Optimality of Certain Channel Aware Scheduling Policies

Rajeev Agrawal and Vijay Subramanian

Mathematics of Communication Networks
Motorola Inc., Arlington Heights, IL 60004.
Email: {ragrawa1,vsubram2}@email.mot.com

Abstract

With an abstraction of serving rate-adaptive sources on a broadcast-type wireless channel as a utility maximization problem, it is shown how one can design many intuitive online scheduling policies based upon the feedback that one obtains at the scheduler. Using a stochastic approximation argument it is then shown that the constructed algorithms converge to optimal solutions of the utility maximization problem over different sets which critically depend on the quality of the feedback information.

1 Introduction

The emergence of the third generation cellular technologies over the last decade has generated a flurry of activity in wireless data. A principal component for enabling wireless data is intelligent scheduling amongst different traffic streams. Exploiting the time-varying nature of wireless channels to acheive system efficiency whilst still maintaining a notion of fairness there have been many different proposals of *opportunitistic* schedulers [6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18] for the broadcast-type channel that exists on the downlink direction of most cellular systems. A key element of every scheduler is to transmit to an appropriate user at a time with a bias to users with good channel conditions so as to not violate the fairness requirements. Similar schedulers can be proposed in the uplink direction as well but since the channel is of a multiple-access nature it is necessary to assume the existence of a centralised scheduler that gathers queue-length information and sends out scheduling commands to the different nodes.

In [11] the proportionally fair algorithm scheduling algorithm was proposed using a $\log(\cdot)$ utility function for different streams. Using a bigger class of utility functions [6, 7] showed how one can easily choose a scheduler with an efficiency-fairness tradeoff in between the two extreme cases, namely, a channel-unaware scheduler and a scheduler that always serves the best user at any given time. The authors of [8, 9] investigate a scheduling algorithm that maximises the minimum weighted throughput of different users and show that one can devise adaptive strategies that asymptotically converge to the optimal solution in a stochastic setting. In [10] the authors consider different fairness criteria and discuss optimal scheduling algorithms. They also discuss some properties of the rate region achievable by a general class of *opportunistic* scheduling algorithms. In [18] optimality properties of the proportionally-fair algorithm are discussed.

The main contributions of this paper are as follows. First, we generalize the scheduling algorithms in the above references to general channels in order to accommodate cases wherein it is better to transmit to multiple users at the same time as is the case with multiple transmit antennae [17]. Next, using a stochastic approximation approach we show that asymptotically it is possible to analyse the performance of the algorithm by means of an ordinary differential equation. Using properties of the differential equation we then show that the algorithms converge to the optimal solution of a related optimization problem. Revisiting schemes that transmit only to user at a time in the above general context we draw a few conclusions on multiuser diversity.

2 The Model

Consider a wireless communication system with d users. The channel conditions are time varying and captured by a stochastic *channel state* $\eta_k \in \mathcal{S}$ at time k, where $\mathcal{S}$ is the channel state space. We assume that $\mathcal{S}$ is a Polish space. Associated with each state $\eta \in \mathcal{S}$ we have a rate-region $\mathcal{R}(\eta) \subset \Re^d_+$. Thus when the channel is in the state η, the users may transmit at any vector of rates $v = (v_1, \ldots, v_d) \in \mathcal{R}(\eta)$. We will assume throughout this paper that $\mathcal{R}(\eta) \subset K \subset \Re^d_+$, $\forall \eta \in \mathcal{S}$ for some compact set K and that the process $\{\eta_k\}$ is ergodic with the stationary distribution γ. For simplicity we shall assume that $\mathcal{R}(\eta)$ is convex, coordinate convex, and closed for every η. Then the steady-state rate region is given by

$$\bar{\mathcal{R}} := \left\{ w \in \Re^k_+ \ : \exists v(\eta) \in \mathcal{R}(\eta) \ \forall \ \eta \in \mathcal{S} \text{ such that } w = \int_{\mathcal{S}} v(\eta)\gamma(d\eta) \right\}. \tag{1}$$

It is easily verified that $\bar{\mathcal{R}}$ is convex, coordinate convex, and compact. Note that $\bar{\mathcal{R}}$ is precisely the set of all achievable steady-state long-term empirical throughput vectors w. See [13] for examples of some specific wireless communication systems including TDMA and CDMA cellular systems and adhoc networks that fit the above mathematical model.

