Dynamic Channel-Sensitive Scheduling Algorithms for Wireless Data Throughput Optimization

Sem Borst and Phil Whiting, *Member, IEEE*

***Abstract*—The relative delay tolerance of data applications, together with bursty traffic characteristics, opens up the possibility for scheduling transmissions so as to optimize throughput. A particularly attractive approach in fading environments is to exploit the variations in the channel conditions and transmit to the user with the currently "best" channel. We show that the "best" user may be identified as the maximum-rate user when feasible rates are weighed with some appropriately determined coefficients. Interpreting the coefficients as shadow prices, or reward values, the optimal strategy may thus be viewed as a revenue-based policy, which always assigns the transmission slot to the user yielding the maximum revenue.**

Calculating the optimal-revenue vector directly is a formidable task, requiring detailed information on the channel statistics. Instead, we present adaptive algorithms for determining the optimal-revenue vector online in an iterative fashion, without the need for explicit knowledge of the channel behavior. Starting from an arbitrary initial vector, the algorithms iteratively adjust the reward values to compensate for observed deviations from the target throughput ratios. The algorithms are validated through extensive numerical experiments. Besides verifying long-run convergence, we also examine the transient performance, in particular the rate of convergence to the optimal-revenue vector. The results show that the target throughput ratios are tightly maintained and that the algorithms are able to track sudden changes in the channel conditions or throughput targets well.

***Index Terms*—High data rate, scheduling, stochastic control, throughput optimization.**

I. Introduction

NEXT-GENERATION wireless networks are expected to support a wide range of services, including high-rate data applications. In contrast to voice users, data applications can usually sustain some amount of packet delay, as long as the throughput over somewhat longer intervals is sufficient. The relative delay tolerance of data applications, together with bursty traffic characteristics, opens up the potential for scheduling transmissions so as to optimize throughput. A coordinated approach along these lines is proposed in [5].

A related approach may be advocated for low-mobility scenarios, such as indoor networks. In such environments, Rayleigh fading frequencies can be quite low and the fading levels can even be anticipated to some extent. For example, fading can be measured by having the base station provide a pilot signal, which can be measured by all the users. These measurements can be fed back to the base station and used to estimate fading levels and, hence, user rates in subsequent slots. With a little simplification, let us suppose that at the start of each slot the base station has perfect knowledge of the maximum feasible rate at which each user can receive and decode a signal with some acceptably low error probability. This is the approach used in the IS-856 [also known as high data rate (HDR)] standard [6].

The above framework allows the base station to schedule transmissions to users when their channel conditions are favorable. The so-called proportional fair algorithm [10] is specifically designed to achieve the latter objective. The key feature is to select users when their rates are near-optimal in a relative sense, so as to optimize throughput performance while ensuring some degree of fairness among users. The proportional fair algorithm is the default scheduling mechanism implemented in current product releases that are based on the IS-856 standard.

The selection of the "best" user depends, of course, on the performance objective that is considered. Depending on the specific situation, there are various performance criteria that might be adopted. In the present paper, we specifically consider throughput optimization relative to prespecified target values. These target values may be set arbitrarily, taking into account the quality-of-service requirements of the users or possibly their current activity levels or locations. For given target ratios, we show that the "best" user may be identified as the maximum-rate user when feasible rates are weighed with some appropriately determined coefficients. Interpreting the coefficients as shadow prices, or reward values, the optimal strategy may thus be viewed as a revenue-based policy. Under such a policy, the transmission slot is always assigned to the user yielding the maximum revenue.

Unfortunately, calculating the optimal-revenue vector (i.e., the revenue vector associated with the optimal strategy) directly is a complicated problem, requiring detailed information on the channel statistics. Although the feasible rates of the users are assumed known slot by slot, the underlying probability distribution that is producing these rates is unknown. Even if it were known, it would not be easy to use since the feasible rates might be dependent, so that the computations would be significantly hampered by the curse of dimensionality.

To avoid these obstacles, we develop adaptive algorithms for determining the optimal-revenue vector online, in an iterative fashion, without the need for explicit knowledge of the channel behavior. Starting from an arbitrary initial vector, the algorithms iteratively adjust the reward values to compensate for observed deviations from the target throughput ratios. The corrections ensure that discrepancies in throughput cannot persist. To ensure convergence to the optimal-revenue vector, the size of the adjustments is gradually reduced.

The algorithms are validated through extensive numerical experiments. Besides verifying long-run convergence, we also ex-

Manuscript received May 29, 2001; revised October 7, 2002.
The authors are with Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974 USA (e-mail: pawhiting@lucent.com).
Digital Object Identifier 10.1109/TVT.2003.810967

amine the transient performance, in particular the rate of convergence to the optimal-revenue vector. The results show that the target throughput ratios are tightly maintained and that the algorithms are able to track sudden changes in the channel conditions or throughput targets well.

Some interesting related algorithms are proposed in [2], [3], [11], [18], and [19], where queue lengths, rather than rewards, are used as weight factors. These algorithms provide throughput guarantees in terms of bounded expected queue lengths (if achievable) rather than target ratios. The abovementioned proportional fair algorithm [10] has a similar structure as well, where the weights are taken reciprocal to the historical average throughputs (with exponential smoothing). The latter algorithm inherits its name from the fact that the achieved-throughput vector is such that, for no single user, the throughput can be improved without reducing the throughputs of the other users by a greater total percentage, which property is referred to as "proportional fairness." A further class of algorithms that are based on a utility maximization formulation are proposed in [1]. Algorithms aimed at optimizing throughput performance subject to additional fairness constraints are described in [12]. The application of the above algorithms opens up two important possibilities for improving network performance, which deserve further investigation. The first is that admission control can be applied by using a *probing technique*, an approach proposed in [4]. In parallel to the actual control algorithm, a "dummy" version of the control would be run with the new user added. The impact of the new user would then be assessed on the basis of the new revenue values as determined by the dummy control. It should be noted that the decisions from the dummy control would *not* be acted on, which means that existing users are unaffected. As an additional benefit, the new revenue values would be immediately available, in case the user is admitted. The second possibility is coordinated operation of base stations in the network, which allows for load sharing and higher throughput for edge users.

The remainder of the paper is organized as follows. In Section II, we present a detailed model description and introduce a class of revenue-based scheduling strategies. We subsequently prove that revenue-based policies optimize throughput relative to prespecified target values, for discrete rate distributions as well as for continuous rates in Sections III and IV, respectively. In Sections V–VII, we develop adaptive online algorithms for determining the optimal-revenue vector in an iterative fashion. In Section VIII, we describe some numerical experiments that we performed to examine the convergence properties of the proposed control algorithms. The results in Sections VII and VIII extend the preliminary results presented in [7]. We make some concluding remarks in Section IX.

II. Model Description

We consider a base station serving M data users. The base station transmits in slots of some fixed duration. In each slot, the base station transmits to exactly one of the users.

We assume that the feasible rates for various users vary over time, according to some stationary discrete-time stochastic process $\{(R_1(n), \ldots, R_M(n)), n = 1, 2, \ldots\}$, with $R_m(n)$ representing the feasible rate for user m in the nth slot. We assume that the base station has perfect knowledge of the maximum feasible rate $R_m(n)$ for user m at the start of the nth slot (see also Remark 2.2 below). Let $(R_1, \ldots, R_M)$ be a random vector with distribution the joint stationary distribution of the feasible rates. Denote $Y_m(n) := X_m(n)R_m(n)$, with $X_m(n)$ a 0–1 variable indicating whether or not the nth slot is assigned to user m. Define $y_m(N) := \mathbb{E}[\sum_{n=1}^{N} Y_m(n)/N]$ as the expected average throughput received by user m after N slots.

Remark 2.1: Notice that we allow for dependence between the feasible rates for the various users. Independence may be a reasonable assumption in the case of an isolated base station serving a group of independent users. In the case of several base stations, however, the feasible rates may vary not only due to independent fading, but also because of the common impact of control actions at adjacent base stations. For example, base stations may transmit at reduced power if there are no backlogged users, inducing strong correlations in interference levels between users. □

We assume that the slot duration (1.67 ms in the IS-856 system) is relatively short compared to the relevant time scales in the traffic patterns and delay requirements of data users. This opens up the possibility for scheduling the data transmissions so as to enhance performance. In particular, scheduling provides a potential mechanism for exploiting variations in the feasible rates so as to optimize throughput.

The M data users may actually be thought of as the subset of active (backlogged) users among a greater population, which may change over time. For scheduling purposes, however, the separation of time scales allows us to think of the subset of active users as nearly static and continuously backlogged. (In practice, flow-control algorithms such as transmission control protocol (TCP) will typically be used to feed data into the base-station buffer at a relatively slow rate, comparable to the actual throughput provided to the user over the wireless link. Thus, the bulk of the backlogs will usually reside at the sender rather than the base-station buffer.)

Depending on the specific situation, there are various performance criteria that might be adopted. One of the most common performance objectives is throughput maximization. This can be achieved by simply assigning each slot to the user with the currently highest feasible rate. The disadvantage is that typically only a few strong users will ever be selected for transmission, causing starvation of all others.

To alleviate that problem, an alternative option is to equalize the (expected) throughput of the various users. This can be achieved easily by assigning each slot to the user with the currently smallest cumulative throughput. The downside is that this strategy does not exploit variations in the feasible rates. Moreover, by insisting on equal throughput, a few weak users may cause the throughput of all others to be dramatically reduced.

A further option is to equalize the proportion of slots allotted to the various users. This can be realized simply by using a round-robin scheme. Again, however, this strategy fails to take advantage of the fluctuations in feasible rates. In addition, some users may end up with extremely low throughput, despite receiving their fair share of the number of slots.

