Dynamic Rate Allocation in Fading Multiple Access Channels

Ali ParandehGheibi, Atilla Eryilmaz, Asuman Ozdaglar, and Muriel Médard

***Abstract*— We consider the problem of rate allocation in a fading Gaussian multiple-access channel (MAC) with fixed transmission powers. Our goal is to maximize a general concave utility function of transmission rates over the throughput capacity region. In contrast to earlier works in this context that propose solutions where a potentially complex optimization problem must be solved in every decision instant, we propose a low-complexity approximate rate allocation policy and analyze the effect of temporal channel variations on its utility performance. To the best of our knowledge, this is the first work that studies the tracking capabilities of an approximate rate allocation scheme under fading channel conditions.**

We build on an earlier work to present a new rate allocation policy for a fading MAC that implements a low-complexity approximate gradient projection iteration for each channel measurement, and explicitly characterize the effect of the speed of temporal channel variations on the tracking neighborhood of our policy. We further improve our results by proposing an alternative rate allocation policy for which tighter bounds on the size of the tracking neighborhood are derived. These proposed rate allocation policies are computationally efficient in our setting since they implement a single gradient projection iteration per channel measurement and each such iteration relies on approximate projections which has polynomial-complexity in the number of users.

I. Introduction

Dynamic allocation of communication resources such as bandwidth or transmission power is a central issue in multiple access channels in view of the time-varying nature of the channel and interference effects. Most of the existing literature on resource allocation in multiple access channels focuses on specific communication schemes such as TDMA (time-division multiple access) [1] and CDMA (code-division multiple access) [2], [3] systems. An exception is the work by Tse *et al.* [4], who introduced the notion of *throughput capacity* for the fading channel with Channel State Information (CSI) and studied dynamic rate allocation policies with the goal of maximizing a linear utility function of rates over the throughput capacity region.

An important literature relevant to our work appears in the context of cross-layer design, where joint scheduling-routing-flow control algorithms have been proposed and shown to achieve utility maximization for concave utility functions while guaranteeing network stability (e.g. [5], [6], [7], [8]). The common idea behind these schemes is to use properly maintained queues to make dynamic decisions about new packet generation as well as rate allocation.

Some of these works ([6], [7]) explicitly address the fading channel conditions, and show that their policies can achieve rates arbitrarily close to the optimal based on a design parameter choice. However, the rate allocation imposed by these schemes requires that a large optimization problem requiring global information be solved over a complex feasible set in every time slot. Clearly, this may not always be possible due to the limitations of the available information, or the processing power, or the complexity intrinsic to the feasible set. In fact, even in the absence of fading, the interference constraints between nearby nodes' transmissions may make the feasible set so complex that the optimal rate allocation problem becomes NP-hard (see [9]).

In the absence of fading, several works have proposed and analyzed approximate randomized and/or distributed rate

This research was partially supported by the National Science Foundation under grant DMI-0545910, and by DARPA ITMANET program.

A. ParandehGheibi is with the Laboratory for Information and Decision Systems, Electrical Engineering and Computer Science Department, Massachusetts Institute of Technology, Cambridge MA, 02139 (e-mail: parandeh@mit.edu)

A. Eryilmaz is with the Electrical and Computer Engineering, Ohio State University, OH, 43210 (e-mail: eryilmaz@ece.osu.edu)

A. Ozdaglar and M. Médard are with the Laboratory for Information and Decision Systems, Electrical Engineering and Computer Science Department, Massachusetts Institute of Technology, Cambridge MA, 02139 (e-mails: asuman@mit.edu, medard@mit.edu)

allocation algorithms for various interference models ([10], [5], [11], [9], [12], [13]), and their effect on the utility maximization is investigated in [9], [14]. However, no similar work exists for fading channel conditions, where the changes in the fading conditions coupled with the inability to solve the optimization problem instantaneously make the solution much more challenging. In fact, it is not even clear what algorithm can be used to achieve close to close-to-optimal performance.

In this work, we propose an approximate gradient projection method and study its tracking capabilities when the channel conditions vary over time. In our algorithm, the solution is updated in every time slot in a direction to increase the utility function at that time slot. But, since the channel may vary between time-slots, the extend of these temporal channel variations become critical to the performance. We explicitly quantify the impact of the speed of fading on the performance of the policy, both for the worst-case and the average behavior. Our results also capture the effect of the degree of concavity of the utility functions on the average performance.

Other than the papers cited above, our work is also related to the work of Vishwanath *et al.* [15] which builds on [4] and takes a similar approach to the rate and power allocation problem for linear utility functions. Other works address different criteria for resource allocation including minimizing the weighted sum of transmission powers [16], and considering Quality of Service (QoS) constraints [17].

The remainder of this paper is organized as follows: In Section II, we introduce the model and describe the capacity region of a multiple-access channel. In Section III, we consider the utility maximization problem in fading channel, and present a rate allocation policy and characterize the tracking neighborhood in terms of the maximum speed of fading. In Section IV, we provide an alternative rate allocation policy and provide a bound on the size of tracking neighborhood as a function of the average speed of fading. Finally, we give our concluding remarks in Section V.

Regarding the notation, we denote by x_i the i-th component of a vector $\boldsymbol{x}$. We denote the nonnegative orthant by $\mathbb{R}^n_+$, i.e., $\mathbb{R}^n_+ = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x} \geq 0\}$. We write $\boldsymbol{x}'$ to denote the transpose of a vector $\boldsymbol{x}$. The exact projection operation on a convex set is denoted by $\mathcal{P}$.

II. System Model

We consider M users sharing the same media to communicate to a single receiver. We model the channel as a Gaussian multiple access channel with flat fading effects

$$Y(n) = \sum_{i=1}^{M} \sqrt{H_i(n)} X_i(n) + Z(n), \tag{1}$$

where $X_i(n)$ are the transmitted waveform with average power P_i, $H_i(n)$ is the channel state corresponding to the i-th user at time slot n, and $Z(n)$ is white Gaussian noise with variance N_0. The channel state process is assumed to be ergodic and bounded. We also assume that the channel states are known to all users and the receiver [1]. Throughout this work we assume that the transmission powers are fixed and no prior knowledge of channel statistics is available.