We assume that the d users are rate-adaptive and need to share the channel described above, fairly and efficiently. The problem that we would like to solve can be translated into the following utility maximization problem:

$$\sup_{w \in \bar{\mathcal{R}}} \sum_{i=1}^{d} U_i\left(w_i\right) \triangleq U(w).$$

For each i we assume that $U_i(\cdot)$ is an increasing, strictly concave, and continuously differentiable utility function on $\Re_+$. From the above observations on $\bar{\mathcal{R}}$, it follows that a maximizer exists and that it is unique.

3 A Gradient Based Scheduling Algorithm

Let $V_k \in \mathcal{R}(\eta_k)$ be the rate selected at time k. Define W_k to be the empirical throughput as follows:

$$\begin{aligned} W_0 &= 0 \\ W_{k+1} &= W_k + \mu_k(V_k - W_k), \quad k > 0. \end{aligned} \tag{2}$$

Broadly speaking we consider two case cases, viz., one, $\mu_k = 1/k, k > 0$ for time-average throughput, and $\mu_k = \mu > 0, \forall k > 0$ for the constant step-size case. In the constant step size we make the additional distinction that the step size μ be small. In all cases we are interested in optimizing $U(W_k)$ as $k \to +\infty$.

To do so, we consider a myopic view of the optimization problem; we optimize $U(W_{k+1})$ by choosing $V_k \in \mathcal{R}(\eta_k)$ appropriately given that $V_0, \ldots, V_{k-1}$ have already been chosen. Thus, we are interested in finding what the next best step is given whatever action was taken in the past. Note that,

$$\begin{aligned} U(W_{k+1}) - U(W_k) &= U(W_k + \mu_k(V_k - W_k)) - U(W_k) \\ &\approx \mu_k \nabla U(W_k)^T (V_k - W_k), \end{aligned}$$

where the last relation holds for $\mu_k \ll 1$. Thus, for small enough μ_k the best choice given the past decisions is to choose a point V_k in the capacity region that satisfies

$$V_k = \arg \max_{v \in \mathcal{R}(\eta_k)} \nabla U(W_k)^T v. \tag{3}$$

Remark We may easily make the following observations:

1. This leads to a gradient-based *scheduling* algorithm.
2. With the convex rate region assumption for each state this is an easy problem to solve.
3. In the case that the region for every state is a simplex we obtain a TDM-type algorithm where only one user is allowed to transmit at a time.

We will also consider a number of other *scheduling* algorithms for choosing the rate vector $V_k = F(W_k, \eta_k) \in \mathcal{R}(\eta_k)$ at time k based on the current throughput W_k and possibly knowledge of the current channel state η_k.

4 Analysis of the constant step-size case

We are interested in studying the behavior of this algorithm (2) for constant step-size case, i.e., $\mu_k = \mu, \forall k > 0$, with a number of different *scheduling* algorithms for choosing the rate vector $V_k = F(W_k, \eta_k) \in \mathcal{R}(\eta_k)$ for small μ. For this pupose define the continuous time process

$$W_\mu(t) := W_{[t/\mu]}, \quad t \geq 0, \quad \text{where } [x] := \sup\{i \in Z : i \leq x\}.$$

Also, define the occupation measure

$$\Gamma_\mu(C \times [0,t]) := \mu \sum_{k=1}^{[t/\mu]} I_C(\eta_k) \text{ for } C \subseteq S,\ C \text{ Borel.}$$

Under the ergodicity condition on the channel state process, this converges to the measure

$$\Gamma(d\eta \times ds) = \gamma(d\eta) ds,$$

where γ is the stationary distribution of $\{\eta_k\}$. We may then express

$$\begin{aligned} W_\mu(t) &= W_\mu(0) + \mu \int_0^{\mu[t/\mu]} (F(W_\mu(s), \eta_{[s/\mu]}) - W_\mu(s)) ds \\ &= W_\mu(0) + \int_{\mathcal{S} \times [0,t]} (F(W_\mu(s), \eta) - W_\mu(s)) \Gamma_\mu(d\eta \times ds) \end{aligned} \tag{4}$$

We shall assume that the initial rate vector $W_\mu(0) \to w_0$ in probability as $\mu \to 0$. Due to the common compact bound on the rate regions, it follows that the family of processes $\{W_\mu\}$ (and consequently $\{W_\mu, \Gamma_\mu\}$) is relatively compact. We shall make the following continuity assumpion on F.