In general, the performance objective will be to maximize some increasing function of the form $H(y_1, \ldots, y_M)$, with $y_m := \liminf_{N\to\infty} y_m(N)$ representing the long-run expected average throughput of user m. Now observe that the set of all feasible throughput vectors must be a convex region by time-sharing arguments. Thus, the throughput vector that maximizes the function $H(\cdot)$ must also maximize some weighted throughput combination.

To formalize the above insight, we now introduce a class of revenue-based scheduling strategies. Suppose there were rewards $w_1, \ldots, w_M$ per bit transmitted to the various users. A revenue-based strategy assigns the nth transmission slot to the user $m^*(n)$ with the current maximum rate-reward product, i.e.,

$$m^*(n) = \arg \max_{m=1,\ldots,M} w_m R_m(n).$$

Clearly, the above principle maximizes the revenue earned in each individual slot and, thus, the total cumulative revenue as well as the average revenue; hence, the term revenue-based strategy. (Usually, exactly how ties are being broken also matters. Regardless of the tie-breaking rule, however, a revenue-based strategy will definitely *not* assign the nth slot to any user k with $w_k R_k(n) < \max_{m=1,\ldots,M} w_m R_m(n)$.)

Now observe that revenue is simply a weighted combination of throughputs. Ignoring some technicalities, we thus conclude that there must exist a revenue-based strategy that maximizes the function $H(\cdot)$. Formally speaking, the optimal-revenue vector is nothing but the gradient to the feasible throughput region around the throughput vector that maximizes the function $H(\cdot)$. Although the optimal-revenue vector remains difficult to determine in general, the above observation does help to limit the search for optimal strategies to the class of revenue-based scheduling strategies.

In the present paper, we specifically consider the problem of maximizing the minimum relative long-run average throughput $\min_{m=1,\ldots,M} y_m/\alpha_m$, where $\alpha_1, \ldots, \alpha_M$ are relative target values for the various users. The optimality criterion above is equivalent to the notion of *weighted max-min fairness*, which is commonly adopted in various sorts of resource-allocation problems. A related resource-sharing concept is embodied in the generalized processor sharing (GPS) paradigm [14], which is at the heart of discriminatory packet-scheduling algorithms such as weighted fair queueing (WFQ). The target values $\alpha_1, \ldots, \alpha_M$ may be set arbitrarily, taking into account the quality-of-service requirements of the users or possibly their current activity levels or locations. For example, the targets may be set lower for users with higher path losses, in order to prevent weak users from dragging down the throughput of all other users. The targets may also be applied to the proportion of slots allotted to the various users (see Remark 2.2 below).

From our earlier observation, we know that, to maximize $\min_{m=1,\ldots,M} y_m/\alpha_m$, we may restrict attention to the class of revenue-based scheduling strategies. Further observe that we may assume that the optimal-throughput vector realizes the target throughput ratios $\alpha_1, \ldots, \alpha_m$ with equality, since one could always reduce the throughputs of users with a surplus. Thus, we conclude that any revenue-based policy that additionally balances the throughputs according to the ratios $\alpha_1, \ldots, \alpha_M$ is, in fact, optimal, which provides the key principle underlying our further approach.

Finally, observe that setting throughput targets is equivalent to normalizing the feasible rates by the corresponding values. In the subsequent analysis, we therefore assume that the throughput targets are discounted for in the rates and take $(\alpha_1, \ldots, \alpha_M) = (1, \ldots, 1)$.

Remark 2.2: In practice, there is always a small probability that a transmission fails because the signal cannot be successfully decoded. The results of the present paper then remain valid if $R_m(n)$ is redefined to represent the expected feasible rate and the 0–1 variable $X_m(n)$ is amended to indicate both which user is selected and whether or not the transmission is successful.

Instead of the (expected) feasible rate, one can also take $R'_m(n) := K_m + R_m(n)$, with the K_m's positive coefficients, to obtain a weighted combination of received rates and slot allocations. By choosing suitable values for the K_m's, one can give weight to balancing the proportion of slots allotted to the various users, besides achieving relative throughput targets. □

Remark 2.3: The results in [12] show that optimizing a throughput function subject to additional fairness constraints in terms of the time fractions received by the various users may induce optimal policies with a different structure. Apparently, imposing additional constraints on the time fractions may give rise to optimal-throughput vectors that are not Pareto-optimal in the absence of these constraints. □

III. Discrete Rate Distribution

In this section, we consider the case where feasible rates $(R_1, \ldots, R_M)$ have a discrete distribution on some bounded set $J \subseteq \mathbb{R}^M$. Since feasible rates are assumed stationary, we restrict attention to the class of stationary policies in order to not blur the presentation with technicalities. The analysis may be readily extended, however, to deal with nonstationary policies.

We first introduce some notation. Let p_j be the stationary probability that the feasible rate vector is $j \in J$. (Note that j is an M-dimensional vector.) We write $R_{ij} = j_i$ for $j = (j_1, \ldots, j_M) \in J$. Let x_{ij}^π be the long-run fraction of time that policy π selects user i for transmission when the feasible rate vector is $j \in J$. Then the minimum average throughput achieved under policy π is $z^\pi = \min_{i=1,\ldots,M} T_i^\pi$ with $T_i^\pi = \sum_{j\in J} p_j R_{ij} x_{ij}^\pi$. Let π^w be the revenue-based strategy corresponding to the vector $w = (w_1, \ldots, w_M)$. Without loss of generality, we assume that $\sum_{i=1}^M w_i = 1$, since only the relative values of the revenues matter.

Lemma 3.1: Policy π is optimal if x_{ij}^π, z^π is an optimal solution to the following linear program:

$$\begin{aligned} \max \quad & z \\ \text{sub} \quad & z \le \sum_{j\in J} p_j R_{ij} x_{ij} \quad i = 1, \ldots, M \\ & \sum_{i=1}^M x_{ij} \le 1 \quad j \in J \\ & x_{ij} \ge 0 \quad i = 1, \ldots, M, j \in J. \end{aligned} \tag{1}$$

Proof: Let x^*_{ij}, z^* be an optimal solution to the linear program above. Now consider the policy that assigns the slot to user i with probability x^*_{ij} when the feasible rate vector is $j \in J$. The minimum average throughput achieved under this policy is $\min_{i=1,\ldots,M} \sum_{j\in J} p_j R_{ij} x^*_{ij} \geq z^*$. Thus, the optimal achievable throughput is at least z^*.

Conversely, for any policy π, x^π_{ij}, z^π are a feasible solution to the above linear program. Thus, the optimal achievable throughput is at most z^* and, hence, exactly z^*. The statement then easily follows. □

It follows from the above lemma, in conjunction with basic linear programming theory [16], that there exists an optimal policy π with at most $|J| + M - 1$ of the variables x^π_{ij} nonzero, which forces most of the variables to be one. Thus, only for a limited number of rate combinations, the slots are shared among several users.

In Section II, we observed that a revenue-based policy that balances the throughputs is optimal. The next theorem shows that the revenue criterion is in fact a necessary optimality condition, in the sense that there exists a revenue vector w^* such that when user i does *not* have the maximum rate-reward product, i.e., $w^*_i R_{ij} < \max_{m=1,\ldots,M} w^*_m R_{mj}$, then $x^\pi_{ij} = 0$, i.e., user i should *not* be selected for transmission. Thus, any optimal strategy *must* be a revenue-based policy associated with w^* (see [2] for a related stability result).

Theorem 3.1: If policy π is optimal, then there exists a vector $w^* \geq 0$ such that

$$x^\pi_{ij}\left[w^*_i R_{ij} - \max_{m=1,\ldots,M} w^*_m R_{mj}\right] = 0 \qquad (2)$$

for all $i = 1, \ldots, M$, $j \in J$.

Proof: By Lemma 3.1 , the x^π_{ij} are an optimal solution to the linear program (1). Now let w^*_i, y^*_j be an optimal solution to the dual problem of (1)

$$\begin{aligned} \min \; & \sum_{j\in J} y_j \\ \text{sub} \; & \sum_{i=1}^{M} w_i \geq 1 \\ & y_j \geq p_j R_{ij} w_i \quad i = 1,\ldots,M, j \in J \\ & w_i \geq 0 \quad i = 1,\ldots,M \\ & y_j \geq 0 \quad j \in J. \end{aligned} \qquad (3)$$

Then the complementary slackness conditions [16] imply $x^\pi_{ij}[y^*_j - p_j R_{ij} w^*_i] = 0$, while optimality forces $y^*_j = p_j \max_{m=1,\ldots,M} w^*_m R_{mj}$, yielding (2). □

The dual problem (3) may be interpreted as follows. The variable $y^*_j = p_j \max w^*_m R_{mj}$ represents the revenue generated in state j, so that the objective function measures the total expected earned revenue. Also, optimality implies $\sum_{i=1}^M w^*_i = 1$. Thus, the dual problem amounts to finding a revenue vector w^* that minimizes the total expected earned revenue, subject to the constraint $\sum_{i=1}^M w^*_i = 1$.

In conclusion, for policy π^{w^*} to balance the throughputs, the revenue vector w^* must minimize the total expected earned revenue, which may also be derived as follows. For any vector w with $\sum_{i=1}^M w_i = 1$, the total expected earned revenue is

$$\begin{aligned} R(w) &= \sum_{i=1}^M w_i T_i^{\pi^w} \geq \sum_{i=1}^M w_i T_i^{\pi^{w^*}} \geq \sum_{i=1}^M w_i \min_{m=1,\ldots,M} T_m^{\pi^{w^*}} \\ &= \sum_{i=1}^M w^*_i \min_{m=1,\ldots,M} T_m^{\pi^{w^*}} = \sum_{i=1}^M w^*_i T_i^{\pi^{w^*}} = R(w^*). \end{aligned}$$

IV. Continuous Rate Distribution

In this section, we consider the case where the feasible rates $(R_1, \ldots, R_M)$ have a continuous distribution on some bounded set $U \subseteq \mathbb{R}^M$.