We model the speed of fading as follows:

$$|H_i(n+1) - H_i(n)| = V_n^i, \quad \text{for all } n,\ i = 1, \ldots, M, \tag{2}$$

where V_n^i is a nonnegative random variable bounded from above by $\hat{v}^i$, and $\{V_n^i\}$ are independent identically distributed (i.i.d.) for fixed i. Under slow fading conditions, the distribution of V_n^i is concentrated around zero.

We first consider the non-fading case where the channel state, $\boldsymbol{H}$, is fixed. The capacity region of the Gaussian multiple-access channel with no power control is described as follows [18]:

$$C_g(\boldsymbol{P}, \boldsymbol{H}) = \Big\{ \boldsymbol{R} \in \mathbb{R}^M_+ : \sum_{i \in S} R_i \leq C\Big(\sum_{i \in S} H_i P_i, N_0\Big), \quad \text{for all } S \subseteq \mathcal{M} = \{1, \ldots, M\} \Big\}, \tag{3}$$

where P_i and R_i are the i-th transmitter's power and rate, respectively. $C(P, N)$ denotes Shannon's formula for the capacity of additive white Gaussian noise (AWGN) channel given by

$$C(P, N) = \frac{1}{2} \log\left(1 + \frac{P}{N}\right) \quad \text{nats.} \tag{4}$$

For a multiple-access channel with fading, but fixed transmission powers P_i, the *throughput* capacity region is obtained by averaging the instantaneous capacity regions with respect to the fading process [19]:

$$C_a(\boldsymbol{P}) = \Big\{ \boldsymbol{R} \in \mathbb{R}^M_+ : \sum_{i \in S} R_i \leq \mathbb{E}_{\boldsymbol{H}}\Big[C\Big(\sum_{i \in S} H_i P_i, N_0\Big)\Big], \quad \text{for all } S \subseteq \{1, \ldots, M\} \Big\}, \tag{5}$$

[1] This assumption is satisfied in practice when the receiver measures the channels and feeds back the channel information to the users.

where $\boldsymbol{H}$ is a random vector with the stationary distribution of the fading process. Let us define the notion of boundary or dominant face for any of the capacity regions defined above.

Definition 1: The *dominant face* or *boundary* of a capacity region, denoted by $\mathcal{F}(\cdot)$, is defined as the set of all M-tuples in the capacity region such that no component can be increased without decreasing others while remaining in the capacity region.

III. Resource Allocation for a Fading Channel

The goal of the dynamic resource allocation problem is to find a rate allocation policy, $\mathcal{R}$, which is a map from the fading state $\boldsymbol{h}$ to the transmission rates, $\mathcal{R}(\boldsymbol{h}) = (\mathcal{R}_1(\boldsymbol{h}), \ldots, \mathcal{R}_M(\boldsymbol{h}))$. In the following we define the optimal rate allocation policy with respect to utility function, $u(\cdot)$.

Definition 2: [Optimal Policy] The optimal rate allocation policy denoted by $\mathcal{R}^*(\cdot)$ is a mapping that satisfies $\mathcal{R}^*(\boldsymbol{H}) \in C_g\big(\mathcal{P}^*(\boldsymbol{H}), \boldsymbol{H}\big)$ for all $\boldsymbol{H}$, such that

$$\mathbb{E}_{\boldsymbol{H}}[\mathcal{R}^*(\boldsymbol{H})] = \boldsymbol{R}^* \in \quad \text{argmax} \quad u(\boldsymbol{R}) \quad \text{subject to} \quad \boldsymbol{R} \in C_a(\boldsymbol{P}) \qquad (6)$$

The utility function $u(\boldsymbol{R})$ is assumed to satisfy the following conditions.

Assumption 1: The following conditions hold:

(a) *The utility function $u(\boldsymbol{R})$ is concave with respect to vector $\boldsymbol{R}$.*

(b) *$u(\boldsymbol{R})$ is monotonically non-decreasing with respect to R_i, for $i = 1, \ldots, M$.*

(c) *There exists a scalar B such that*

$$\|\boldsymbol{g}\| \leq B, \quad \textit{for all } \boldsymbol{g} \in \partial u(\boldsymbol{R}) \textit{ and all } \boldsymbol{R},$$

where $\partial u(\boldsymbol{R})$ denotes the subdifferential of u at $\boldsymbol{R}$, i.e., the set of all subgradients [2] of u at $\boldsymbol{R}$.

(d) *If $\boldsymbol{R}^\dagger = \text{argmax}_{\boldsymbol{R} \in C_g(\boldsymbol{P}, \boldsymbol{H})} u(\boldsymbol{R})$, then there exists a positive scalar A such that*

$$|u(\boldsymbol{R}^\dagger) - u(\boldsymbol{R})| \geq A\|\boldsymbol{R}^\dagger - \boldsymbol{R}\|^2, \quad \textit{for all } \boldsymbol{R} \in C_g(\boldsymbol{P}, \boldsymbol{H}).$$

Assumption 1(c) imposes a bound on subgradients of the utility function. In this paper, it is sufficient to have a utility function with bounded subgradient only in a neighborhood of optimal solution, but weakening Assumption 1(c) may require unnecessary technical details. Assumption 1(d) is a strong concavity type assumption which is satisfied for most of the utility functions. In fact, strong concavity of the utility implies Assumption 1(d), but it is not necessary.