A1. For each $w \in \Re_+^d$ and $C_w := \{\eta : F \text{ is continuous at } (w, \eta)\}$, $\gamma(C_w) = 1$.

Then we have the following theorem based on Lemma 1 Part (c) of [2].

Theorem 4.1 *Under assumption A1, it follows that any limit point* (W, Γ) *of* $\{W_\mu, \Gamma_\mu\}$ *satisfies*

$$\begin{aligned} W(t) &= w_0 + \int_{\mathcal{S}\times[0,t]} (F(W(s), \eta) - W(s))\gamma(d\eta)ds \qquad (5) \\ &= w_0 + \int_0^t (\bar{F}(W(s)) - W(s))ds. \qquad (6) \end{aligned}$$

where

$$\bar{F}(w) := \int_{\mathcal{S}} F(w, \eta)\gamma(d\eta)$$

In case $\bar{F}$ *is continuous,* W *satisfies the ODE*

$$\dot{W} = \bar{F}(W) - W. \qquad (7)$$

Extensions: Assuming that $\bar{F}(w)$ is continuously differentiable and mixing conditions hold for the state process η_k, we can derive a central-limit theorem type convergence result[19] for the error process

$$\Xi_\mu(t) = \frac{1}{\sqrt{\mu}}(W_\mu(t) - W(t)) \qquad (8)$$

based upon [2, Thm 2., p. 969]. Define

$$L_\mu(t) = \sqrt{\mu}\sum_{k=1}^{[t/\mu]} \left(F(W(k\mu), \eta_k) - \bar{F}(W(k\mu))\right).$$

Assume **C.1** that $L_\mu \Longrightarrow L$, where L is a zero-mean Brownian motion. Mixing conditions on η_k will imply this. Additionally assuming **C.2** that $F(w, \eta)$ is continuously differentiable in η with bounded derivative $\partial_w F(w, \eta)$. We then have

Theorem 4.2 *Assume* **C.1-C.2** *and that the solution to (7) exists for all* $t \geq 0$*, and that* $\Xi_\mu(0) \to \xi_0$ *in probability. Then* $\Xi_\mu \Longrightarrow \Xi$ *satisfying*

$$\Xi(t) = \xi_0 + L(t) + \int_0^t \partial\bar{F}(W(s))\Xi(s)ds. \qquad (9)$$

5 Some Gradient-type Scheduling Algorithms and their F, $\bar{F}$

We now consider some specific choices of scheduling algorithms $V_k = F(W_k, \eta_k)$. All cases are based on a gradient type algorithm. The first case we consider is where, the scheduler knows the exact channel state η_k.

5.1 Complete knowledge of the current channel state

This is the ideal case of complete knowledge where we know the current state η_k and optimize for the current rate V_k over the set $\mathcal{R}(\eta)$. Thus,

$$F(w, \eta) = \arg\max_{u\in\mathcal{R}(\eta)} \nabla U(w)^T u$$

We define a compact and convex set $\mathcal{Q} \subset \Re^d_+$ to be *strictly*-convex if for all $a \geq 0, \sum_{i=1}^d a_i = 1$ there is a unique maximiser of $a^T u$ in $\mathcal{Q}$.

Proposition 5.1 *Under strict-convexity of $\bar{\mathcal{R}}$*

$$\bar{F}(w) = \arg\max_{u\in\bar{\mathcal{R}}} \nabla U(w)^T u. \quad (10)$$

The proposition follows from the lemma below.