We first introduce some notation. Let $p(u)$ be the stationary density of the feasible rate vector, i.e., the probability that the feasible rates are in some set $V \subseteq U$ is $\int_{u\in V} p(u)du$. We write $R_i(u) = u_i$ for $u = (u_1, \ldots, u_M) \in U$. Let $x^\pi_i(u)$ be the long-run fraction of time that policy π selects user i for transmission when the feasible rate vector is $u \in U$.

Lemma 4.1: Policy π is optimal if $x^\pi_i(u)$, z^π are an optimal solution to the following mathematical program:

$$\begin{aligned} \max \; & z \\ \text{sub} \; & z \leq \int_{u\in U} p(u) R_i(u) x_i(u) du \quad i = 1,\ldots,M \\ & \sum_{i=1}^M x_i(u) \leq 1 \quad u \in U \\ & x_i(u) \geq 0 \quad i = 1,\ldots,M, u \in U. \end{aligned} \qquad (4)$$

The proof of the above lemma is similar to that of Lemma 3.1.

In Section II, we reasoned that a revenue-based policy that balances the throughputs is optimal. The next theorem shows that the revenue principle is in fact a necessary optimality criterion, in the sense that there exists a revenue vector w^* such that if user i does *not* have the maximum rate-reward product on some set of nonzero measure, then user i should not be selected for transmission on that set. Thus, in the above sense, any optimal strategy *must* be a revenue-based policy associated with w^*.

Theorem 4.1: If policy π is optimal, then there exists a vector $w^* \geq 0$ such that

$$\int_{u\in U} x^\pi_i(u) \left[w^*_i R_i - \max_{m=1,\ldots,M} w^*_m R_m(u)\right] p(u)du = 0 \quad (5)$$

for all $i = 1, \ldots, M$.

Proof: By Lemma 4.1 , the $x^\pi_i(u)$ are an optimal solution to the mathematical program (4). Now let w^*_i, $y^*(u)$ be an optimal solution to the following "dual" problem of (4):

$$\begin{aligned} \min \; & \int_{u\in U} y(u) du \\ \text{sub} \; & \sum_{i=1}^M w_i \geq 1 \\ & y(u) \geq p(u) R_i(u) w_i \quad i = 1,\ldots,M, u \in U \\ & w_i \geq 0 \quad i = 1,\ldots,M \\ & y(u) \geq 0 \quad u \in U. \end{aligned} \qquad (6)$$

Then the complementary slackness conditions [16] yield $x_i^\pi(u)[y^*(u) - p(u)R_i(u)w_i^*] = 0$, while optimality requires $y^*(u) = p(u) \max_{m=1,\ldots,M} w_m^* R_m(u)$, giving (5). (Although strong duality does not directly apply, the complementary slackness properties may be derived via discretization.) □

V. Adaptive Algorithms

In the previous two sections, we concluded that revenue-based policies optimize throughput relative to pre-specified target values. However, calculating the optimal-revenue vector directly is a complicated problem, requiring detailed information on the channel statistics in the form of the joint stationary distribution of the feasible rates $(R_1, \ldots, R_M)$. Instead, we develop adaptive scheduling algorithms for determining the optimal-revenue vector online in an iterative fashion, without the need for explicit knowledge of the channel behavior. Specifically, in the nth slot, a revenue vector $w(n)$ is used for selecting a user for transmission, i.e., the nth transmission slot is assigned to the user $m^*(n)$ identified as $m^*(n) = \arg\max_{m=1,\ldots,M} w_m(n)R_m(n)$. Starting from an arbitrary initial vector $w(1)$, the algorithms iteratively adjust the reward values to compensate for observed deviations from the target throughput ratios. The corrections ensure that discrepancies in throughput cannot persist. To ensure convergence to the optimal-revenue vector w^*, the size of the adjustments is gradually reduced.

In the next two sections, we assume that the distribution of the feasible rates is modulated by some underlying stochastic process $S(n)$, which may be interpreted as the channel state. The evolution of the process $S(n)$ is governed by a discrete-time irreducible Markov chain with a finite discrete state space $\mathcal{S}$. When the channel state is $s \in \mathcal{S}$, the feasible rates have some continuous M-dimensional distribution $F_s(\cdot)$ on $\mathcal{R} \subseteq [R_{\min}, R_{\max}]^M$, $0 < R_{\min} < R_{\max} < \infty$, with zero probability measure in any set of Lebesgue measure zero. In practice, the feasible rates will typically have to be selected from a limited set of discrete values. However, we may adhere to the above assumptions by simply adding a small random perturbation. By choosing the sufficiently small random perturbation, the true achieved throughputs should be arbitrarily close to the perturbed ones.

Denote by $\mathcal{W} := \{w \in \mathbb{R}_+^M : \sum_{m=1}^M w_m = 1\}$ the set of all price vectors. For any $w \in \mathcal{W}$, denote by $\Xi_m(w)$ the expected average throughput per slot received by user m under price vector w in stationarity. Define $\Xi_{\text{ave}}(w) := (1/M)\sum_{m=1}^M \Xi_m(w)$, $\Xi_{\min}(w) := \min_{m=1,\ldots,M} \Xi_m(w)$, and $\Xi_{\max}(w) := \max_{m=1,\ldots,M} \Xi_m(w)$ as the average, the minimum, and the maximum expected throughput per slot under price vector w over all users, respectively.

The above assumptions ensure that the expected throughput vector $\big(\Xi_1(w), \ldots, \Xi_M(w)\big)$ is completely determined by the price vector w (without the need to specify a tie breaking rule). The assumptions further imply that the expected throughput vector $\big(\Xi_1(w), \ldots, \Xi_M(w)\big)$ is a continuous function of the price vector w.

To facilitate the presentation, we assume that the optimal price vector w^* is unique. The analysis may be readily modified for the case where there is a whole range of optimal price vectors.

VI. Two Users

We first focus on the case of two users. In the next section, we consider the situation with an arbitrary number of users.

A. Algorithm Description

Before describing the algorithm in detail, we first introduce some useful notation. With minor abuse of notation, we write $w = w_1$, so that $w_2 = 1 - w$. Denote $\Delta Y(n) := Y_1(n) - Y_2(n)$ and define $U(N) := \sum_{n=1}^N \Delta Y(n)$ as the difference in cumulative throughput between users 1 and 2 after N slots. The absolute difference $|U(N)|$ is referred to as the throughput *gap*. We say that the throughput gap *widens* in the Nth slot if $|U(N)| > \max_{n=1,\ldots,N-1} |U(n)|$. User 1 is said to be *leading* if $U(N) > 0$ and is referred to as *lagging* otherwise (vice versa for user 2). We say that a *crossover* occurs in the Nth slot if the leading and lagging users exchange positions, which means that the throughput gap changes sign, i.e., $U(N)U(N-1) < 0$.

The algorithm may now be described as follows. In every slot, the user with the maximum price-rate product, at the current price value, is selected for transmission. Thus, the nth slot is assigned to user 1 if $w(n)R_1(n) > (1 - w(n))R_2(n)$ and to user 2 otherwise (ties being broken arbitrarily).To drive the price sequence $w(n)$ toward the optimal value w^*, the price is adjusted over time on the basis of the observed throughput realizations. As long as the throughput gap does *not* widen, the price is left unaltered. However, if the throughput gap *does* widen, then the price is changed in favor of the deficit user; thus, at the expense of the surplus user. The price of the leading user is decreased by $\delta_{k(n)}$, while the price of the lagging user is simultaneously increased by the same amount.

To ensure convergence, a *reset* is triggered at every crossover. The step size $\delta_{k(n)}$ is then reduced by incrementing $k(n)$, with $\{\delta_k, k = 1, 2, \ldots\}$ a predetermined convergent sequence (e.g., $\delta_{k+1} = \delta_1 \rho^k$ with $\rho < 1$ or $\delta_k = \delta_1 k^{-\beta}$ with $\beta > 1$).

B. Convergence Proof

We now proceed to demonstrate convergence of the above-described algorithm. We first state an important assumption.

Assumption 6.1 (Large-Deviations Assumption): Let $X_m^N(s, w)$ be a random variable representing the average throughput per slot obtained by user $m$over a period of N slots under price-vector w, given that the initial state of the Markov chain is s. Given a price vector $w \in \mathcal{W}$ and $\xi > 0$, there exist numbers $C_m^\xi(w), D_m^\xi(w) > 0$ such that for any initial state s

$$\mathbb{P}\{|\, X_m^N(s, w) - \Xi_m(w)\,| > \xi\} \leq C_m^\xi(w)\, e^{-D_m^\xi(w)N}$$

$m = 1, 2.$

It may be verified that the above assumption is satisfied for the feasible-rate process described earlier.

