A Generalized Gradient Scheduling Algorithm in Wireless Networks for Variable Rate Transmission

Xiaolu Zhang, Meixia Tao and Chun Sum Ng
Department of Electrical and Computer Engineering
National University of Singapore
Singapore 117576
Email: {zhangxiaolu, mxtao, elengcs}@nus.edu.sg.

Abstract—Average transmission rate and rate oscillation are two important performance metrics for most wireless services. Both are often needed to be optimized in multi-user scheduling and resource management. In this paper we introduce a utility function that increases with average rate but decreases with rate variance. It is capable of facilitating resource allocation with flexible combinations of the two performance metrics. A generalized gradient scheduling algorithm (GGSA) is then developed to maximize the proposed utility. It is shown that the best scheduler should maximize the sum of concave functions of instantaneous transmission rates in order to maximize the utility of average rate and rate oscillation. The scheduler reduces to the traditional gradient scheduling algorithm when the rate variance term in the new utility function is omitted. We analyze the dynamics of average transmission rates and rate variances using ordinary differential equation and show that GGSA is asymptotically optimal under the condition that the transmission rate vector, after an appropriate scaling, converges to a fixed vector as time goes into infinity.

I. INTRODUCTION

We consider the multi-user scheduling problem over a time-varying wireless fading channel. Multiple users are served by a common base station. In each time frame a scheduler determines the transmission parameters of all users, such as power and rate, according to channel state information and system preferences.

Many existing scheduling schemes are formulated as a utility-based throughput allocation problem, such as [1]–[4]. In [1], the scheme is in essence restricted to a linear utility function. It is shown that the sum capacity of a wireless system is maximized when only the user with the maximum achievable rate is chosen for transmission. However, this scheme could result in significant unfairness among users with asymmetric channels. In order to address the fairness concerns, the *proportional fair* (PF) scheduling algorithm is proposed in [2] and [5] by using a log-form utility function of average transmission rate. The PF scheduling is then generated by the authors in [3] to *gradient scheduling* algorithm, which applies to any increasing and concave utility function of average transmission rate. The gradient scheduling algorithm suggests that at any time instant, the best scheduler should maximize the sum of weighted transmission rates (Throughout this paper, we will use the term “transmission rate” to refer to the instantaneous transmission rate in a time frame) and the weight depends on system preferences and channel conditions.

Although the gradient scheduling algorithm is developed from a myopic view of the optimization problem (maximizing the utility function of average transmission rate), Stolyar in [4] proves its asymptotic optimality for a general model, which allows for simultaneous transmission of multiple users and the set of scheduling decisions to be discrete (as in [2]).

The gradient algorithms in [2], [4], [6] can improve efficiency in terms of average transmission rate but also result in high rate oscillation. As we know, many services such as wireless audio and video transmission generally have a fairly low tolerance of rate oscillation. They expect a balance between average rate and rate oscillation, both of which are important performance metrics for these applications. This motivates the search for a new criterion for multi-user scheduling by modifying the utility function in a way that penalizes rate oscillation. Rate oscillation can be measured using a statistical rate variance. We demonstrate later that a utility function that increases with average rate but decreases with rate variance can be used to facilitate the choice of the combinations of average rate and rate oscillation. A gradient based scheduling algorithm is developed to maximize the proposed utility function. The proposed algorithm reduces to the traditional gradient algorithm in [3] and [4] when we omit the rate variance term in the new utility function so that the utility is a function of the average rate only. Thus, we refer to our algorithm as the *generalized gradient scheduling algorithm* (GGSA). It can be shown that at each time frame, the best scheduler of GGSA is not to maximize the sum of weighted transmission rates (as in [3] and [4]), but the sum of concave functions of transmission rates. This coincides with the idea in [7] that maximizing the sum of concave functions of transmission rates can obtain a balance between average transmission rate and rate oscillation. Next, we analyze the dynamics of transmission rate using ordinary differential equation and show that GGSA is asymptotically optimal under the condition that the transmission rate vector, after an appropriate scaling, converges to a fixed vector as time goes into infinity.