Definition 3: [Greedy Policy] A *greedy* rate allocation policy, denoted by $\bar{\mathcal{R}}$, is given by

$$\bar{\mathcal{R}}(\boldsymbol{H}) = \quad \text{argmax} \quad u(\boldsymbol{R}) \quad \text{subject to} \quad \boldsymbol{R} \in C_g(\boldsymbol{P}, \boldsymbol{H}) \qquad (7)$$

i.e., for each channel state, the greedy policy chooses the rate vector that maximizes the utility function over the corresponding capacity region.

Note that the greedy policy is not necessarily optimal for general concave utility functions, i.e., the expected achieved rate does not maximize the utility over the throughput capacity region. However, the performance difference, i.e., utility difference between the expected rates assigned by the greedy and the optimal policy, is bounded and the bounds can be characterized in terms of channel variations and the structure of the utility function [20].

The maximization problem in (7) is a convex program and the optimal solution can be obtained by iterative methods such as the gradient projection method with approximate projection studied in [21]. The k-th iteration of this method is given by

$$\boldsymbol{R}^{k+1} = \tilde{\mathcal{P}}(\boldsymbol{R}^k + \alpha^k \boldsymbol{g}^k), \quad \boldsymbol{g}^k \in \partial u(\boldsymbol{R}^k), \qquad (8)$$

where $\boldsymbol{g}^k$ is a subgradient of u at $\boldsymbol{R}^k$, and α^k denotes the stepsize. $\tilde{\mathcal{P}}$ denotes the approximate projection operator which is defined in the following.

Definition 4: Let $X = \{\boldsymbol{x} \in \mathbb{R}^n | A\boldsymbol{x} \leq \boldsymbol{b}\}$ where A has non-negative entries. Let $\boldsymbol{y} \in \mathbb{R}^n$ violate the constraint $\boldsymbol{a}_i'\boldsymbol{x} \leq b_i$, for $i \in \{i_1, \ldots, i_l\}$. The approximate projection of $\boldsymbol{y}$ on X, denoted by $\tilde{\mathcal{P}}$, is given by

$$\tilde{\mathcal{P}}(\boldsymbol{y}) = \mathcal{P}_{i_1}(\ldots(\mathcal{P}_{i_{l-1}}(\mathcal{P}_{i_l}(\boldsymbol{y})))),$$

where $\mathcal{P}_{i_k}$ denotes the exact projection on the hyperplane $\{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{a}_{i_k}'\boldsymbol{x} = b_{i_k}\}$.

An example of approximate projection on a two-user multiple-access capacity region is illustrated in Figure 1. Note that the result of projection is not necessarily unique. However it is pseudo-nonexpansive, i.e., the distance between any feasible point and the projected point is smaller than its distance to the original point. Under Assumption 1 and specific

[2] The vector $\boldsymbol{g}$ is a subgradient of a concave function $f : D \to \mathbb{R}$ at x_0, if and only if $f(x) - f(x_0) \leq \boldsymbol{g}'(x - x_0)$ for all $x \in D$.

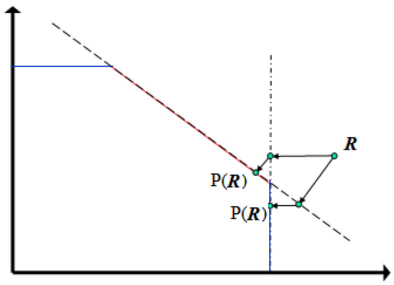

Fig. 1. Approximate projection of $\boldsymbol{R}$ on a two-user MAC region

stepsize rules, we established the convergence of the iterations in (8) to the optimal solution in (7) by using the pseudo-nonexpansiveness of the approximate projection (see [21], Proposition 2). We also showed by exploiting the polymatroid structure of the capacity region, each iteration in (8) can be computed in $O(M^3 \log M)$ time. However, for each channel state, finding even a "near-optimal" solution of the problem in (7) requires a large number of iterations, making the online evaluation of the greedy policy impractical. In the following section, we introduce an alternative rate allocation policy, which implements a single gradient projection iteration of the form (8) per time slot.

IV. Approximate Rate Allocation Policy

In this section, we assume that the channel state information is available instantly at each time slot n, and the computational resources are limited such that a single iteration of the gradient projection method in (8) can be implemented in each time slot.

Definition 5: [Approximate Policy] Given some fixed integer $k \geq 1$, we define the *approximate* rate allocation policy, $\widetilde{\mathcal{R}}$, as follows:

$$\widetilde{\mathcal{R}}\big(\boldsymbol{H}(0)\big) = \bar{\mathcal{R}}\big(\boldsymbol{H}(0)\big),$$
$$\widetilde{\mathcal{R}}\big(\boldsymbol{H}(n)\big) = \boldsymbol{R}_t^{\tau}, \quad \text{for all } n \geq 1, \tag{9}$$

where

$$\tau = \underset{0 \leq j < k-1}{\operatorname{argmax}} \; u(\boldsymbol{R}_t^j), \quad t = \left\lfloor \frac{n-1}{k} \right\rfloor, \tag{10}$$

and $\boldsymbol{R}_t^j \in \mathbb{R}^M$ is given by the following gradient projection iterations

$$\begin{aligned} \boldsymbol{R}_t^0 &= \tilde{P}_t\left[\widetilde{\mathcal{R}}\big(\boldsymbol{H}(kt)\big)\right], \\ \boldsymbol{R}_t^{j+1} &= \tilde{P}_t\left[\boldsymbol{R}_t^j + \alpha^j \boldsymbol{g}^j\right], \quad j = 1, \dots, k-1, \end{aligned} \tag{11}$$

where $\boldsymbol{g}^j$ is a subgradient of $u(\cdot)$ at $\boldsymbol{R}^j$, α^j denotes the stepsize and $\tilde{P}_t$ is the approximate projection on $C_g(\boldsymbol{P}, \boldsymbol{H}(kt))$.