Lemma 5.1 *Under strict-covexity of $\bar{\mathcal{R}}$*

$$\int_{\mathcal{S}} \arg\max_{u\in\mathcal{R}(\eta)} a^T u \gamma(d\eta) = \arg\max_{u\in\bar{\mathcal{R}}} a^T u$$

Proof: Let $u^*(\eta)$ be a maximizer in the LHS. Let $u^* := \int_{\mathcal{S}} u^*(\eta)\gamma(d\eta)$. Clearly $u^* \in \bar{\mathcal{R}}$ and thus $a^T u^* \leq \max_{u\in\bar{\mathcal{R}}} a^T u$. Now for any $u = \int_{\mathcal{S}} u(\eta)\gamma(d\eta) \in \bar{\mathcal{R}}$ with $u(\eta) \in \mathcal{R}(\eta)$, $a^T u^* = \int_{\mathcal{S}} a^T u^*(\eta)\gamma(d\eta) \geq \int_{\mathcal{S}} a^T u(\eta)\gamma(d\eta) = a^T u$. Thus, $a^T u^* \geq \max_{u\in\bar{\mathcal{R}}} a^T u$. Consequently, $a^T u^* = \max_{u\in\bar{\mathcal{R}}} a^T u$, and hence, $u^* = \arg\max_{u\in\bar{\mathcal{R}}} a^T u$, since there is a unique maximizer of the RHS by the strict-convexity assumption.

5.2 No knowledge of the current channel state

In this section we devise a scheduling algorithm that is based just on knowledge of the steady state rate region and not the current rate. Let $\bar{R}_i$ denote the intercept of the rate region $\bar{\mathcal{R}}$ along the i-th dimension. This is simply the throughput user i would get if it was the only user scheduled for all times. At time k the scheduling policy picks the user $i^*_k = i^*(W_k)$ for transmission based on only W_k as follows:

$$i^*_k = i^*(W_k) := \arg\max_i \dot{U}_i(W_{k,i})\bar{R}_i.$$

The resulting rate vector $U_k = F(W_k, \eta_k)$ is given by

$$F(w,\eta) = R_{i^*(w)}(\eta) e^{i^*(w)}$$

where $R_i(\eta)$ is the intercept along the i-th dimension of $\mathcal{R}(\eta)$ and e^i is the unit vector in the i-th direction. The following is easily verified.

Proposition 5.2

$$\bar{F}(w) = \int_{\mathcal{S}} F(w,\eta)\gamma(d\eta) = \bar{R}_{i^*(w)} e^{i^*(w)} = \arg\max_{u\in\bar{\mathcal{R}}^S} \nabla U(w)^T u. \quad (11)$$

where $\bar{\mathcal{R}}^S$ is the largest simplex inscribed in $\bar{\mathcal{R}}$.

Note that $F(w,\eta)$ defined here will not satisfy **A1**.

5.3 TDM type schedulers with knowledge of channel state

In this subsection we concentrate on time division multiplexing (TDM) type algorithms that only serve one user at a time but still use the time-varying capacity for each user. At time k the scheduling policy picks the user $i^*_k = i^*(W_k, \eta_k)$ for transmission based on both W_k and η_k as follows:

$$i^*_k = i^*(W_k, \eta_k) := \arg\max_i \dot{U}_i(W_{k,i}) R_i(\eta_k).$$

where $R_i(\eta)$ is the intercept along the i-th dimension of $\mathcal{R}(\eta)$ as defined before. The resulting rate vector $U_k = F(W_k, \eta_k)$ is given by

$$F(w,\eta) = R_{i^*(w)}(\eta) e^{i^*(w,\eta)} = \arg\max_{u\in\mathcal{R}^S(\eta)} \nabla U(w)^T u. \quad (12)$$

In other words, the capacity region for every channel state $\eta \in \mathcal{S}$ is restricted to the largest simplex inscribed in $\mathcal{R}(\eta)$ which we refer to by $\bar{\mathcal{R}}^S(\eta)$. This limited view of the rate region results if we place the restriction that the users send only the rates that they can achieve as feedback. The average rate region which is defined in a manner similar to that of $\bar{\mathcal{R}}$ in (1), is denoted by $\bar{\mathcal{R}^S}$. Note that this is in general not equal to $\bar{\mathcal{R}}^S$, the largest simplex inscribed in $\bar{\mathcal{R}}$. In general, $\bar{\mathcal{R}^S} \supset \bar{\mathcal{R}}^S$. As in the case of complete information, the proposition below follows from Lemma 5.1.