The rest of the paper is organized as follows. In Section II, we introduce basic notations and conventions. In Section III, the system and channel model are presented. Section IV describes the GGSA for variable rate transmission. We study the dynamics of average transmission rates and rate variances

in Section V. Numerical results are given in Section VI. Finally, Section VII concludes the paper.

II. BASIC NOTATIONS AND CONVENTIONS

We use standard notations $\mathbb{R}$ and $\mathbb{R}_+$ to denote the sets of real and real nonnegative numbers. Correspondingly, $\mathbb{R}^N$ denotes a standard vector space, with elements $\mathbf{a} \in \mathbb{R}^N$ being column vectors $\mathbf{a} := (a_1, \dots, a_N)^T$.

For vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^N$,

$$\mathbf{a}^T \cdot \mathbf{b} \doteq \sum_{i=1}^N a_i b_i$$

and

$$\mathbf{a} \times \mathbf{b} \doteq (a_1 b_1, \dots, a_N b_N)^T.$$

The Euclidean norm $|\mathbf{a}| \doteq \sqrt{\mathbf{a}^T \cdot \mathbf{a}}$.

The gradient of a function $U(\mathbf{a}, \mathbf{b})$ with respect to $\mathbf{a}$ is denoted by

$$\nabla_{\mathbf{a}} U(\mathbf{a}, \mathbf{b}) \doteq \left(\frac{\partial U(\mathbf{a}, \mathbf{b})}{\partial a_1}, \dots, \frac{\partial U(\mathbf{a}, \mathbf{b})}{\partial a_N} \right)^T.$$

III. SYSTEM MODEL

Consider a wireless system with a base station (BS) and a finite set of users, denoted by $\mathcal{N} = \{1, 2, \dots, N\}$. The communication link between each user and the base station is modelled as a time-varying fading channel. And the transmission is centrally coordinated by BS on a time frame basis. In time frame k , the channel state between BS and user i is denoted as $g_i(k) \in \mathcal{G}$, where $\mathcal{G}$ is the channel state space. Here we assume that $\{g_i(k), k = 1, 2, \dots\}$ is ergodic and stationary with the distribution γ .

We consider the situation where each user always has data for transmission, and are interested in optimizing the pair of the average rates and the rate variances $(\mathbf{m}, \boldsymbol{\sigma}^2) := (m_1, \dots, m_N, \sigma_1^2, \dots, \sigma_N^2)$ so as to maximize a utility function $U(\mathbf{m}, \boldsymbol{\sigma}^2)$. Here, rate variance is chosen to quantify the rate oscillation. The utility function U represents the system preferences. Since the system expects a high average transmission rate and a low rate oscillation, the utility is assumed to be increasing in $\mathbf{m}$ but decreasing in $\boldsymbol{\sigma}^2$. Mathematically, the problem can be expressed as

$$\begin{aligned} \max \quad & U(\mathbf{m}, \boldsymbol{\sigma}^2) \\ \text{s.t.} \quad & (\mathbf{m}, \boldsymbol{\sigma}^2) \in \mathcal{V}, \end{aligned} \quad (1)$$

where we further assume that U is continuously differentiable in $\mathbf{m}$ and $\boldsymbol{\sigma}^2$ but not necessarily concave. The steady-state average rate and rate variance (ARRV) region $\mathcal{V}$ is defined as

$$\begin{aligned} \mathcal{V} := \{ & (\mathbf{m}, \boldsymbol{\sigma}^2) \in \mathbb{R}_+^{2N} : \exists \mathbf{r}(\mathbf{g}) \in \mathcal{R}(\mathbf{g}) \ \forall \mathbf{g} \in \mathcal{G} \\ & \text{such that } \mathbf{m} = \int_{\mathcal{G}} \mathbf{r}(\mathbf{g}) \gamma(d\mathbf{g}), \text{ and} \\ & \boldsymbol{\sigma}^2 = \int_{\mathcal{G}} [\mathbf{r}(\mathbf{g}) \times \mathbf{r}(\mathbf{g}) - \mathbf{m} \times \mathbf{m}] \ \gamma(d\mathbf{g}) \}, \end{aligned}$$

where $\mathbf{r}$ is the transmission rate vector, and $\mathcal{R}(\mathbf{g})$ is the achievable rate region when the network channel state is $\mathbf{g}$.