For $k = 1$, (11) reduces to taking only one gradient projection iteration at each time slot. For $k > 1$, the proposed rate allocation policy essentially let the channel state change for a block of k consecutive time slots, and then takes k iterations of the gradient projection method with approximate projection. Note that to compute the policy at time slot n, we are using the channel state information at time slots $kt, k(t-1), \ldots$ Hence, in practice the channel measurements need to be computed every k time slots.

There is a tradeoff in choosing k, because taking only one gradient projection step may not be sufficient to get close enough to the greedy policy's operating point. Moreover, for large k the new operating point of the greedy policy can be far from the previous one, and k iterations may be insufficient again. So the parameter k should be chosen optimally to obtain the best performance for the approximate policy.

Before stating the main result, let us introduce some definitions and lemmas.

Definition 6: Let Q be a polyhedron described by a set of linear constraints, i.e.,

$$Q = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} \leq \boldsymbol{b}\}. \tag{12}$$

Define the *expansion* of Q by δ, denoted by $\mathcal{E}_\delta(Q)$, as the polyhedron obtained by relaxing all the constraints in (12), i.e., $\mathcal{E}_\delta(Q) = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} \leq \boldsymbol{b} + \delta\boldsymbol{1}\}$, where $\boldsymbol{1}$ is the vector of all ones.

Definition 7: Let X and Y be two polyhedra described by a set of linear constraints. Let $\mathcal{E}_d(X)$ be an *expansion* of X by relaxing its constraints by d. The distance $d_H(X, Y)$ between X and Y is defined as the minimum scalar d such that $X \subseteq \mathcal{E}_d(Y)$ and $Y \subseteq \mathcal{E}_d(X)$.

The next lemma shows that if the distance between two capacity regions is small, the distance between the optimal solutions of maximizing the utility function over these regions is also small.

Lemma 1: Let $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ be two different channel states. Also, let $\boldsymbol{R}_1^*$ and $\boldsymbol{R}_2^*$ be the optimal solution of maximizing a utility function over $C_g(\boldsymbol{P}, \boldsymbol{H}_1)$ and $C_g(\boldsymbol{P}, \boldsymbol{H}_2)$, respectively. If the utility satisfies Assumption 1, and

$$d_H\big(C_g(\boldsymbol{P}, \boldsymbol{H}_1), C_g(\boldsymbol{P}, \boldsymbol{H}_2)\big) \leq \delta$$

Then, we have

$$\|\boldsymbol{R}_1^* - \boldsymbol{R}_2^*\| \leq \delta^{\frac{1}{2}} \left[\delta^{\frac{1}{2}} + \left(\frac{B}{A}\right)^{\frac{1}{2}}\right]. \tag{13}$$

Proof: See Appendix I. ∎

In the following lemma, we translate the model for the speed of fading in terms of channel state variations into changes in the corresponding capacity regions.

Lemma 2: Let $\{[H_i(n)]_{i=1,\dots,M}\}$ be the fading process that satisfies condition in (2). We have

$$d_H\Big(C_g\big(\boldsymbol{P}, \boldsymbol{H}(n+1)\big), C_g\big(\boldsymbol{P}, \boldsymbol{H}(n)\big)\Big) \le W_n, \qquad (14)$$

where $\{W_n\}$ are nonnegative independent identically distributed random variables bounded from above by $\hat{w} = \frac{1}{2}\sum_{i=1}^{M} \hat{v}^i P_i$, where $\hat{v}^i$ is an upperbound on the process $\{V_n^i\}$ and P_i is the i-th user's transmission power.

Proof: By Definition 7 we have

$$\begin{aligned} & d_H\Big(C_g\big(\boldsymbol{P}, \boldsymbol{H}(n+1)\big), C_g\big(\boldsymbol{P}, \boldsymbol{H}(n)\big)\Big) \\ &= \max_{S\subseteq\mathcal{M}} \frac{1}{2}\left|\log\Big(1 + \frac{\sum_{i\in S}(H_i(n+1) - H_i(n))P_i}{1+\sum_{i\in S} H_i(n)P_i}\Big)\right| \\ &\le \max_{S\subseteq\mathcal{M}} \frac{\sum_{i\in S}|H_i(n+1) - H_i(n)|P_i}{2(1+\sum_{i\in S} H_i(n)P_i)} \\ &\le \frac{1}{2}\sum_{i=1}^{M}|H_i(n+1) - H_i(n)|P_i = \frac{1}{2}\sum_{i=1}^{M} V_n^i P_i. \qquad (15)\end{aligned}$$

Therefore, (14) is true for $W_n = \frac{1}{2}\sum_{i=1}^{M} V_n^i P_i$. Since V_n^i's are i.i.d. and bounded above by $\hat{v}_n^i$, W_n's are i.i.d. and bounded from above by $\frac{1}{2}\sum_{i=1}^{M} \hat{v}^i P_i$.

∎

The following lemma by Nedić and Bertsekas [22] addresses the convergence rate of the gradient projection method with constant stepsize.

Lemma 3: Let rate allocation policies $\bar{\mathcal{R}}$ and $\widetilde{\mathcal{R}}$ be given by Definition 3 and Definition 5, respectively. Also, let Assumption 1 hold and the stepsize α^n be fixed to some positive constant α. Then for a positive scalar ϵ we have

$$u\Big(\widetilde{\mathcal{R}}\big(\boldsymbol{H}(n)\big)\Big) \ge u\Big(\bar{\mathcal{R}}\big(\boldsymbol{H}(kt)\big)\Big) - \frac{\alpha B^2 + \epsilon}{2}, \qquad (16)$$

where k satisfies

$$k \ge \left\lfloor \frac{\|\boldsymbol{R}_t^0 - \bar{\mathcal{R}}\big(\boldsymbol{H}(kt)\big)\|^2}{\alpha\epsilon} \right\rfloor. \qquad (17)$$

Proof: See Proposition 2.3 of [22]. ∎

We next state our main result, which shows that the approximate rate allocation policy given by Definition 5 tracks the greedy policy within a neighborhood which is quantified as a function of the maximum speed of fading, the parameters of the utility function, and the transmission powers.