Proposition 5.3 *Under assumption that the corresponding steady state capacity region $\bar{\mathcal{R}^S}$ is strictly convex,*

$$\bar{F}(w) = \arg\max_{u \in \bar{\mathcal{R}^S}} \nabla U(w)^T u. \tag{13}$$

5.4 TDM type schedulers with imperfect knowledge of channel state

Note that schedulers considered in Sections 5.2, 5.3 were both TDM type. In this subsection we generalize these two cases to TDM type algorithms that only serve one user at a time and which use (possibly) imperfect knowledge of the time-varying rate-region. Let $R_i(\eta)$ be the intercept along the i-th dimension of $\mathcal{R}(\eta)$ as defined before. This is the maximum rate user i may transmit at when the channel is in state η. Let the (possibly) imperfect knowledge of the maximum rate user i may be served at when the channel state is η be denoted by $\hat{R}_i(\eta)$. In the case of perfect knowledge of the maximum rates, $\hat{R}_i(\eta) = R_i(\eta)$ but, in general, $\hat{R}_i(\eta) \neq R_i(\eta)$ At time k the scheduling policy picks the user $i_k^* = i^*(W_k, \eta_k)$ for transmission based on both W_k and η_k as follows:

$$i_k^* = i^*(W_k, \eta_k) := \arg\max_i \dot{U}_i(W_{k,i}) \hat{R}_i(\eta_k).$$

The resulting rate vector $U_k = F(W_k, \eta_k)$ is given by

$$F(w, \eta) = R_{i^*(w)}(\eta) e^{i^*(w,\eta)} = \sum_{i=1}^{d} R_i(\eta) e^i 1_{\{\dot{U}_i(w_i)\hat{R}_i(\eta) > \dot{U}_i(w_i)\hat{R}_i(\eta) \forall j \neq i\}}. \tag{14}$$

In case all of the above indicators are 0, i.e., there are multiple maximizers $i^*(w, \eta)$, we assume that one of the above indicators will be set to one and others to 0.

This model actually captures a number of interesting cases including the following:

1. Knowledge only of the steady state maximum rates, i.e. $\hat{R}_i(\eta) = \bar{R}_i$, as described in the previous section 5.2.

2. Knowledge only of the steady state maximum rates with *dithering*. Since the model with knowledge only of the steady state rates leads to a F that violates assumption A1, dithering may be deliberately introduced to smooth it out. The channel state η in this model may include some auxilliary random variable $\epsilon(\eta) = (\epsilon_1, \ldots, \epsilon_d)$ used for dithering. In this case $\hat{R}_i(\eta) = \bar{R}_i + \epsilon(\eta)$. Assuming that under the stationary distribution γ, $\epsilon(\eta)$ is uniform on $[-\delta, \delta]^d$ and independent of $R_i(\eta)$, it may be shown that F satisfies assumption A1 and that $\bar{F}(w)$ is continuous in w. Moreover as $\delta \to 0$, F and $\bar{F}$ converge to the corresponding functions for the case without dithering (at their points of continuity).

3. Predicted maximum rate vectors, i.e, $\hat{R}_i(\eta)$ is a prediction of $R_i(\eta)$. For instance, we may not know $R_i(\eta_k)$, but instead estimate it by using the IIR filter:

   $$\hat{R}_{k+1,i} = \alpha \hat{R}_{k,i} + (1 - \alpha) R_i(\eta_k)$$

   We may "fold in" $\hat{R}_{k,i}$ into the channel state process η_k.

4. TDM type scheduler with complete knowledge of the current channel state, i.e. $\hat{R}_i(\eta) = R_i(\eta)$, as described in the previous section 5.3

6 Solution of the ODE

Next we investigate the solution of the differential equation (7) reproduced below.

$$\dot{W} = \bar{F}(W) - W. \tag{15}$$

Note that for several of the algorithms presented in the previous section, $\bar{F}$ takes the form

$$\bar{F}(w) = \arg\max_{u \in \mathcal{Q}} \nabla U(w)^T u,$$

for some convex, coordinate convex, and compact subset $\mathcal{Q} \subset \bar{\mathcal{R}}$. Because of the strict concavity of the utility function U and the above, it is easy to verify the following lemma (see [3, Prop 2.1.1 p. 175, Prop. A8 p. 540]).

Lemma 6.1 $x^* = \arg\max_{x \in \mathcal{Q}} U(x)$ *if and only if*

$$\nabla U(x^*)^T (x - x^*) \le 0 \quad \forall x \in \mathcal{Q}. \tag{16}$$

Let $w^* := \arg\max_{x \in \mathcal{Q}} U(x)$ denote the optimal throughput.