IV. GENERALIZED GRADIENT SCHEDULING

In this section we develop a generalized gradient scheduling algorithm to solve the problem in (1). Let $X_i(k)$ and $Y_i(k)$ denote the average rate and rate variance, respectively, for user i up to the k -th time frame. If $r_i(k)$ is the selected transmission rate of user i at time frame k , then $X_i(k)$ and $Y_i(k)$ can be updated as:

$$\begin{aligned} X_i(k+1) &= \frac{1}{k} \sum_{l=1}^k r_i(l) \\ &= X_i(k) + \frac{1}{k} [r_i(k) - X_i(k)]; \\ Y_i(k+1) &= \frac{1}{k} \sum_{l=1}^k [r_i(l) - X_i(l)]^2 \\ &= Y_i(k) + \frac{1}{k} \left\{ [r_i(k) - X_i(k)]^2 - Y_i(k) \right\}. \end{aligned}$$

If we use a small fixed step size μ to replace $1/k$, then the recursive forms can be rewritten as

$$X_i(k+1) = X_i(k) + \mu [r_i(k) - X_i(k)]; \quad (2)$$

$$Y_i(k+1) = Y_i(k) + \mu \left\{ [r_i(k) - X_i(k)]^2 - Y_i(k) \right\}. \quad (3)$$

Instead of finding the optimal solution to the problem in (1) directly, we consider a simple gradient based scheduling policy which attempts to adapt the current rates such that we have the largest first order change in the utility, namely

$$\begin{aligned} & U(\mathbf{X}(k+1), \mathbf{Y}(k+1)) - U(\mathbf{X}(k), \mathbf{Y}(k)) \\ &= U(\mathbf{X}(k) + \mu[\mathbf{r}(k) - \mathbf{X}(k)], \mathbf{Y}(k) + \mu\{[\mathbf{r}(k) - \mathbf{X}(k)] \times \\ & \quad [\mathbf{r}(k) - \mathbf{X}(k)] - \mathbf{Y}(k)\}) - U(\mathbf{X}(k), \mathbf{Y}(k)) \\ &= \mu \nabla_{\mathbf{m}} U(\mathbf{X}(k), \mathbf{Y}(k))^T \cdot [\mathbf{r}(k) - \mathbf{X}(k)] + \\ & \quad \mu \nabla_{\boldsymbol{\sigma}^2} U(\mathbf{X}(k), \mathbf{Y}(k))^T \cdot \{[\mathbf{r}(k) - \mathbf{X}(k)] \times \\ & \quad [\mathbf{r}(k) - \mathbf{X}(k)] - \mathbf{Y}(k)\}. \end{aligned} \quad (4)$$

In (4), the last equality holds when the step size μ is sufficiently small. The best scheduler of the system in state $\mathbf{g}$ is to choose a transmission rate vector $\mathbf{r}(k)$ that satisfies

$$\begin{aligned} \mathbf{r}(k) &= \arg \max_{\mathbf{v}(k) \in \mathcal{R}(\mathbf{g})} \nabla_{\mathbf{m}} U(\mathbf{X}(k), \mathbf{Y}(k))^T \cdot \mathbf{v}(k) + \\ & \quad \nabla_{\boldsymbol{\sigma}^2} U(\mathbf{X}(k), \mathbf{Y}(k))^T \cdot \\ & \quad [(\mathbf{v}(k) - \mathbf{X}(k)) \times (\mathbf{v}(k) - \mathbf{X}(k))] \\ &= \arg \max_{\mathbf{v} \in \mathcal{R}(\mathbf{g})} \sum_{i=1}^N [a_i(k) v_i^2(k) + b_i(k) v_i(k)] \end{aligned} \quad (5)$$