Theorem 1: Let Assumption 1 hold and the rate allocation policies $\bar{\mathcal{R}}$ and $\widetilde{\mathcal{R}}$ be given by Definition 3 and Definition 5, respectively. Let $k = \lfloor (\frac{2B}{Aw'})^{\frac{2}{3}} \rfloor$ and fix the stepsize to $\alpha = \big(\frac{16A}{B^2}\big)^{\frac{1}{3}} w'^{\frac{2}{3}}$ in Eq. (11), where $w' = \hat{w}^{\frac{1}{2}}\big(\hat{w}^{\frac{1}{2}} + (\frac{B}{A})^{\frac{1}{2}}\big)$, $\hat{w}$ is the upperbound on W_n as defined in Lemma 2, A and B are system parameters depending on the structure of utility function as in Assumption 1(c),(d). Then, we have

$$\|\widetilde{\mathcal{R}}\big(\boldsymbol{H}(n)\big) - \bar{\mathcal{R}}\big(\boldsymbol{H}(n)\big)\| \le 2\theta = 2(\frac{2B}{A})^{\frac{2}{3}} w'^{\frac{1}{3}}. \qquad (18)$$

Proof: First, we show that

$$\|\widetilde{\mathcal{R}}\big(\boldsymbol{H}(n)\big) - \bar{\mathcal{R}}\big(\boldsymbol{H}(kt)\big)\| \le \theta = (\frac{2B}{A})^{\frac{2}{3}} w'^{\frac{1}{3}}, \qquad (19)$$

where $t = \lfloor \frac{n-1}{k} \rfloor$. The proof is by induction on t. For $t = 0$ the claim is trivially true. Now suppose that (19) is true for some positive t. Hence, it also holds for $n = k(t+1)$ by induction hypothesis, i.e.,

$$\|\boldsymbol{R}_{t+1}^0 - \bar{\mathcal{R}}\big(\boldsymbol{H}(kt)\big)\| \le \theta. \qquad (20)$$

On the other hand, by Lemma 1 and Lemma 2 we have

$$\|\bar{\mathcal{R}}\big(\boldsymbol{H}(k(t+1))\big) - \bar{\mathcal{R}}\big(\boldsymbol{H}(kt)\big)\| \le kw' \le \theta. \qquad (21)$$

Therefore, by triangle inequality we have the following

$$\|\boldsymbol{R}_{t+1}^0 - \bar{\mathcal{R}}\big(\boldsymbol{H}(k(t+1))\big)\| \le 2\theta. \qquad (22)$$

After plugging the corresponding values of α and θ, it is straightforward to show that (17) holds for $\epsilon = \alpha B^2$. Thus, we can apply Lemma 3 to show

$$\left| u\Big(\widetilde{\mathcal{R}}\big(\boldsymbol{H}(n)\big)\Big) - u\Big(\bar{\mathcal{R}}\big(\boldsymbol{H}(k(t+1))\big)\Big)\right| \le \alpha B^2. \qquad (23)$$

By Assumption 1(d) we can write

$$\|\widetilde{\mathcal{R}}\big(\boldsymbol{H}(n)\big) - \bar{\mathcal{R}}\big(\boldsymbol{H}(k(t+1))\big)\| \le (\frac{\alpha B^2}{A})^{\frac{1}{2}} = \theta. \qquad (24)$$

Therefore, the proof of (19) is complete by induction.

Again by applying Lemma 1 and Lemma 2 we have

$$\|\bar{\mathcal{R}}\big(\boldsymbol{H}(n)\big) - \bar{\mathcal{R}}\big(\boldsymbol{H}(kt)\big)\| \le kw' \le \theta, \qquad (25)$$

and the desired result directly follows from (19) and (25) by triangle inequality. ∎

It is straightforward to show that the parameters k and α in Theorem 1 are designed such that the smallest tracking neighborhood, θ, is obtained for the approximate policy presented in Definition 5 with constant stepsize. The proof is by parameterizing θ, the size of the tracking neighborhood, in terms of k and minimizing $\theta(k)$ by relaxing k to be a real and differentiating with respect to k. We eliminate the full proof for brevity. Theorem 1 provides a bound on the size of the tracking neighborhood as a function of the maximum speed of fading, denoted by $\hat{w}$, which may be too conservative. It is of interest to provide a rate allocation policy and a bound on the size of its tracking neighborhood as a function of the average speed of fading. The next section addresses this issue.

V. Improved Approximate Rate Allocation Policy

In this section, we design an efficient rate allocation policy that tracks the greedy policy within a neighborhood characterized by the average speed of fading which is typically much smaller than the maximum speed of fading. We consider policies which can implement one gradient projection iteration per time slot.

Unlike the approximate policy given by (9) which uses the channel state information once in every k time slots, we present an algorithm which uses the channel state information in all time slots. Roughly speaking, this method takes fixed a number of gradient projection iterations only after the change in the channel state has reached a certain threshold.

Definition 8: [Improved Approximate Policy] Let $\{W_n\}$ be the sequence of nonnegative random variables as defined in Lemma 2, and γ be a positive constant. Define the sequence $\{T_i\}$ as

$$\begin{aligned} T_0 &= 0, \\ T_{i+1} &= \min\{t \mid \sum_{n=T_i}^{t-1} W_n \geq \gamma\}. \end{aligned} \tag{26}$$