Proposition 6.1 *Under the assumption that $\mathcal{Q} \subseteq \bar{\mathcal{R}}$ is strictly convex, it follows that w^* is the unique equilibrium point of the differential equation (7) and $W(t) \to w^*$ as $t \to \infty$ starting with any initial state $W(0) = w_0 \in \mathcal{Q}$.*

Proof: w' is an equilibrium point of (7) if and only if $w' \in \mathcal{Q}$ and

$$\begin{aligned} \bar{F}(w') = w' \quad &\Leftrightarrow \quad \arg\max_{w \in \mathcal{Q}} \nabla U(w')^T w = w' \\ \Leftrightarrow \nabla U(w')^T w \le \nabla U(w')^T w', \quad \forall x \in \mathcal{Q} \quad &\Leftrightarrow \quad \nabla U(w')^T (w - w') \le 0 \\ \Leftrightarrow w' \quad &= \quad w^*. \end{aligned}$$

To show the convergence of the solution of the differential equation to this equilibrim point, we use the utility function U as a natural choice for a Lyapunov function.

$$\begin{aligned} \frac{d}{dt} U(W(t)) &= \nabla U(W(t))^T \dot{W}(t) = \nabla U(W(t))^T (\bar{F}(W(t)) - W(t)) \\ &= \nabla U(W(t))^T (\arg\max_{w \in \mathcal{Q}} \nabla U(W(t))^T w - W(t)) \\ &= \max_{w \in \mathcal{Q}} \nabla U(W(t))^T w - \nabla U(W(t))^T W(t)) \ge 0 \quad \text{for } W(t) \in \mathcal{Q} \end{aligned}$$

with equality above iff $W(t) = w^*$. Thus $U(W(t))$ is strictly increasing with t unless $W(t) = w^*$. The function U is thus a Lyapunov function for the differential equation (7) and the proposition follows (see [5, Theorem 5.1(58)]).

Remark: Thus, the equilibrium point is now the unique sum utility maximiser restricted to the subset $\mathcal{Q} \subset \bar{\mathcal{R}}$. From this it is quite clear that knowledge of the current rates, and even better, the current rate-region, leads to a better solution than what can be obtained merely by knowing the average rates.

7 Analysis of the decreasing step size case

In this section we consider the decreasing step size (stochastic approximation) case of (2), i.e.,

$$W_{k+1} = W_k + \mu_k(V_k - W_k) = W_k + \mu_k(F(W_k, \eta_k) - W_k), \tag{17}$$

with the usual assumptions on the sequence $\mu_k \geq 0$, i.e. that $\sum_{k=1}^{\infty} \mu_k = \infty$, $\sum_{k=1}^{\infty} \mu_k^2 < \infty$, $\sum_{k=1}^{\infty} |\mu_{k+1} - \mu_k| < \infty$. We can use Theorem 3.3 of [4] to obtain the following result.

Theorem 7.1 *Under assumption A1, W_k converges to w^* almost surely as $k \to \infty$.*

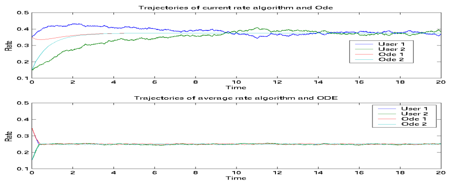
Extensions: With an *i.i.d.* evolution for the state process η_k we [19] can relax assumption **A1** to cover cases where $\bar{F}(\cdot)$ is a member of a certain class of set-valued functions and prove the result in Thm. 7.1.

8 Algorithm Simulations and comparison with Numerical Analysis of the ODE

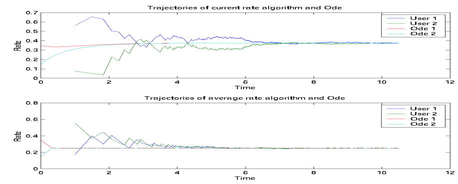
In this section using simulations and numerical analysis we provide evidence for the results obtained in this paper. For simplicity we consider 2 user scenarios and simplex-type rate-regions $\mathcal{R}(\eta)$ for each state η.