where

$$a_i(k) = \frac{\partial U(\mathbf{X}(k), \mathbf{Y}(k))}{\partial \sigma_i^2},$$

and

$$b_i(k) = \frac{\partial U(\mathbf{X}(k), \mathbf{Y}(k))}{\partial m_i} - 2X_i(k) \frac{\partial U(\mathbf{X}(k), \mathbf{Y}(k))}{\partial \sigma_i^2}.$$

Remark: Let us observe the best scheduler in (5). When U strictly increases in m_i and decreases in σ_i^2 for all i (it is a natural assumption in practice), we have $a_i(k) < 0$

and $b_i(k) > 0$ for all i and k . Hence, unlike the gradient scheduling algorithms in [3] and [4] which maximize the sum of weighted transmission rates at each time frame, the best scheduler of GGSA is to maximize the sum of concave functions of transmission rate

$$u_i(r_i) := a_i r_i^2 + b_i r_i. \quad (6)$$

The advantage of GGSA is that it can jointly optimize the average rates and rate oscillations. Consider a simple case of GGSA corresponding to a utility function

$$U(\mathbf{m}, \boldsymbol{\sigma}) = \sum_{i=1}^N [\log(m_i) - \alpha \log(\sigma_i^2)], \quad (7)$$

where $\alpha \geq 0$ is a parameter that reflects the sensitivity of the communication service to rate oscillation. It is easy to see that when $\alpha = 0$, it is just the same as the utility in the PF algorithm. Thus, we refer to it as PF-GGSA. The best choice of the PF-GGSA is to select a rate vector that satisfies

$$\mathbf{r} = \arg \max_{\mathbf{v} \in \mathcal{R}(\mathbf{g})} \sum_{i=1}^N u_i(v_i) \quad (8)$$

where

$$u_i(x) = -\frac{\alpha}{Y_i} \left[x^2 - 2 \left(X_i + \frac{Y_i}{2\alpha X_i} \right) x \right].$$

It is noted that the function u_i increases in the region $[0, X_i + Y_i/(2\alpha X_i)]$ but decreases in the region $[X_i + Y_i/(2\alpha X_i), +\infty)$. This provides the key mechanism that the maximum transmission rate of user i chosen by the best scheduler is always bounded by $X_i + Y_i/(2\alpha X_i)$. Therefore, the shared resource (e.g., time or frequency) can be allocated to other users with lower transmission rate. The parameter α reflects the impact of rate variance on the communication services. A larger value of α results in more sensitive to rate oscillation. Correspondingly, the selected transmission rate is restricted within a small region.

V. ASYMPTOTIC ANALYSIS OF GGSA

We study the dynamics of average rates and rate variances under the GGSA algorithm when the step size μ approaches zero. For this purpose we define a fluid sample path (FSP) under GGSA. For a fixed step size μ , let $\mathbf{X}^\mu(k)$ and $\mathbf{Y}^\mu(k)$ be the realizations of an average rate and rate variance vector process. For a given channel state realization $\mathbf{g}^\mu = (\mathbf{g}^\mu(k), k = 0, 1, 2, \dots)$ and a fixed initial state $\mathbf{X}^\mu(0)$ and $\mathbf{Y}^\mu(0)$, $\mathbf{X}^\mu(k)$ and $\mathbf{Y}^\mu(k)$ are uniquely determined. Consider the following continuous time process $(\mathbf{x}^\mu(t), \mathbf{y}^\mu(t))$,

$$\begin{aligned} \mathbf{x}^\mu(t) &:= \mathbf{X}^\mu(\lfloor t/\mu \rfloor), \quad t \geq 0 \\ \mathbf{y}^\mu(t) &:= \mathbf{Y}^\mu(\lfloor t/\mu \rfloor), \quad t \geq 0, \end{aligned}$$

where $\lfloor x \rfloor := \sup\{i \in \mathbb{Z} : i \leq x\}$.