Define the *improved approximate* rate allocation policy, $\widehat{\mathcal{R}}$, with parameters γ and k, as the following:

$$\begin{aligned} \widehat{\mathcal{R}}\big(\boldsymbol{H}(0)\big) &= \bar{\mathcal{R}}\big(\boldsymbol{H}(0)\big), \\ \widehat{\mathcal{R}}\big(\boldsymbol{H}(n)\big) &= \boldsymbol{R}_t^{\tau}, \quad \text{for all } n \geq 1, \end{aligned} \tag{27}$$

where

$$t = \max\{i \mid T_i < n\}, \tag{28}$$

$$\tau = \underset{0\leq j<k-1}{\arg\max}\ u(\boldsymbol{R}_t^j), \tag{29}$$

and $\boldsymbol{R}_t^j \in \mathbb{R}^M$ is given by the following gradient projection iterations

$$\begin{aligned} \boldsymbol{R}_t^0 &= \tilde{P}_t\left[\widehat{\mathcal{R}}\big(\boldsymbol{H}(T_t)\big)\right], \\ \boldsymbol{R}_t^{j+1} &= \tilde{P}_t\left[\boldsymbol{R}_t^j + \alpha^j \boldsymbol{g}^j\right], \quad j = 1, \ldots, k-1, \end{aligned} \tag{30}$$

where $\boldsymbol{g}^j$ is a subgradient of $u(\cdot)$ at $\boldsymbol{R}^j$, α^j denotes the stepsize and $\tilde{P}_t$ is the approximate projection on $C_g(\boldsymbol{P}, \boldsymbol{H}(T_t))$.

Theorem 2: Let t be as defined in (28), and let $\bar{w}$ denote the expected value of W_n. If $k = \frac{\gamma}{\bar{w}}$, then we have

$$\lim_{n\to\infty} \frac{n}{tk} = 1, \quad \text{with probability } 1. \tag{31}$$

Proof: The sequence $\{T_i\}$ is obtained as the random walk generated by W_n's cross the threshold level γ. Since W_n's are positive random variables, we can think of the threshold crossing as a renewal process, denoted by $N(\cdot)$, with inter arrivals W_n.

We can rewrite the limit as follows

$$\lim_{n\to\infty} \frac{n - N(t\gamma) + N(t\gamma)}{tk} = \lim_{n\to\infty} \frac{n - N(t\gamma)}{tk} + \bar{w}\frac{N(t\gamma)}{t\gamma}. \tag{32}$$

Since the random walk will hit the threshold with probability 1, the first term goes to zero with probability 1. Also, by Strong law for renewal processes the second terms goes to 1 with probability 1 (see [23], p.60). ■

Theorem 2 essentially guarantees that the number of gradient projection iterations is the same as the number of channel measurements in the long run with probability 1.

Theorem 3: Let Assumption 1 hold and the rate allocation policies $\bar{\mathcal{R}}$ and $\widehat{\mathcal{R}}$ be given by Definition 3 and Definition 8, respectively. Also, let $\gamma = c(\frac{B}{A})^{\frac{3}{4}}\bar{w}^{\frac{1}{4}}$, and $k = \lfloor \frac{\gamma}{\bar{w}} \rfloor$ and fix the stepsize to $\alpha = \frac{A\gamma^2}{B^2}$ in (30), where $c \geq 1$ is a constant satisfying the following equation

$$\frac{(c^2-1)^8}{2^8 c^4} = \hat{w}. \tag{33}$$

Then

$$\|\widehat{\mathcal{R}}\big(\boldsymbol{H}(n)\big) - \bar{\mathcal{R}}\big(\boldsymbol{H}(n)\big)\| \leq 2\gamma + (\frac{\gamma B}{A})^{\frac{1}{2}}. \tag{34}$$

Proof: We follow the line of proof of Theorem 1. First, by induction on t we show that

$$\|\widehat{\mathcal{R}}\big(\boldsymbol{H}(n)\big) - \bar{\mathcal{R}}\big(\boldsymbol{H}(T_t)\big)\| \leq \gamma, \tag{35}$$

where t is defined in (28). The base is trivial. Similar to (20), by induction hypothesis we have

$$\|\boldsymbol{R}_{t+1}^0 - \bar{\mathcal{R}}\big(\boldsymbol{H}(T_t)\big)\| \leq \gamma. \tag{36}$$

By definition of T_i in (26) we can write

$$d_H\Big(C_g\big(\boldsymbol{P}, \boldsymbol{H}(T_{t+1})\big), C_g\big(\boldsymbol{P}, \boldsymbol{H}(T_t)\big)\Big) \leq \gamma. \tag{37}$$

Thus, by Lemma 1, we have

$$\|\bar{\mathcal{R}}\big(\boldsymbol{H}(T_{t+1})\big) - \bar{\mathcal{R}}\big(\boldsymbol{H}(T_t)\big)\| \leq \gamma^{\frac{1}{2}}\big(\gamma^{\frac{1}{2}} + (\frac{B}{A})^{\frac{1}{2}}\big). \tag{38}$$

Therefore, by combining (36) and (38) by triangle inequality we obtain

$$\|\boldsymbol{R}^0_{t+1} - \bar{\mathcal{R}}\big(\boldsymbol{H}(T_{t+1})\big)\| \leq 2\gamma + (\frac{\gamma B}{A})^{\frac{1}{2}}. \qquad (39)$$

Using the fact that $\bar{w} \leq \hat{w} = \frac{(c^2-1)^8}{2^8 c^4}$, after a few steps of straightforward manipulations we can show that

$$\|\boldsymbol{R}^0_{t+1} - \bar{\mathcal{R}}\big(\boldsymbol{H}(T_{t+1})\big)\|^2 \leq \Big(2\gamma + (\frac{\gamma B}{A})^{\frac{1}{2}}\Big)^2 \leq c^4 \frac{\gamma B}{A}. \quad (40)$$

Now by plugging in (17) the values of α and γ in terms of system parameters we can verify that

$$k = \Big\lfloor \frac{\gamma}{\bar{w}} \Big\rfloor = \Bigg\lfloor \frac{c^4 \frac{\gamma B}{A}}{A \frac{\gamma^2}{B^2} A\gamma^2} \Bigg\rfloor \geq \Bigg\lfloor \frac{\|\boldsymbol{R}^0_{t+1} - \bar{\mathcal{R}}\big(\boldsymbol{H}(T_{t+1})\big)\|^2}{\alpha\epsilon} \Bigg\rfloor. \qquad (41)$$