For the first scenario we assume that the state process takes values in $[0, 1]$ with the states chosen *i.i.d.* with uniform probability. Given that the state is $\eta \in [0, 1]$ we assume that the maximum rate that user 1 gets is given by $R_1(\eta) = 1 - \eta$ and the maximum rate that user 2 gets is given by $R_2(\eta) = \eta$. Note that choice of the state distribution ensures that we satisfy the continuity condition A1. The rest are easily verified for this simple case. For this case it can be shown that the boundary of the steady-state capacity region $\bar{\mathcal{R}}$ is given by $\{(R_1^a, R_2^a) = (\frac{1+2a}{2(1+a)^2}, \frac{a^2+2a}{2(1+a)^2}), a \in [0, +\infty]\}$ satisfying the strict convexity assumption. Now assuming that the utility functions of the two users are the same it is clear that the optimal operating point in the steady-state capacity region is given by $(0.375, 0.375)$ and the same in the inscribed simplex region is given by $(0.25, 0.25)$. In Figure 1(a) with a $\mu = 0.001$ and a simulation of the algorithm for $20,000$ steps we show that the stochastic process of the rates using the $\log(\cdot)$ utility function is well approximated by the ODE solution. In Figure 1(b) the (asymptotic) convergence for the decreasing step-size case of the same utility function is demonstrated. Since it suffices to consider the ODE to predict the behaviour of the algorithm we demonstrate how the different algorithms converge to the different optimal points in Figure 3(a). Considering the average rate based algorithm we can argue that the trajectories will approach a fixed curve (straight-line from $\mathbf{0}$ if all the utility functions are the same) and then go along the curve to the optimal solution.

For the second scenario we assume the same state process but assume that given that the state is $\eta \in [0, 1]$, the maximum rate that user 1 gets stays the same as before but that for user 2 is now given by $R_2(\eta) = 0.5\eta$. The boundary of the steady-state capacity region is now given by $\{(R_1^a, R_2^a) = (\frac{2(1+a)}{(2+a)^2}, \frac{4a+a^2}{4(2+a)^2}), a \in [0, +\infty]\}$. Assuming a $\sqrt{\cdot}$ utility function we can show that the optimal point in the steady-state capacity region is approximately given by $(0.4053, 0.1703)$ and the optimal point in the inscribed simplex region is given by $(1/3, 1/12) \approx (0.3333, 0.0833)$. In Figure 2 with a $\mu = 0.0001$ and a simulation of the algorithm for $100,000$ steps we show that the ODE very accurately predicts the stochastic process. Once again in Figure 3(b) we show how the different algorithms converge to the different optimal points.



(a) Comparison of the constant step-size algorithm with ODE.



(b) Comparison of the decreasing step-size algorithm with ODE.

Figure 1: Trajectories of the different algorithms and comparison with the respective ODEs for the first scenario.

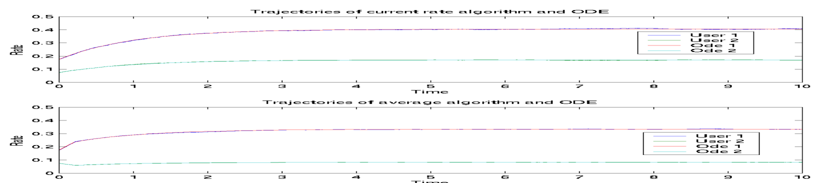


Figure 2: Trajectories of the different algorithms and comparison with the respective ODEs for the second scenario.

9 Conclusions

Based upon a stochastic approximation approach we showed that a general class of "gradient-like" *opportunistic* scheduling algorithms converge to the optimal solution of a related optimization problem. Using this we showed that the gains of multiuser diversity critically depend upon shape of the steady-state region based upon the amount of information that can be fed-back to the scheduler about the rate-regions in different channel states. This knowledge is critical, for instance, in systems with multiple transmit antennae where it is optimal to transmit to more than one user.

References

[1] J. A. Bucklew and T. G. Kurtz, "Mixed time-scale recursive algorithms," *IEEE Transactions on Signal Processing*, Vol. 49, No. 8, pp. 1824-1830, Aug. 2001.

