbol{A}^{(2)}(\nu))^{-1}\right)^2\right)\right]$$

must be equal to zero. For every ν, the matrix in $\mathtt{Tr}(.)$ is a positive semi-definite matrix and hence has a positive trace. Since the expectation of a positive random variable is zero, it must be zero a.e., or $\boldsymbol{A}^{(1)}(\nu) = \boldsymbol{A}^{(2)}(\nu)$ a.e.. This implies that $\boldsymbol{S}_k^{(1)} = \boldsymbol{S}_k^{(2)}$ for the case of infinite number of fading states.

References

[1] I.E. Telatar, "Capacity of multi-antenna Gaussian channels," *Bell Labs Technical Memorandum*, 1995.

[2] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels", *IEEE J. Select. Areas Commun.*, vol. 21, No. 5, pp. 684-702, June 2003.

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

[4] D. Tse and S. Hanly,"Multi-access fading channels-Part I: polymatroid structure, optimal resouce allocation and throuput capacities" *IEEE Trans. Inform. Theory*, vol.44, no.7, pp.2796-2815, Nov. 1998.

[5] M. Mohseni, R. Zhang and John Cioffi, "Optimized transmission of fading multiple-access and broadcast channels with multiple antennas," *submitted to IEEE J. Select. Areas Commun.*, Sept. 2005. Available at http://www.stanford.edu/ engp9824/Websites/JSAC.pdf

[6] Syed A. Jafar and A. J. Goldsmith, "Vector MAC capacity region with Covariance Feedback", *Proc. Int. Symp. Inform. Theory,*, 2001.

[7] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge Univeristy Press, 2004.

[8] S. Shamai and A.D. Wyner, "Information theoretic considerations for symmetric, cellular, multiple access fading channels-part I," *IEEE Trans. on Information Theory*, vol.43, No.6, pp.1877-1894, Nov. 1997.