Hence, we can apply Lemma 3 for $\epsilon = A\gamma^2$, and conclude

$$\bigg| u\Big(\widehat{\mathcal{R}}\big(\boldsymbol{H}(n)\big)\Big) - u\Big(\bar{\mathcal{R}}\big(\boldsymbol{H}(T_{t+1})\big)\Big) \bigg| \leq \alpha B^2. \qquad (42)$$

By exploiting Assumption 1(d) we have

$$\|\widehat{\mathcal{R}}\big(\boldsymbol{H}(n)\big) - \bar{\mathcal{R}}\big(\boldsymbol{H}(T_{t+1})\big)\| \leq (\frac{\alpha B^2}{A})^{\frac{1}{2}} = \gamma. \qquad (43)$$

Therefore, the proof of (35) is complete by induction. Similarly to (38) we have

$$\|\bar{\mathcal{R}}\big(\boldsymbol{H}(n)\big) - \bar{\mathcal{R}}\big(\boldsymbol{H}(T_t)\big)\| \leq \gamma^{\frac{1}{2}}\big(\gamma^{\frac{1}{2}} + (\frac{B}{A})^{\frac{1}{2}}\big), \qquad (44)$$

and (34) follows immediately from (35) and (44) by invoking triangle inequality. ■

Theorem 2 and Theorem 3 guarantee that the presented rate allocation policy tracks the greedy policy within a small neighborhood while with probability 1, only one gradient projection iteration is computed per time slot. The neighborhood is characterized in terms of the average behavior of channel variations and vanishes as the fading speed decreases.

VI. Conclusion

We study the problem of rate allocation in a fading multiple access channel with no power control from an information theoretic point of view. Our goal is to approximate the optimal rate allocation policy, which yields an average rate that maximizes a general concave utility function of transmission rates over the throughput capacity region of the multiple-access channel.

We present a dynamic rate allocation policy which takes a block of channel measurements and implements the same number of gradient projection iterations with approximate projection at the end of each block. This rate allocation policy tracks the greedy policy within a small neighborhood whose size decreases as a function of the maximum speed of fading.

In order to provide a bound on the tracking neighborhood in terms of average speed of fading, we present an alternative rate allocation policy. This policy adaptively selects variable block lengths for channel measurements using feedback information about the current channel states. It implements a fixed number of gradient iterations at the end of these blocks. We show that the ratio of the total number of channel measurements and the number of gradient iterations converges to 1 with probability one. We also provide a bound on the size of the neighborhood with which the new policy tracks the greedy policy as a function of the average speed of fading. The proposed dynamic rate allocation policies are efficiently implementable since they require a single gradient projection step per channel measurement and the projection can be done in time polynomial in the number of users.

References

[1] X. Wang and G.B. Giannakis. Energy-efficient resource allocation in time division multiple-access over fading channels. Preprint, 2005.

[2] S.J. Oh, Z. Danlu, and K.M. Wasserman. Optimal resource allocation in multiservice CDMA networks. *IEEE Transactions on Wireless Communications*, 2(4):811–821, 2003.

[3] J.B. Kim and M.L. Honig. Resource allocation for multiple classes of DS-CDMA traffic. *IEEE Transactions on Vehicular Technology*, 49(2):506–519, 2000.

[4] D. Tse and S. Hanly. Multiaccess fading channels part I: Polymatroid structure, optimal resource allocation and throughput capacities. *IEEE Transactions on Information Theory*, 44(7):2796–2815, 1998.

[5] X. Lin and N. Shroff. The impact of imperfect scheduling on cross-layer rate control in multihop wireless networks. In *Proceedings of IEEE Infocom*, Miami, FL, March 2005.

[6] A. Eryilmaz and R. Srikant. Fair resource allocation in wireless networks using queue-length based scheduling and congestion control. In *Proceedings of IEEE Infocom*, volume 3, pages 1794–1803, Miami, FL, March 2005.

[7] M.J. Neely, E. Modiano, and C. Li. Fairness and optimal stochastic control for heterogeneous networks. In *Proceedings of IEEE Infocom*, pages 1723–1734, Miami, FL, March 2005.

[8] A. Stolyar. Maximizing queueing network utility subject to stability: Greedy primal-dual algorithm. *Queueing Systems*, 50(4):401–457, 2005.

[9] A. Eryilmaz, A. Ozdaglar, and E. Modiano. Polynomial complexity algorithms for full utilization of multi-hop wireless networks. In *Proceedings of IEEE Infocom*, Anchorage, AL, May 2007.

[10] L. Tassiulas. Linear complexity algorithms for maximum throughput in radio networks and input queued switches. In *Proceedings of IEEE Infocom*, pages 533–539, 1998.

[11] E. Modiano, D. Shah, and G. Zussman. Maximizing throughput in wireless networks via gossiping. In *ACM SIGMETRICS/IFIP Performance*, 2006.

[12] S. Sanghavi, L. Bui, and R. Srikant. Distributed link scheduling with constant overhead, 2007. Technical Report.

[13] C. Joo, X. Lin, and N. Shroff. Performance limits of greedy maximal matching in multi-hop wireless networks. In *Proceedings of IEEE Conference on Decision and Control*, New Orleans LA, December 2007.

[14] A. Eryilmaz, A. Ozdaglar, D. Shah, and E. Modiano. Randomized algorithms for optimal control of wireless networks, 2007. ICCOPT Conference, Hamilton CA.

[15] S. Vishwanath, S.A. Jafar, and A. Goldsmith. Optimum power and rate allocation strategies for multiple access fading channels. In *Proceedings of IEEE VTC*, 2001.

[16] D. Yu and J.M. Cioffi. Iterative water-filling for optimal resource allocation in OFDM multiple-access and broadcast channels. In *Proceedings of IEEE GLOBECOM*, 2006.