Let $Y_m^{w^*+\epsilon}(n)$ be random variables representing the throughput that user m would receive in the nth slot if the price were fixed at $w^* + \epsilon$, $m = 1, 2$. Define $\Delta Y^{w^*+\epsilon}(n) := Y_1^{w^*+\epsilon}(n) - Y_2^{w^*+\epsilon}(n)$ as the difference in throughput between users 1 and 2 in the nth slot. Define $\xi := \Xi_1(w^* + \epsilon) - \Xi_2(w^* + \epsilon) > 0$ as the difference in expected throughput between users 1 and 2 in stationarity. For all $N \geq n_0 \geq 0$, the events

$$\sum_{n=n_0}^{N} Y_1^{w^*+\epsilon}(n) \geq (N - n_0 + 1)\left(\Xi_1(w^* + \epsilon) - \frac{\xi}{4}\right)$$

and

$$\sum_{n=n_0}^{N} Y_2^{w^*+\epsilon}(n) \leq (N - n_0 + 1)\left(\Xi_2(w^* + \epsilon) + \frac{\xi}{4}\right)$$

imply the event

$$\sum_{n=n_0}^{N} \Delta Y^{w^*+\epsilon}(n) \geq \frac{(N - n_0 + 1)\xi}{2}.$$

Assumption 6.1 then implies that there exist numbers $C, D > 0$ such that

$$\mathbb{P}\left\{\sum_{n=n_0}^{N} \Delta Y^{w^*+\epsilon}(n) < \frac{(N - n_0 + 1)\xi}{2}\right\} \leq 2Ce^{-D(N-n_0+1)}$$

which means that

$$\sum_{n=n_0}^{N} \Delta Y^{w^*+\epsilon}(n) \to \infty \tag{7}$$

with probability (wp) 1 as $N \to \infty$.

The next theorem establishes almost-sure convergence to the optimal-revenue vector.

Theorem 6.1: For the scheduling algorithm described above, the price sequence $w(n)$ converges to the optimal price w^* wp 1 and, consequently, the sequence $z(n)$ converges to the optimal value $z^{\pi^{w^*}}$ wp 1.

In preparation for the proof of the above theorem, we first present two lemmas.

Lemma 6.1: The price sequence $w(n)$ cannot get permanently trapped in either of the intervals $[0, w^* - \epsilon]$ or $[w^* + \epsilon, 1]$.

Proof: We only prove the statement for the interval $[w^* + \epsilon, 1]$. The statement for the interval $[0, w^* - \epsilon]$ follows from symmetry considerations.

The idea of the proof is as follows. As long as the price remains in favor of user 1, the throughput difference continues to have a positive drift and will wander off to infinity. As a result, the price will keep decreasing in fixed steps and will eventually turn negative, which is not possible.

To formalize the above idea, suppose that, at some point in time, let us say the n_0-th slot, the price value enters the interval $[w^* + \epsilon, 1]$ to get permanently trapped there, i.e., $w(n) \geq w^* + \epsilon$ for all $n \geq n_0$. Then $Y_1(n) \geq Y_1^{w^*+\epsilon}(n)$ and $Y_2(n) \leq Y_2^{w^*+\epsilon}(n)$ for all $n \geq n_0$, so that $\Delta Y(n) \geq \Delta Y^{w^*+\epsilon}(n)$ for all $n \geq n_0$. Hence, (7) implies that $\sum_{n=n_0}^{N} \Delta Y(n) \to \infty$ wp 1 as $N \to \infty$. Consequently, the throughput gap $U(N) = U(n_0 - 1) + \sum_{n=n_0}^{N} \Delta Y(n) \to \infty$ wp 1 as $N \to \infty$ as well, which means that: (i) only a finite number of crossovers occur and (ii) the throughput gap will widen infinitely many times in favor of user 1. Thus, (i) the step size $\delta_{k(n)}$ will only be reduced a finite number of times and (ii) the price will be decreased infinitely many times and increased only finitely many times. Hence, the price will eventually turn negative, which is not possible. □

Lemma 6.2: The price sequence $w(n)$ cannot move from the interval [0, $w^* + \epsilon$] to the interval [$w^* + 2\epsilon$, 1] infinitely often. Similarly, $w(n)$ cannot move from the interval [$w^* - \epsilon$, 1] to the interval [0,$w^* - 2\epsilon$] infinitely often.

Proof: We only prove the first statement. The second one follows from symmetry considerations.

The idea of the proof is as follows. In order for the price sequence to move from the interval [0, $w^* + \epsilon$] to the interval $[w^* + 2\epsilon, 1]$, it must cross the interval $[w^* + \epsilon, w^* + 2\epsilon]$ from left to right. For that to happen, the algorithm must make a number of ϵ-wrong moves. By an ϵ-wrong move, we mean that the price is increased while the current price is at least ϵ above the optimal value w^*. As will be shown below, the expected number of ϵ-wrong moves before a crossover occurs is finite. However, as crossovers occur, the step size will get smaller and smaller and the required number of ϵ-wrong moves for the interval to be crossed will get larger and larger. As a result, it will eventually become increasingly unlikely for the interval to be crossed.

To make the above idea precise, we first introduce some helpful terminology. A crossover is referred to as an upward turn in case user 2 takes over the lead from user 1. Otherwise, a crossover is called a downward turn. Let $K(t)$ and $L(t)$ be the total number and the total size of ϵ-wrong moves, respectively, between the tth upward turn and the subsequent downward turn.

Note that the value of the step size between the tth upward turn and the subsequent downward turn is at most δ_{2t}. Once the value of δ_{2t} has dropped below $\epsilon/2$, we must have $L(t) > \epsilon/2$ in order for the interval $[w^* + \epsilon, w^* + 2\epsilon]$ to be crossed between the tth upward turn and the subsequent downward turn.

Also, note that the interval can be crossed at most once between the tth upward turn and the subsequent downward turn and cannot be crossed from left to right otherwise. Thus, in order for the interval to be crossed infinitely often, we must have $\sum_{t=0}^{\infty} L(t) = \infty$.

Now suppose that, at some point in time, let us say the n_0-th slot, the price value increases to enter the interval $[w^* + \epsilon, w^* + 2\epsilon]$ for the first time, between the tth upward turn and the subsequent downward turn in the N^*-th slot. Then $w(n) \geq w^* + \epsilon$ for all $n = n_0, \ldots, N^*$. As a result, $Y_1(n) \geq Y_1^{w^*+\epsilon}(n)$ and $Y_2(n) \leq Y_2^{w^*+\epsilon}(n)$ for all $n = n_0, \ldots, N^*$, so that $\Delta Y(n) \geq \Delta Y^{w^*+\epsilon}(n)$ for all $n = n_0, \ldots, N^*$ and thus

$$\sum_{n=n_1}^{n_2} \Delta Y(n) \geq \sum_{n=n_1}^{n_2} \Delta Y^{w^*+\epsilon}(n)$$

for all $n_0 \leq n_1 \leq n_2 \leq N^*$. Hence, (7) implies that $\sum_{n=n_0}^{N} \Delta Y(n)$ reaches only finitely many decreasing ladder heights for $N = n_0, \ldots, N^*$. Consequently, the throughput gap $U(N) = U(n_0 - 1) + \sum_{n=n_0}^{N} \Delta Y(n)$ widens only finitely many times in favor of user 2 for $N = n_0, \ldots, N^*$. Thus, the price is increased only finitely many times before the next downward turn occurs, i.e., $\mathbb{E}K(t) \leq K^* < \infty$, and

$$\sum_{t=0}^{\infty} \mathbb{E}L(t) \leq \sum_{t=0}^{\infty} \mathbb{E}K(t)\delta_{2t} \leq K^* \sum_{t=0}^{\infty} \delta_{2t} < \infty$$

which implies that $\sum_{t=0}^{\infty} L(t) < \infty$ wp 1. □

We conclude the section with the proof of Theorem 6.1.

Proof of Theorem 6.1: Lemma 6.1 implies that the sequence $w(n)$ spends infinitely many times in the interval $[w^* - \epsilon, 1]$ wp 1. Lemma 6.2 shows that the sequence $w(n)$ returns only finitely many times from the interval $[w^* - \epsilon, 1]$ to the interval $[0, w^* - 2\epsilon]$ wp 1. Combining these two statements, we find that the sequence $w(n)$ spends only finitely many times in the interval $[0, w^* - 2\epsilon]$ wp 1. Similarly, we have that the sequence $w(n)$ spends only finitely many times in the interval $[w^* + 2\epsilon, 1]$ wp 1. Hence, for any $\epsilon > 0$, the sequence $w(n)$ will eventually enter the interval $[w^* - 2\epsilon, w^* + 2\epsilon]$ wp 1, to never leave it again. Thus, the sequence $w(n)$ converges to the optimal price w^* wp 1.

By continuity, the sequence $\mathbb{E}[Y_m(n)]$ converges to $\Xi_m(w)$, $m = 1, 2$. The convergence of $z(n)$ then follows immediately. □

VII. Arbitrary Number of Users

We now turn to the situation with an arbitrary number of users. In principle, the algorithm for the case of two users, described in the previous section, may be extended to several users. The main subtlety lies in identifying a proper rule for when to trigger a reset. If a reset is triggered at every crossover of any pair of users, then resets may occur too rapidly. In that case, two leapfrogging users may cause the step size to be reduced quickly, while still far removed from the other users. The price sequence may then get trapped in a bias region and never reach the optimal point. A better rule is to trigger a reset only when every user has become leading or lagging. Some care is then required, though, to show that resets occur frequently enough compared to wrong moves, because otherwise the price sequence may continue to visit a bias region indefinitely.

A. Algorithm Description

In the remainder of the section, we consider a related but somewhat different algorithm, which may be described as follows. The algorithm makes price updates based on sample periods of predetermined ever-increasing size. Thus, the price updates occur at predetermined slots $K(n)$, instead of randomly determined slots as before, with $L(n) := K(n+1) - K(n)$ the length of the nth sample period. In every slot of the nth sample period, the price vector $w(n)$ is used for selecting a user for transmission. (From now on we use n to index sample periods, rather than transmission slots as before.)