A pair of vector-functions $(\mathbf{x} = (\mathbf{x}(t), t \geq 0), \mathbf{y} = (\mathbf{y}(t), t \geq 0))$ is called a fluid sample path (FSP) [4], if there

exist a sequence of positive values of μ such that as $\mu \rightarrow 0$, a sequence of sample path $(\mathbf{x}^\mu(t), \mathbf{y}^\mu(t))$ satisfy

$$(\mathbf{x}^\mu(t), \mathbf{y}^\mu(t)) \rightarrow (\mathbf{x}(t), \mathbf{y}(t)) \quad u. o. c.$$

Here, *u. o. c.* means uniform on compact sets convergence of function.

Some properties of FSPs are described in the following lemmas based on which we establish Theorem 1.

Lemma 1: For any FSP, $\mathbf{x}$ and $\mathbf{y}$ are lipschitz continuous in $[0, \infty)$.

Proof: Recall the update equations (2) and (3) in Section IV. Applying these equations iteratively, we have

$$\begin{aligned} X_i^\mu(k) &= \sum_{j=1}^{k-1} \mu(1-\mu)^j r_i^\mu(k-j) + (1-\mu)^k X_i^\mu(0) \\ &\leq \max\{\bar{r}, X_i^\mu(0)\}, \quad \forall k \geq 0 \\ Y_i^\mu(k) &= \sum_{j=1}^{k-1} \mu(1-\mu)^j [r_i^\mu(k-j) - X_i^\mu(k-j)]^2 \\ &\quad + (1-\mu)^k Y_i^\mu(0) \\ &\leq \max\{\max\{\bar{r}^2, [x_i^\mu(0)]^2\} + \bar{r}^2 Y_i^\mu(0)\}, \quad \forall k \geq 0, \end{aligned}$$

where $\bar{r}$ is the upper bound of the rate $r_i^\mu(k)$, $\forall i, k, \mu > 0$. Then, we have

$$\begin{aligned} |X_i^\mu(k) - X_i^\mu(k-1)| &= |\mu r_i^\mu(k) - \mu X_i^\mu(k-1)| \\ &\leq \mu[r_i^\mu(k) + X_i^\mu(k-1)] \\ &\leq \mu[2\bar{r} + X_i^\mu(0)]. \end{aligned} \quad (9)$$

Similarly, we can prove

$$\begin{aligned} |Y_i^\mu(k) - Y_i^\mu(k-1)| &\leq \mu[2\max\{\bar{r}^2, [x_i^\mu(0)]^2\} + 2\bar{r}^2 + Y_i^\mu(0)]. \end{aligned} \quad (10)$$

The inequalities (9) and (10) imply that $(\mathbf{x}, \mathbf{y})$ is Lipschitz continuous in $[0, \infty)$. $\blacksquare$

Under the assumption that the chosen rate vector $\mathbf{r}(k) = \mathbf{F}(\mathbf{X}(k), \mathbf{Y}(k), \mathbf{g}(k))$ is continuous for almost all $(\mathbf{X}(k), \mathbf{Y}(k), \mathbf{g}(k))$, we have the Lemma 2.