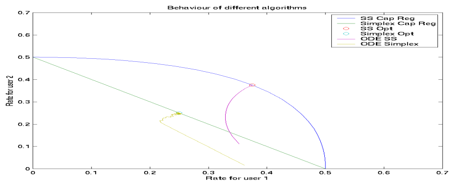
[2] J. A. Bucklew, T. G. Kurtz, and W. A. Sethares, "Weak convergence and local stability properties of fixed step size recursive algorithms," *IEEE Transactions on Information Theory*, Vol. 39, No. 3, pp. 966-978, May 1993.

[3] D. P. Bertsekas, "Nonlinear programming," Athena Scientific, Belmont, Massachusetts, 1995.

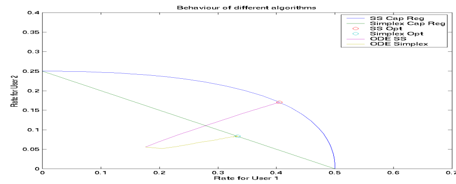
[4] T. G. Kurtz, "Averaging for martingale problems and stochastic approximations," *Applied Stochastic Analysis, Proceedings of the US-Frnch Workshop, Lecture Notes in Control and Information Sc., 177*, pp. 186-209, 1992.

[5] M. Vidyasagar, *Nonlinear System Analysis*. Prentice Hall, 1978.

[6] R. Agrawal, A. Bedekar, R. La, R. Pazhyannur, and V. Subramanian, "A Class and Channel-Condition based Weighted Proportionally Fair Scheduler for EDGE/GPRS," in *Proceedings of ITCOM'01*, Denver, August 2001.



(a) Trajectories of ODEs for the first scenario.



(b) Trajectories of ODEs for the second scenario.

Figure 3: Convergence of the ODEs to the optimal solutions.

[7] R. Agrawal, A. Bedekar, R. La, and V. Subramanian, "A Class and Channel-Condition based Weighted Proportionally Fair Scheduler," in *Proceedings of ITC 2001*, Salvador, Brazil, December 2001.

[8] X. Liu, E. K. P. Chong, and N. B. Shroff, "Opportunistic transmission scheduling with resource-sharing constraints in wireless networks," *IEEE Journal on Selected Areas in Communications*, 19(10), Oct 2001.

[9] X. Liu, E. K. P. Chong, and N. B. Shroff, "A framework for opportunistic scheduling in Wireless Networks," to appear in *Computer Networks*.

[10] S. C. Borst and P. A. Whiting, "Dynamic Rate Control Algorithms for HDR Throughput Optimization," in *Proceedings of IEEE Infocom 2001*, pp. 976-985.

[11] A. Jalali, R. Padovani, and R. Pankaj, "Data throughput of CDMA-HDR: A high efficiency, high data rate personal communications system," in *Proceedings of 50th IEEE Vehicular Technology Conference 2000*, pp. 1854-1858.

[12] M. Andrews, S. Borst, F. Dominique, P. Jelenkovic, K. Kumaran, K. Ramakrishna, and P. Whiting, "Dynamic bandwidth allocation algorithms for high-speed data wireless networks," *Bell Labs Technical Report*, 2000.

[13] R. Leelahakriengkrai and R. Agrawal, "Scheduling in Multimedia Wireless Networks," Technical Report ECE-99-3, ECE Dept., University of Wisconsin - Madison, July 1999. To appear in *IEEE Trans. on Vehicular Technology.*

[14] R. Leelahakriengkrai and R. Agrawal, "Scheduling in Multimedia Wireless Networks," *17th International Teletraffic Congress*, Salvador da Bahia, Brazil, December 2-7, 2001.

[15] S. Shakkottai and A. Stolyar, "Scheduling algorithms for a mixture of real-time and non-real time data in HDR," in *Proceedings ITC-17 2001*, Salvado da Bahia, Brasil, pp. 793-804.

[16] L. Tassiulas and A. Ephremides, "Dynamic server allocation to parallel queues with randomly varying connectivity," *IEEE Trans. on Information Theory* Vol. 30, 1993, pp. 466-478.

[17] H. Viswanathan and S. Venkatesan, "The impact of antenna diversity in packet data systems with scheduling," *IEEE Trans. Comm.*, submitted for publication.

[18] D. N. C. Tse, "Multiuser diversity and proportionally fair scheduling," in preparation.

[19] R. Agrawal and V. G. Subramanian, "Optimality of certain channel aware scheduling algorithms," in preparation.