To drive the price sequence $w(n)$ toward the optimal point w^*, the price is adjusted over time on the basis of the observed throughput realizations. The *direction* in which the price vector is modified at the nth update is determined by a random vector $v\big(w(n)\big)$, based on the throughput obtained during the nth sample period when the price vector $w(n)$ is used. The *size* of the nth update is $\delta(n) = \delta_{k(n)}$, with $\{\delta_k, k = 1, 2, \ldots\}$ a predetermined convergent sequence. Thus, at the $(n+1)$-th update, the price vector is recursively determined as

$$w(n+1) = w(n) - \delta(n) v\big(w(n)\big).$$

To ensure convergence, the step size $\delta(n)$ is reduced by incrementing $k(n)$ every time a reset is triggered. Intuitively, resets should occur far away from the optimal point w^*rarely, but occur readily once the price vector is close to w^*. It remains to specify the exact rules for (i), how to determine the update direction $v\big(w(n)\big)$, and (ii), when to trigger a reset.

(i) For every user, the empirical average throughput over the sample period is computed. The users are then partitioned into two groups: (1) those with above-average throughput and (2) those with below-average throughput. The prices of the above-average users are decreased, while the prices of the below-average users are increased. As the sample size grows, so that with high probability the empirical average throughputs line up with the true expected throughputs, this ensures that the price vector gets closer to the optimal point w^* in some appropriate sense, as will be shown later.

Formally, the procedure may be described as follows. Denote by X_m the throughput received by user m during a particular sample period in which price vector w is used. Define $X_{\rm ave} := (1/M) \sum_{m=1}^{M} X_m$ as the average throughput over all users. Denote by $\Omega^- := \{m : X_m \leq X_{\rm ave}\}$ and $\Omega^+ := \{m : X_m > X_{\rm ave}\}$ the groups of below-average and strictly above-average users, respectively. Then the price update direction $v(w)$ is determined as

$$v_i(w) = \frac{w_i}{\sum\limits_{m \in \Omega^-} w_m} \quad i \in \Omega^- \tag{8}$$

$$v_j(w) = \frac{-w_j}{\sum\limits_{m \in \Omega^+} w_m} \quad j \in \Omega^+. \tag{9}$$

Note that Ω^- is always nonempty, since it is impossible for all users to have strictly above-average throughput. However, Ω^+ may be empty in the case that all users have exactly equal throughput. In that case, the price vector is simply left unaltered.

Also note that the price ratios within both Ω^- and Ω^+ are maintained. This ensures that the expected throughput of the below-average users increases, while the expected throughput of the above-average users decreases, as may be deduced from Lemma 7.1 below.

Note that the above price update cannot be applied in the case that price values of some of the users in Ω^+ are zero. To prevent that situation from happening, the price process will be restricted to the set $\mathcal{W}_\nu := \{w \in \mathcal{W} : w_m \geq \nu \text{ for all } m = 1, \ldots, M\}$, with $\nu := R_{\min}/(R_{\min} + (M-1)R_{\max})$. It is easily verified that if $w_m \leq \nu$, then $\Xi_m(w) = 0$, which implies that $w^* \in \mathcal{W}_\nu$. In order to restrict the price process to the set $\mathcal{W}_\nu$, the update is truncated at the boundary if necessary.

(ii) To ensure convergence, a reset is triggered under the condition that every user has been a member of Ω^+ at least once during a consecutive sequence of updates. Once the reset has occurred, the next one is not triggered until every user has been a member of Ω^+ at least once again.

The next lemma shows that the above price update increases the throughputs of the users in Ω^- and decreases the throughputs of the users in Ω^+.

Lemma 7.1: Let $w, w' \in \mathcal{W}$ be two price vectors and $\Theta^-, \Theta^+ \subseteq \{1, \ldots, M\}$ two groups of users such that for all $i \in \Theta^-$, $w'_i/w_i \geq w'_k/w_k$ for all $k \neq i$ and for all $j \in \Theta^-$, $w'_j/w_j \leq w'_k/w_k$ for all $k \neq j$. Then

$$\Xi_i(w') \geq \Xi_i(w) \quad i \in \Theta^-$$
$$\Xi_j(w') \leq \Xi_j(w) \quad j \in \Theta^+.$$

Proof: First consider a user $i \in \Theta^-$. For any given rate, vector $(R_1, \ldots, R_M) \in \mathcal{R}$, $w_i R_i = \max\limits_{k=1,\ldots,M} w_k R_k$ implies

$w_i' R_i = \max_{k=1,\ldots,M} w_k' R_k$. In other words, if user i is selected under the old price vector w, then so is user i under the new price vector w'. Thus, the throughput of user i must increase (in fact sample path wise). Similarly, the throughput of a user $j \in \Theta^+$ must decrease. □

B. Convergence Proof

We now proceed to prove convergence of the above-described algorithm. We first discuss a few important assumptions.

Large-Deviations Assumption: As described above, the algorithm works by making price updates based on samples of ever increasing size. To ensure convergence, we need that, as the sample size grows, a "correct" price update direction is selected with sufficiently high probability. Given a price vector $w \in \mathcal{W}$, user m is called ξ-below-average (respectively, ξ-above-average) if $\Xi_m(w) < \Xi_{\text{ave}}(w) - \xi$ (respectively, $\Xi_m(w) > \Xi_{\text{ave}}(w) + \xi$). We say that the price update direction is "ξ-right" if all the ξ-below-average users belong to Ω^- and have their price increased and all the ξ-above-average users belong to Ω^+ and have their price decreased. (Otherwise, the price direction is "ξ-wrong.") This ensures that the price vector gets closer to the optimal point w^* in some appropriate sense, as will be shown later. Now remember that, at each update, the prices of the empirical below-average users are increased, while the prices of the empirical above-average users are decreased. Thus, for the price update direction to be "correct," it is critical that the empirical average throughputs line up with the true expected throughputs. This then motivates the following assumption.

Assumption 7.1 (Large-Deviations Assumption): Let $X_m^n(w)$ be a random variable representing the average throughput per slot obtained by user m over a period of $L(n)$ slots under price vector w in stationarity. Given a price vector $w \in \mathcal{W}$ and $\xi > 0$, there exist a ζ-neighborhood $N_\zeta^\xi(w)$ of w and numbers $C_m^\xi(w), D_m^\xi(w) > 0$ such that

$$\mathbb{P}\{|\, X_m^n(w') - \Xi_m(w)\,| > \xi\} \le C_m^\xi(w)\, e^{-D_m^\xi(w) L(n)}$$

for all $w' \in N_\zeta^\xi(w)$, $m = 1, \ldots, M$.

In Appendix I, we prove that the above assumption is satisfied for the feasible-rate process described earlier.

Boundary Conditions: We further require that, when a correct price direction is selected, the update cannot be truncated to an arbitrarily small size. The following assumption implies that if a correct price direction is chosen then, for small enough step size δ, the price sequence will stay away from the boundary.

Assumption 7.2: There exist positive constants $\delta^* > 0, \xi > 0$ such that for all price vectors $w \in \mathcal{W}_\nu$, for any ξ-right direction $v(w)$, and for any $\delta \in (0, \delta^*)$

$$w + \delta v(w) \in \mathcal{W}_\nu.$$

To check that the above assumption is satisfied, it suffices to verify that extremely low prices cannot be decreased and that extremely high prices cannot be increased. First, consider a user i with a price $w_i < R_{\min}/(R_{\min} + (M-1)R_{\max})$. Then the throughput of user i is zero and, thus, certainly ξ-below-average for some $\xi > 0$, which means that the price of user i is increased if the price direction is right. Similarly, the throughput of a user j with a price $w_j > R_{\max}/(R_{\min} + R_{\max})$ is ξ-above-average for some $\xi > 0$, so that the price of user j is decreased if the price direction is right.

Function $T(\cdot)$: As indicated above, we also need that when a correct price update direction is selected, the price vector gets closer to the optimal point w^* by some definite amount. To measure distance from w^*, we introduce a function $T(\cdot)$, which attains a unique minimum at w^*. Define $\Gamma_\epsilon := \{w \in \mathcal{W} : \Xi_{\max}(w) - \Xi_{\min}(w) \le \epsilon\}$ as an "ϵ-neighborhood" of w^*. The following assumption implies that, if a correct price update direction is chosen, then outside Γ_ϵ the reduction in the value of $T(\cdot)$ for small enough step size δ is at least δ times some constant of proportionality η.

Assumption 7.3: There exist positive constants $\delta^* > 0$, $\eta > 0$, $\xi > 0$ such that for all price vectors $w \notin \Gamma_\epsilon$, for any ξ-right direction $v(w)$, and for any $\delta \in (0, \delta^*)$

$$T\big(w + \delta v(w)\big) \le T(w) - \delta\eta.$$

We will consider two alternative choices for the function $T(\cdot)$. The first one is

$$T(w) := \Xi_{\max}(w) - \Xi_{\min}(w)$$

i.e., the maximum difference in expected throughput between any pair of users. By definition, $T(w^*) = 0$ and $T(w) \ge 0$ for all $w \ne w^*$, with strict inequality in the case that the optimal price vector w^* is unique.

The second function that we will consider is

$$T(w) := \sum_{m=1}^{M} w_m \Xi_m(w)$$

i.e., the total expected revenue earned. As found in Section III, the optimal price vector w^* minimizes that quantity over all vectors in the set $\mathcal{W}$, i.e., $T(w^*) \le T(w)$ for all $w \in \mathcal{W}$, $w \ne w^*$, with strict inequality in the case that w^* is unique.