Lemma 2: The family of FSPs $(\mathbf{x}, \mathbf{y})$ satisfies the ordinary differential equations (ODE),

$$\dot{\mathbf{x}} = \bar{\mathbf{F}}(\mathbf{x}, \mathbf{y}) - \mathbf{x} \quad (11)$$

$$\dot{\mathbf{y}} = \bar{\mathbf{F}}^2(\mathbf{x}, \mathbf{y}) - 2\mathbf{x} \times \bar{\mathbf{F}}(\mathbf{x}, \mathbf{y}) - \mathbf{x} \times \mathbf{x} - \mathbf{y} \quad (12)$$

where

$$\bar{\mathbf{F}}(\mathbf{x}, \mathbf{y}) = \int_{\mathcal{G}} \mathbf{F}(\mathbf{x}(s), \mathbf{x}(s), \mathbf{g}) \gamma(d\mathbf{g}) \quad (13)$$

$$\begin{aligned} \bar{\mathbf{F}}^2(\mathbf{x}, \mathbf{y}) &= \int_{\mathcal{G}} \mathbf{F}(\mathbf{x}(s), \mathbf{x}(s), \mathbf{g}) \\ &\quad \times \mathbf{F}(\mathbf{x}(s), \mathbf{x}(s), \mathbf{g}) \gamma(d\mathbf{g}). \end{aligned} \quad (14)$$

Proof: Based on Lemma 1 of [8], we can obtain

$$\begin{aligned}
\mathbf{x}(t) &= \mathbf{x}(0) + \int_{\mathcal{G} \times [0,t]} [\mathbf{F}(\mathbf{x}(s), \mathbf{x}(s), \mathbf{g}) - \mathbf{x}(s)] \gamma(d\mathbf{g}) ds \\
&= \mathbf{x}(0) + \int_0^t [\bar{\mathbf{F}}(\mathbf{x}(s), \mathbf{y}(s)) - \mathbf{x}(s)] ds \\
\mathbf{y}(t) &= \mathbf{y}(0) + \int_{\mathcal{G} \times [0,t]} \{ [\mathbf{F}(\mathbf{x}(s), \mathbf{x}(s), \mathbf{g}) - \mathbf{x}(s)] \times \\
&\quad [\mathbf{F}(\mathbf{x}(s), \mathbf{x}(s), \mathbf{g}) - \mathbf{x}(s)] - \mathbf{y} \} \gamma(d\mathbf{g}) ds \\
&= \mathbf{y}(0) + \int_0^t [\bar{\mathbf{F}}^2(\mathbf{x}(s), \mathbf{y}(s)) - 2\mathbf{x}(s) \times \\
&\quad \bar{\mathbf{F}}(\mathbf{x}(s), \mathbf{y}(s)) + \mathbf{x}(s) \times \mathbf{x}(s) - \mathbf{y}(s)] ds
\end{aligned}$$

Since $(\mathbf{x}, \mathbf{y})$ of an FSP is lipschitz continuous, its derivatives exist at almost all points and, therefore, for every regular point $t \geq 0$, $(\mathbf{x}, \mathbf{y})$ satisfies the ODEs (11) and (12). ■

Lemma 3: Suppose $(\mathbf{x}, \mathbf{y})$ is a stationary FSP, namely

$$\mathbf{x}(t) \equiv \mathbf{x}^*, \mathbf{y}(t) \equiv \mathbf{y}^*, \quad \forall t \geq 0. \quad (15)$$

Then $(\mathbf{x}, \mathbf{y})$ is a solution to the problem in (1).

Proof: For simplicity, we define

$$\bar{\mathbf{H}}(\mathbf{x}, \mathbf{y}) := \bar{\mathbf{F}}^2(\mathbf{x}, \mathbf{y}) - 2\mathbf{x} \times \bar{\mathbf{F}}(\mathbf{x}, \mathbf{y}) - \mathbf{x} \times \mathbf{x}.$$

When $\mathbf{x}(t) \equiv \mathbf{x}^*$, we have $\bar{\mathbf{F}}(\mathbf{x}, \mathbf{y}) = \mathbf{x}$ from (11). It follows that

$$\bar{\mathbf{H}}(\mathbf{x}, \mathbf{y}) = \bar{\mathbf{F}}^2(\mathbf{x}, \mathbf{y}) - [\bar{\mathbf{F}}(\mathbf{x}, \mathbf{y})] \times [\bar{\mathbf{F}}(\mathbf{x}, \mathbf{y})],$$