[17] K. Seong, R. Narasimhan, and J. Cioffi. Scheduling for fading multiple access channels with heterogeneous QoS constraints. In *Proceedings of International Symposium on Information Theory*, 2007.

[18] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. John Wiley and Sons, Inc., New York, New York, 1991.

[19] S. Shamai and A.D. Wyner. Information theoretic considerations for symmetric, cellular, multiple-access fading channels part I. *IEEE Transactions on Information Theory*, 43(6):1877–1894, 1997.

[20] A. ParandehGheibi, A. Eryilmaz, A. Ozdaglar, and M. Medard. Resource allocation in fading multiple-access channel. Preprint, 2007.

[21] A. Parandehgheibi, A. Ozdaglar, M. Medard, and A. Eryilmaz. Utility maximization in multiple access channels. In *Asilomar Conference on Signals, Systems, and Computers*, 2007.

[22] A. Nedić and D.P. Bertsekas. *Convergence Rate of Incremental Subgradient Algorithms*. Stochastic Optimization: Algorithms and Applications (S. Uryasev and P. M. Pardalos, Editors), Kluwer Academic Publishers, 2000.

[23] R. Gallager. *Discrete Stochastic Processes*. Kluwer Academic Publishers, London, United Kingdom, 1996.

APPENDIX I
PROOF OF LEMMA 1

Without loss of generality assume that $u(\boldsymbol{R}_2^*) \geq u(\boldsymbol{R}_1^*)$. To simplify the notations for capacity regions, let C_1 be a *polymatroid*, i.e.,

$$C_1 = \left\{ \boldsymbol{R} \in \mathbb{R}_+^M : \sum_{i \in S} R_i \leq f(S), \text{ for all } S \subseteq \mathcal{M} \right\}, \quad (45)$$

for some submodular function $f(S)$, and C_2 be an *expansion* of C_1 by δ. We first show that for every $\boldsymbol{R} \in \mathcal{F}(C_2)$, there exists a vector $\boldsymbol{R}' \in \mathcal{F}(C_1)$ such that $\|\boldsymbol{R} - \boldsymbol{R}'\| \leq \delta$, where $\mathcal{F}(\cdot)$ denotes the dominant face of a capacity region as in Definition 1.

Assume R is a vertex of C_2. Then the polymatroid structure of C_2 implies that R is the intersection of M constraints corresponding to a chain of subsets of $\mathcal{M}$. Hence, there is some $k \in \mathcal{M}$ such that $R_k = f(\{k\}) + \delta$. Choose $\boldsymbol{R}'$ as follows

$$R_i' = \begin{cases} R_i - \delta, & i = k \\ R_i, & \text{otherwise.} \end{cases} \quad (46)$$

$\boldsymbol{R}'$ is obviously in a δ-neighborhood of $\boldsymbol{R}$. Moreover, the constraint corresponding to the set $\mathcal{M}$ is active for $\boldsymbol{R}'$, so we just need to show that R' is feasible in order to prove that it is on the dominant face. First, let us consider the sets S that contain k. We have

$$\sum_{i \in S} R_i' = \sum_{i \in S} R_i - \delta \leq f(S). \quad (47)$$

Second, consider the case that $k \notin S$.

$$\begin{aligned} \sum_{i \in S} R_i' &= \sum_{i \in S \cup \{k\}} R_i' - R_k + \delta \\ &\leq f(S \cup \{k\}) + \delta - R_k \\ &\leq f(S) + f(\{k\}) + \delta - R_l \\ &= f(S). \end{aligned}$$

where the first inequality come from (47), and the second inequality is valid because of the submodularity of the function $f(\cdot)$.

The previous argument establishes that the claim is true for each vertex $\boldsymbol{R}_j$ of the dominant face. But every other point $\boldsymbol{R}$ on the dominant face can be represented as a convex combination of the vertices, i.e.,

$$\boldsymbol{R} = \sum_j \alpha_j \boldsymbol{R}_j, \qquad \sum_j \alpha_j = 1, \alpha_j \geq 0.$$

Using the convexity of the norm function, it is quite straightforward to show that the desired $\boldsymbol{R}'$ is given by

$$\boldsymbol{R}' = \sum_j \alpha_j \boldsymbol{R}_j',$$

where $\boldsymbol{R}_j'$ is obtained for each $\boldsymbol{R}_j$ in the same manner as in (46).

So we have shown that there exists some $\boldsymbol{R}$ on the dominant face of $C_1 = C_g(\boldsymbol{P}, \boldsymbol{H}_1)$ such that $\|\boldsymbol{R}_2^* - \boldsymbol{R}\| \leq \delta$. Thus, from the hypothesis and the fact that $u(\boldsymbol{R}_2^*) \geq u(\boldsymbol{R}_1^*) \geq u(\boldsymbol{R})$, we have

$$u(\boldsymbol{R}_2^*) - u(\boldsymbol{R}) = |u(\boldsymbol{R}_2^*) - u(\boldsymbol{R})| \leq B\|\boldsymbol{R}_2^* - \boldsymbol{R}\| \leq B\delta. \quad (48)$$

Now suppose that $\|\boldsymbol{R}_1^* - \boldsymbol{R}\| > (\frac{B}{A}\delta)^{\frac{1}{2}}$, hence by Assumption 1(d) we have

$$u(\boldsymbol{R}_1^*) - u(\boldsymbol{R}) = |u(\boldsymbol{R}_1^*) - u(\boldsymbol{R})| > B\delta. \quad (49)$$

By subtracting (48) from (49) we obtain $u(\boldsymbol{R}_2^*) < u(\boldsymbol{R}_1^*)$ which is a contradiction. Therefore, $\|\boldsymbol{R}_1^* - \boldsymbol{R}\| \leq (\frac{B}{A}\delta)^{\frac{1}{2}}$, and the desired result follows immediately by invoking the triangle inequality.