In Appendix II, we prove that Assumption 7.3 is indeed satisfied for the above two $T(\cdot)$ functions. In contrast to the first $T(\cdot)$ function, the second is also suitable to show that Assumption 7.3 is satisfied for various alternative options to select a price update direction. For example

$$v_{i^*} = 1 - \beta > 0 \quad i^* = \arg\min_{m=1,\ldots,M} X_m \tag{10}$$

$$v_{j^*} = -1 \quad j^* = \arg\max_{m=1,\ldots,M} X_m \tag{11}$$

and $v_k = \beta_n/(M-2)$ for all $k \ne i^*, j^*$, for β_n a given positive sequence with $\lim_{n\to\infty} \beta_n = 0$. In the sequel, this will be referred to as the "update-extreme" algorithm, as opposed to the procedure described earlier, which will be called the "move-to-average" algorithm.

The next theorem establishes almost-sure convergence to the optimal-revenue vector w^* for the move-to-average algorithm. The proof for the update-extreme algorithm is mostly similar, except for a somewhat different notion of a correct price-update direction.

Theorem 7.1: The price sequence $w(n)$ converges to the optimal price vector w^* wp 1 and, consequently, the sequence $z(n)$ converges to the optimal value $z^{\pi^{w^*}}$ wp 1.

In preparation for the proof of the above theorem, we first introduce some terminology and present some auxiliary lemmas. We say that the nth sample is "ξ-right" if, for every user, the empirical average throughput is within ξ from the true expected throughput, i.e., $|\ X_m(n) - \Xi_m(w(n))\ | \leq \xi$ for all $m = 1, \ldots, M$. Otherwise the sample is "ξ-wrong."

Lemma 7.2: For any fixed $\xi > 0$, the total number of ξ-wrong samples is finite wp 1.

Proof: Consider some price vector $w \in \mathcal{W}$. By continuity of $\Xi_m(w)$ as a function of w, there exists for any $\eta > 0$ a β-neighborhood $N_\beta^\eta(w)$ of w such that

$$|\ \Xi_m(w') - \Xi_m(w)\ | \leq \eta \tag{12}$$

for all $w' \in N_\beta^\eta(w)$, $m = 1, \ldots, M$.

Now suppose that $w(n) = w' \in N_\beta(w)$ and that

$$|\ X_m(n) - \Xi_m(w)\ | \leq \theta \tag{13}$$

for all $m = 1, \ldots, M$. Then, using (12)–(13), taking $\eta = \theta = \xi/2$,

$$\begin{aligned} &|X_m(n) - \Xi(w(n))\ | \\ &\leq |\ X_m(n) - \Xi_m(w)\ | + |\ \Xi_m(w(n)) - \Xi_m(w)\ | \\ &\leq \eta + \theta = \xi \end{aligned}$$

for all $m = 1, \ldots, M$.

In conclusion, if $w(n) \in N_\beta^{\xi/2}(w)$, then the event (13) implies that the nth sample is ξ-right. Thus, the probability that the nth sample is ξ-wrong is then

$$\begin{aligned} \sigma(n) \leq& 1 - \mathbb{P}\{|\ X_m(n) - \Xi_m(w)\ | \\ &\leq \frac{\xi}{2} \text{ for all } m = 1, \ldots, M\}. \end{aligned} \tag{14}$$

The Large-Deviations Assumption 7.1 implies that there exist a ζ-neighborhood $N_\zeta^\xi(w)$ of w and numbers $C_m^\xi(w) > 0$, $D_m^\xi(w) > 0$ such that if $w(n) \in N_\zeta^\xi(w)$, then

$$\mathbb{P}\left\{|\ X_m(n) - \Xi_m(w)\ | > \frac{\xi}{2}\right\} \leq C_m^\xi(w)\, e^{-D_m^\xi(w)L(n)} \tag{15}$$

for all $m = 1, \ldots, M$.

Define $N^\xi(w) := N_\beta^{\xi/2}(w) \cap N_\zeta^\xi(w)$. Combining (14) and (15), if $w(n) \in N^\xi(w)$, then

$$\sigma(n) \leq \sum_{m=1}^{M} C_m^\xi(w)\, e^{-D_m^\xi(w)L(n)}.$$

Since $\mathcal{W}$ is a compact set, there exists a finite covering of such sets $N^{\xi^{(k)}}(w^{(k)})$, $k = 1, \ldots, K$. Thus, deconditioning

$$\sigma(n) \leq \sum_{k=1}^{K}\sum_{m=1}^{M} C_m^{\xi^{(k)}}(w^{(k)})\, e^{-D_m^{\xi^{(k)}}(w^{(k)})L(n)}.$$

As $L(n) = n^\beta$, with $\beta > 0$, we have $\sum_{n=1}^\infty \sigma(n) < \infty$. The statement then follows from the Borel-Cantelli lemma. □

By definition, if the nth sample is $\xi/2$-right, then

$$|\ X_m(n) - \Xi_m(w(n))\ | \leq \frac{\xi}{2}$$

for all $m = 1, \ldots, M$, which also implies

$$|\ X_{\text{ave}}(n) - \Xi_{\text{ave}}(w(n))\ | \leq \frac{\xi}{2}.$$

Hence, if user i is ξ-below average, i.e., $\Xi_i\big(w(n)\big) < \Xi_{\text{ave}}\big(w(n)\big) - \xi$, then $X_i(n) \leq X_{\text{ave}}(n)$, i.e., $i \in \Omega^-(n)$. Similarly, if user j is ξ-above average, i.e., $\Xi_j\big(w(n)\big) > \Xi_{\text{ave}}\big(w(n)\big) + \xi$, then $X_j(n) > X_{\text{ave}}(n)$, i.e., $j \in \Omega^+(n)$. Consequently, if a sample is $\xi/2$-right, then the move-to-average algorithm will select a ξ-right price update direction. The above lemma thus implies that from a certain time N on no ξ-wrong price updates will occur. It suffices to prove convergence starting from the state of the process at that time. Now observe that we may simply view the state of the process at that time as the initial state, which we allowed to be completely arbitrary. To prove convergence, we may thus assume that no ξ-wrong price updates occur at all.

Lemma 7.3: The total number of resets is infinite wp 1.

Proof: Assume that the total number of resets were finite, let us say K, and that the Kth reset occurs at the Nth price update. Assumption 7.2 ensures that the price update is never truncated to less than size δ^*, unless the price direction were ξ^*-wrong, which we may assume does not occur. Thus

$$\delta(n) \geq \min\{\delta^*, \delta_K\} \tag{16}$$

for all $n \geq N$. In view of the reset condition, there must also be some user i that belongs to either $\Omega^-(n)$ or $\Omega^+(n)$ for all $n \geq N$. Let us say $\Omega^-(n)$, thus, starting from the Nth update, the price of user i is constantly increased, i.e.,

$$w_i(n+1) \geq w_i(n) + \nu\delta(n) \tag{17}$$

for all $n \geq N$.

Combining (16) and (17), we conclude that $w_i(n) \to \infty$ as $n \to \infty$, which is not possible. □

Lemma 7.4: The price sequence $w(n)$ cannot converge to a point outside Γ_ϵ.

Proof: Assume that the price sequence does converge to a point outside Γ_ϵ; let us say w. Define $\xi := (\Xi_{\max}(w) - \Xi_{\text{ave}}(w))/2 \geq \epsilon/2M > 0$. By continuity of $\Xi_m(w)$ as a function of w, there exist a β-neighborhood $N_\beta(w)$ of w and a user i such that i is ξ-below average for all $w' \in N_\beta(w)$. Thus, if $w(n) \in N_\beta(w)$, then $i \in \Omega^-(n)$, unless the nth price update were ξ-wrong, which we may assume does not occur.

Now, since $w(n)$ converges to w, there exists an N such that $w(n) \in N_\beta(w)$ for all $n \geq N$. Thus, user i belongs to $\Omega^-(n)$ for all $n \geq N$. In other words, user i does not belong to $\Omega^+(n)$ for any $n \geq N$. That implies that no resets occur after the Nth price update, which contradicts Lemma 7.3. □

Lemma 7.5: The price sequence $w(n)$ visits Γ_ϵ infinitely often.

Proof: Assume that the price sequence visits Γ_ϵ only finitely often. Lemma 7.4 then implies that the total size of the price updates must be infinite, i.e.,

$$\sum_{n=1}^{\infty} \delta(n) = \infty. \tag{18}$$

For compactness, denote $T_n := T\big(w(n)\big)$. Lemma 7.3 implies that, at a certain time N, the step size $\delta(N)$ falls below δ^*. Assumption 7.3 then gives that

$$T_{n+1} \leq T_n - \eta\delta(n) \tag{19}$$

for all $n \geq N$, unless the nth price update was ξ^*-wrong, which we may assume does not occur.

Combining (18) and (19), we conclude that $T_n \to -\infty$ as $n \to \infty$, which is not possible. □

Lemma 7.6: The price sequence $w(n)$ cannot move from Γ_ϵ to outside $\Gamma_{2\epsilon}$ infinitely often.

Proof: Let Θ be the minimum distance between Γ_ϵ and any point outside $\Gamma_{2\epsilon}$.

Lemma 7.3 implies that at a certain time N the step size $\delta(N)$ falls below Θ. From time N on, for the price sequence to move from Γ_ϵ to outside $\Gamma_{2\epsilon}$, at least price update is required from a point $w \notin \Gamma_\epsilon$ to a point w' with $T(w') \geq T(w)$. Assumption 7.3 then implies that that price update must be ξ^*-wrong, which we may assume does not occur. □

The proof of Theorem 7.1 may now be completed as follows.