By (5), (13) and (14), we have

$$\begin{aligned}
&(\bar{\mathbf{F}}(\mathbf{x}, \mathbf{y}), \bar{\mathbf{H}}(\mathbf{x}, \mathbf{y})) \\
&= \arg \max_{(\mathbf{v}, \mathbf{u}) \in \mathcal{V}} \nabla_{\mathbf{m}} U(\mathbf{x}, \mathbf{y})^T \cdot \mathbf{v} + \nabla_{\sigma^2} U(\mathbf{x}, \mathbf{y})^T \cdot \mathbf{u}.
\end{aligned}$$

When (15) holds, we have

$$(\bar{\mathbf{F}}(\mathbf{x}, \mathbf{y}), \bar{\mathbf{H}}(\mathbf{x}, \mathbf{y})) = (\mathbf{x}^*, \mathbf{y}^*),$$

This implies that $(\mathbf{x}^*, \mathbf{y}^*)$ is a solution to the following problem

$$\max_{(\mathbf{v}, \mathbf{u}) \in \mathcal{V}} \nabla_{\mathbf{m}} U(\mathbf{x}^*, \mathbf{y}^*)^T \cdot \mathbf{v} + \nabla_{\sigma^2} U(\mathbf{x}^*, \mathbf{y}^*)^T \cdot \mathbf{u},$$

which means that $(\mathbf{x}^*, \mathbf{y}^*)$ is a maximal point of the set $\mathcal{V}$. In addition, $(\nabla_{\mathbf{m}} U(\mathbf{x}^*, \mathbf{y}^*), \nabla_{\sigma^2} U(\mathbf{x}^*, \mathbf{y}^*))$ is normal to the set $\mathcal{V}$. Therefore, $(\mathbf{x}, \mathbf{y})$ is a solution to the problem in (1). ■

It follows from Lemma 1, 2 and 3 that we have Theorem 1.

Theorem 1: Suppose FSP, $(\mathbf{x}, \mathbf{y})$ satisfies $(\mathbf{x}(t), \mathbf{y}(t)) \rightarrow (\mathbf{x}^*, \mathbf{y}^*)$ as $t \rightarrow \infty$, then $(\mathbf{x}^*, \mathbf{y}^*)$ is a solution to the problem in (1).

Proof: The proof is analogous to [6, Proof of Theorem 1]. ■

Theorem 1 reveals that GGSA is asymptotically optimal under the condition that the transmission rate vector converges to a fixed vector as time goes into infinity. Whether the average rate and rate variance vectors will converge is, however, not known. We shall study the convergence properties using numerical results in the next section.

VI. NUMERICAL RESULTS

In this section, we present numerical results to illustrate the performance of the GGSA.

We concentrate on a time division multiplexing (TDM) system where each time frame can be accessed by all the N users in an adaptive time-sharing fashion. The channels of different users are statistically independent and follows the exponential distribution, denoted by

$$f_i(g) = \begin{cases} \frac{1}{\bar{g}_i} \exp\left(-\frac{g}{\bar{g}_i}\right) & g \geq 0 \\ 0 & \text{otherwise} \end{cases}, \quad (16)$$

where $\bar{g}_i$ is the average channel gain of user i . Let p_i denote the transmit power allocated to or from user i . The achievable transmission rate of user i in the absence of other users can be expressed as

$$c_i = \log_2 \left(1 + \frac{p_i \bar{g}_i}{\beta N_0} \right), \quad (17)$$

where N_0 is the noise power spectral density, and β is the SNR gap [9]. When uncoded QAM constellation is used, β is a constant related to a given bit-error-rate (BER) constraint, given by $\beta = -\ln(5 \cdot \text{BER})/1.5$. Let $\rho(\mathbf{g}) = (\rho_1, \rho_2, \dots, \rho_N)$ denote the time-sharing adaptation policy with respect to the network channel gain $\mathbf{g}$, where ρ_i represents the fraction of the frame duration allocated to user i . The actual transmission rate of user i in each time frame, r_i , can therefore be written as

$$r_i = \rho_i c_i = \rho_i \log_2 \left(1 + \frac{p_i \bar{g}_i}{\beta N_0} \right). \quad (18)$$