Proof of Theorem 7.1: Combining Lemmas 7.5 and 7.6, we conclude that the sequence $w(n)$ spends only finitely many times outside the region $\Gamma_{2\epsilon}$ wp 1. Hence, for any $\epsilon > 0$, the sequence $w(n)$ will eventually enter the region $\Gamma_{2\epsilon}$ wp 1, to never leave it again. Thus, the sequence $w(n)$ converges to the optimal price vector w^* wp 1.

By continuity, the sequence $\mathbb{E}[Y_m(n)]$ converges to $\Xi_m(w)$ for all $m = 1, \ldots, M$. The convergence of $z(n)$ then immediately follows. □

Remark 7.1: In the present paper, we focus on establishing almost-sure convergence to the optimal-revenue vector w^*. This critically relies on the step sizes $\{\delta_k, k = 1, 2, \ldots\}$ being a convergent sequence. As an alternative, the step sizes may be kept fixed at some given value δ. We expect that the price sequence will then continue to oscillate around w^*, but with smaller amplitudes for smaller values of δ. Observe, however, that there is an inherent trade-off between the accuracy achieved on the one hand and the speed the convergence, and thus the responsiveness to changing conditions, on the other hand. The value of δ may then be used to find the right balance between these two conflicting objectives. □

VIII. Numerical Results

In this section, we describe some numerical experiments that we conducted to investigate the convergence properties of the proposed control algorithms. Besides verifying long-run convergence, we also examine the transient performance, in particular the rate at which the prices converge to the optimal values.

In the first three experiments, we consider continuous rate distributions. In the fourth experiment, we assume a discrete distribution in which the feasible rates are determined by a fading process via the signal-to-noise ratio (SNR). The fading process is modeled using a discrete number of sinusoidal oscillators as described by Jakes' model [9].

In the final three experiments, we examine how well the throughput ratios are maintained and how well the algorithms are able to track changes in the channel conditions or throughput targets.

A. Two Users With Exponential Rates

In the first experiment, we consider a model of two users with independent rates.

The feasible rate for user i is governed by a conditional exponential distribution on some interval $[R_{\min}, R_{\max}]$, i.e.,

$$F_i(r) = G_i^{-1}[1 - e^{-\gamma_i(r - R_{\min})}], \quad r \in [R_{\min}, R_{\max}]$$

with $G_i = 1 - e^{-\gamma_i(R_{\max} - R_{\min})}$ a normalization coefficient, $i = 1, 2$. We take $[R_{\min}, R_{\max}] = [10, 400]$ Kbits/s and assume

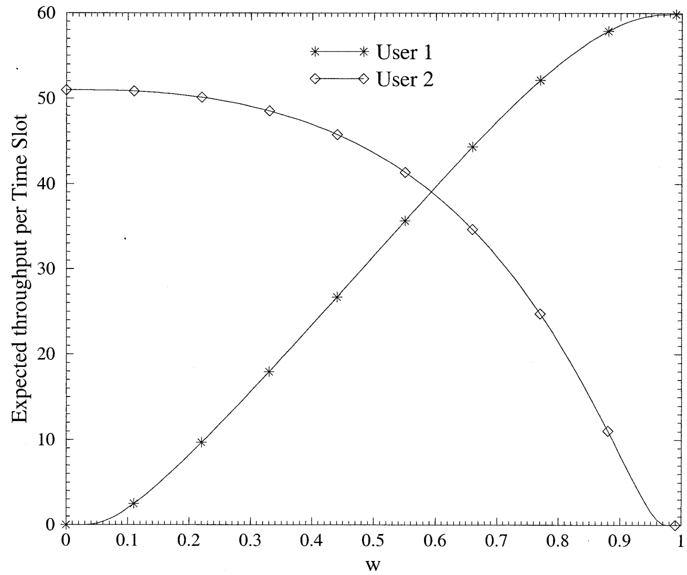

Fig. 1. Normalized expected throughput $\Xi_i(w)$ as function of w.

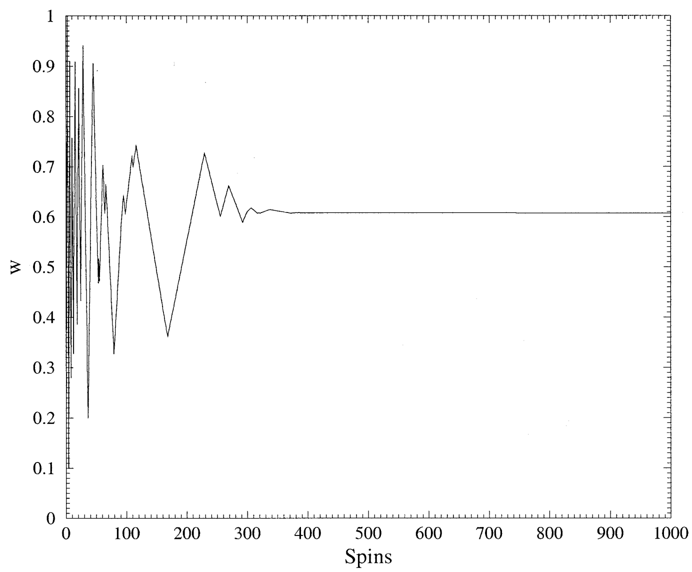

Fig. 2. Price trajectory for two users over 1000 slots.

$(\gamma_1, \gamma_2) = (0.02, 0.01)$. Thus, the feasible rate for user 2 is about twice as large in distribution as for user 1. The throughput target for user 2 is also set twice as large as for user 1, i.e., $(\alpha_1, \alpha_2) = (1, 2)$.

The values of $\Xi_i(w)$ for these parameters as a function of w are plotted in Fig. 1. From this figure, we see that the optimal price is $w^* \approx 0.6$, which may be more precisely determined as $w^* \approx 0.593$ using bisection.

We ran the control algorithm described in Section VI for 1000 slots. We used step sizes $\delta_{k+1} = \rho^k \delta_1$, with initial value $\delta_1 = 0.5$ and reduction factor $\rho = 0.9$. The resulting price trajectories are graphed in Fig. 2 for a period of 1000 slots. Observe that the prices converge to the optimal values in roughly 300 slots, which corresponds to about 0.3 s of operation.

We repeated the above experiment for nongeometric step sizes $\delta_k = \delta_1 k^{-\beta}$, with β successively chosen as 1.5, 2.0,

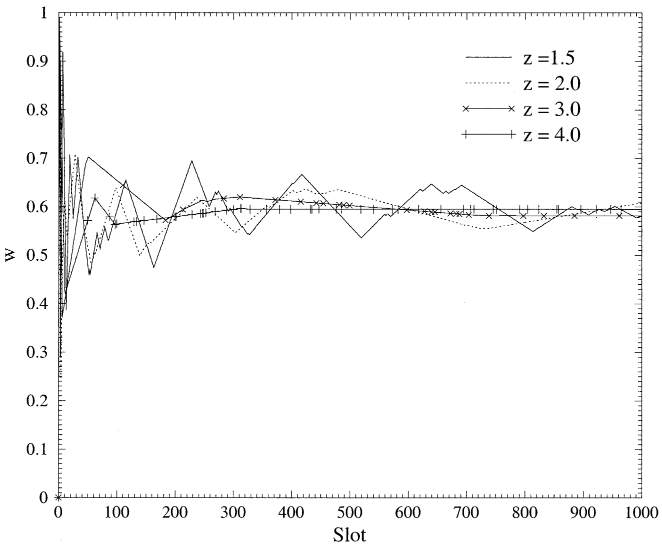

Fig. 3. Price trajectories for two users versus w^* (nongeometric step sizes).

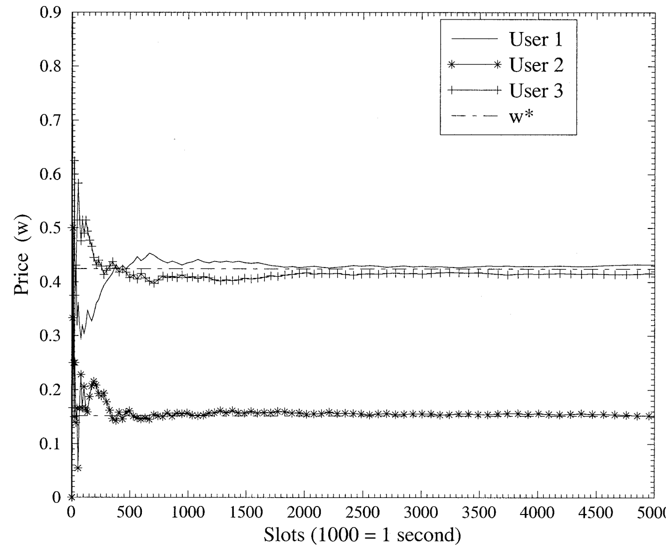

Fig. 4. Price trajectories for three users over 5000 slots versus w^* (move-to-average algorithm).

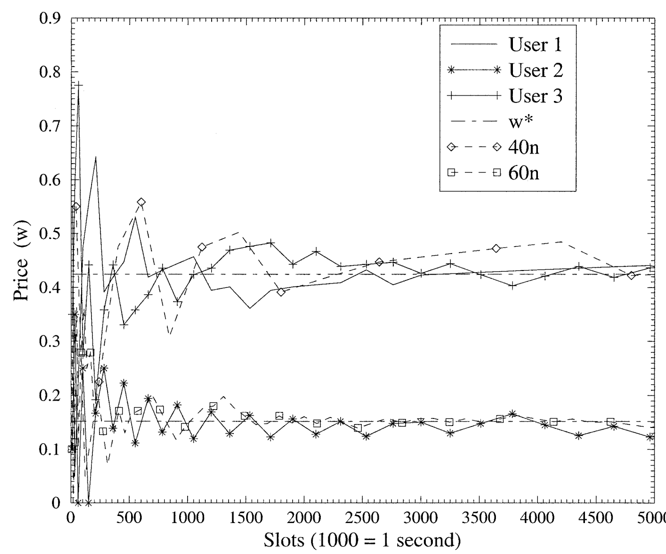

Fig. 5. Price trajectories for three users over 5000 slots versus w^* (update-extreme algorithm).