A wireless network with four symmetric users is considered. The target bit-error-rate is set to be 10^{-5} , which results in a SNR gap of 8.2dB. The average received SNR is varied from 0 dB to 35 dB. We assume that the feedback channel is error free and has no delay. Each run comprises 50000 time frames. The PF-GGSA scheduler defined in (8) is used for simplicity.

Fig. 1 compares the average rates and the rate variances by varying the parameter α in the set $\{0, 0.05, 0.1, 0.2\}$. It can be shown that the GGSA has a flexible balance between the average rate and the rate oscillation through adjusting the parameter α . When $\alpha = 0$, the GGSA is equivalent to the PF algorithm in [2], and it can obtain the maximum average transmission rate (in the symmetric case) while it suffers a high rate oscillation.

Figs. 2 and 3 present the trajectories of the GGSA with different starting points and step sizes, μ . When the step size is sufficiently small, the trajectories with different starting points and different step sizes always converge to the same optimal point. The speed of convergence depends on the step size. A small step size corresponds to a slow speed of convergence. When the step size is not small enough, the trajectories oscillate around the optimal points.

VII. CONCLUSIONS

We proposed a GGSA for wireless networks supporting variable rate transmission. This algorithm can maximize a utility function that increases with average transmission rate but

Fig. 1. Average rate and rate variance of user one versus average received SNR

Fig. 2. Trajectories of the average rate of user one with different starting points and step sizes

decreases with rate variance. The utility function is capable of facilitating the resource allocation with flexible combinations of average rate and rate oscillation, which are both important performance metrics for variable rate transmission. The GGSA suggests that the best scheduler should maximize the sum of concave functions of instantaneous transmission rate at each time frame. We studied the asymptotic performance of the algorithm, and showed that if the transmission rates converge, then the convergence point maximizes the desired utility function. Numerical results show a good convergence performance of the algorithm in time division multiplexing systems when the step size is small enough.

Fig. 3. Trajectories of the rate variance of user one with different starting points and step sizes

REFERENCES

- [1] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 1995.
- [2] A. Jalali, R. Padovani, and R. Pankaj, "Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system," in *Proc. Veh. Technol. (VTC)*, 2000.
- [3] R. Agrawal and V. Subramanian, "Optimality of certain channel aware scheduling policies," in *Proc. 40th Annual Allerton Conf. Comm., Control, Comput.*, Monticello, IL, 2002, pp. 1532–1541.
- [4] A. L. Stolyar, "On the asymptotic optimality of the gradient scheduling algorithm for multiuser throughput allocation," *Operation Research*, vol. 53, no. 1, pp. 12–25, Jan. 2005.
- [5] H. Viswanathan, S. Venkatesan, and H. Huang, "Downlink capacity evaluation of cellular networks with known-interference cancellation," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 802–811, June 2003.
- [6] M. Andrews, L. Qian, and A. Stolyar, "Optimal utility based multi-user throughput allocation subject to throughput constraints," in *Proc. IEEE INFOCOM '05*, 2005.
- [7] X. Zhang, M. Tao, and C. S. Ng, "Time sharing policy in wireless networks for variable rate transmission," in *Proc. of IEEE ICC'07*, Glasgow, UK, June 2007.
- [8] J. A. Bucklew, T. Kurtz, and W. A. Sethares, "Weak convergence and local stability properties of fixed step size recursive algorithms," *IEEE Trans. on Infor. Theory*, vol. 39, no. 3, pp. 966–978, May 1993.
- [9] X. Qiu and K. Chawla, "On the performance of adaptive modulation in cellular systems," *IEEE Trans. on Commun.*, vol. 47, no. 6, pp. 884–895, June 1999.