3.0, and 4.0. Note that the sum of the price changes is still convergent, although the step sizes decay more slowly than before. The corresponding price trajectories are shown in Fig. 3 for a period of 1000 slots. We see that convergence is considerably slower for smaller values of β, i.e., slower decay of the step sizes.

B. Three Users

In the second experiment, we consider a scenario with three users. As before, the feasible rate for user i follows a conditional exponential distribution on the interval [10, 400] with parameters $(\gamma_1, \gamma_2, \gamma_3) = (0.02, 0.01, 0.02)$. Thus, the feasible rate for user 2 is about twice as large in distribution as for users 1 and 3.

The target throughput ratios for the three users are set equal, i.e., $(\alpha_1, \alpha_2, \alpha_3) = (1, 1, 1)$. The optimal-revenue vector is $w^* \approx (0.424, 0.152, 0.424)$, as may be determined using numerical integration and two-dimensional bisection. Observe that the optimal price for users 1 and 3 is higher than for user 1, as is required in order to obtain equal throughput since the feasible rate for user 2 is stochastically larger.

We ran the two control algorithms described in Section VII for 5000 slots, or approximately 5 s of operation, with $L(n) = 10n$ slots for the nth update. This amounts to roughly 30 price updates. The initial revenue vector is set to $w(1) = (0.3, 0.6, 0.1)$. We used step sizes $\delta_k = k^{-2}$, $k = 1, 2, \ldots$. The resulting price trajectories are depicted as the solid curves in Figs. 4 and 5. The revenue vector for the update-extreme algorithm after 30 price updates is $w(30) \approx (0.441, 0.123, 0.436)$, quite close to the optimal one.

We repeated the above experiment for the update-extreme algorithm using $40n$ and $60n$ slots for the nth update, with the same power series for δ_k. The corresponding price trajectories are reproduced as the the dashed lines in Fig. 5 for user 1 in the first case and user 2 in the second (with similar results for the remaining prices.) As expected, we see that using fewer samples per price update leads to a slower and "noisier" convergence to the optimal-revenue vector w^*.

C. Eight Users

In the third experiment, we consider a situation with eight users. As before, the feasible rate for user i follows a conditional exponential distribution on the interval [10, 400]. The exponents were chosen uniformly at random in [0.01, 0.05] and turned out to be approximately (0.0489, 0.0263, 0.0139, 0.0480, 0.0220, 0.0107, 0.0461, 0.0128).

The target throughput ratios are again set equal for all users. As before, we expect that a larger value of the exponent γ, inducing smaller feasible rates, requires a higher price in order to obtain equal throughput.

We ran the two control algorithms described in Section VII for 15 000 slots, or approximately 15 s of operation, with $L(n) = 30n$ slots for the nth update. This amounts to roughly 55 price updates. The initial revenue vector is set at random. We used step sizes $\delta_k = k^{-2}$, $k = 1, 2, \ldots$. The resulting price trajectories are graphed in Figs. 6 and 7.

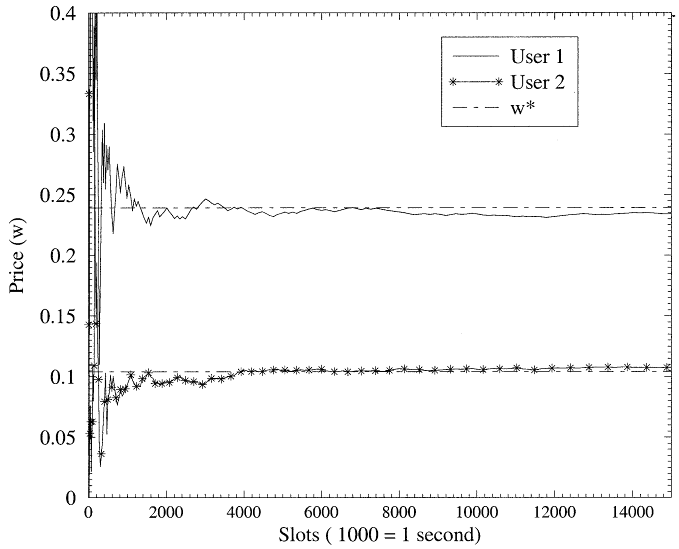

Fig. 6. Price trajectories for eight users with 15 000 slots versus w^* (move-to-average algorithm).

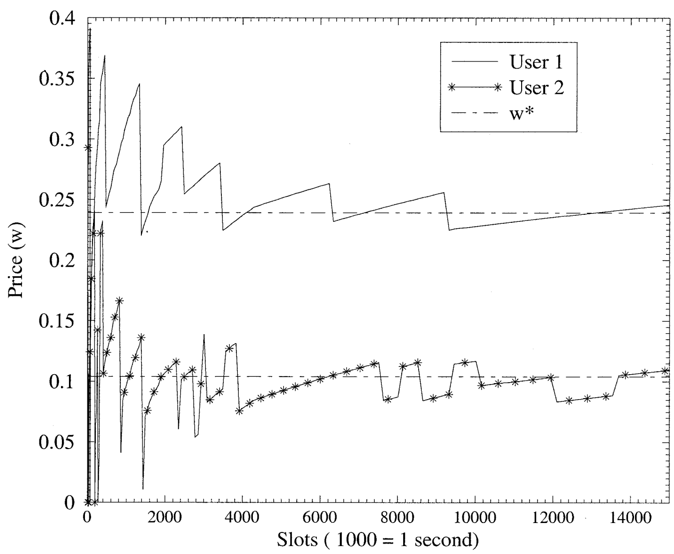

Fig. 7. Price trajectories for eight users with 15 000 slots versus w^* (update-extreme algorithm).

D. Discrete Rates Driven by a Fading Process

We now consider a case with discrete rates governed by independent fading processes, as described by Jakes' model [9]. The mean received powers of user 1, 2, and 3 are -15.0 dB, 0.0 dB, and -10.0 dB, respectively. The feasible rates per slot then follow from Table I, using fading realizations as shown in Fig. 8.

The throughput target for user 2 is set twice as large as for users 1 and 3, i.e., $(\alpha_1, \alpha_2, \alpha_3) = (1, 2, 1)$. We ran the two control algorithms described in Section VII for 10 000 slots, with $L(n) = n$ slots for the nth update. We used step sizes $\delta_k = k^{-3/2}$ and $\delta_k = k^{-2}$, $k = 1, 2, \ldots$.

As explained earlier, the discrete rate values are perturbed by adding a small uniformly distributed random variable to obtain

TABLE I
FEASIBLE RATE PER SLOT AS FUNCTION OF SNR

Signal-to-Noise Ratio (dB)	Rate (bits)
$-5.0 <$ SNR	1000
$-10.0 <$ SNR ≤ -5.0	500
$-20.0 <$ SNR ≤ -10.0	250
$-30.0 <$ SNR ≤ -20.0	100
SNR ≤ -30.0	30

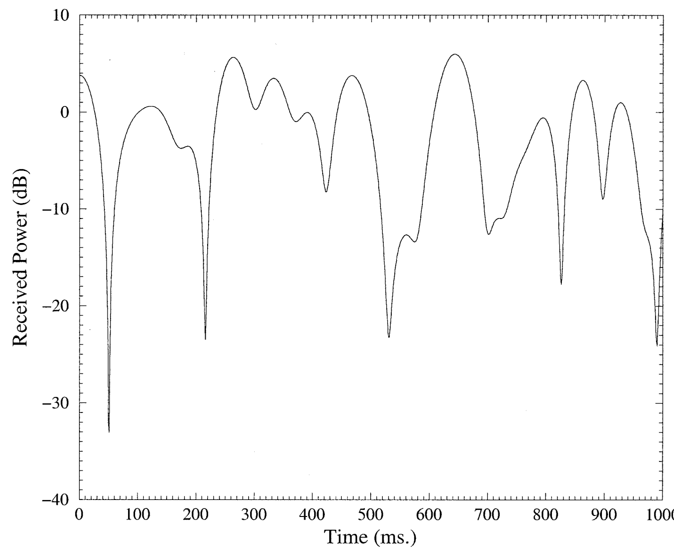

Fig. 8. Fading process with unit power.

a continuous version of the problem. Thus, we ensure that the optimal control algorithm is determined by the revenue vector only.

The empirical average throughputs are depicted in Figs. 9 and 10. The achieved throughputs under the update-extreme algorithm are approximately 130 bits per slot for both users 1 and 3 and 270 bits per slot for user 2, quite close to the target ratios. Under the move-to-average algorithm, the realized throughputs are reasonably close to the target ratios too, provided the step size is reduced sufficiently slowly, as in Fig. 9.

The corresponding price trajectories are displayed in Figs. 11 and 12. We see that under the update-extreme algorithm, the prices converge to the optimal values in about 5 s. Under the move-to-average algorithm, the prices converge fairly quickly too, unless the step size is reduced so quickly that the process gets essentially overdamped.

E. Comparison With a Forcing Scheme

We now compare the revenue-based algorithms with a forcing scheme. The forcing scheme assigns the nth transmission slot to the user $m^*(n)$ with the current minimum normalized throughput, i.e.,

$$m^*(n) = \arg\min_{m=1,\ldots,M} \frac{y_m(n)}{\alpha_m}.$$

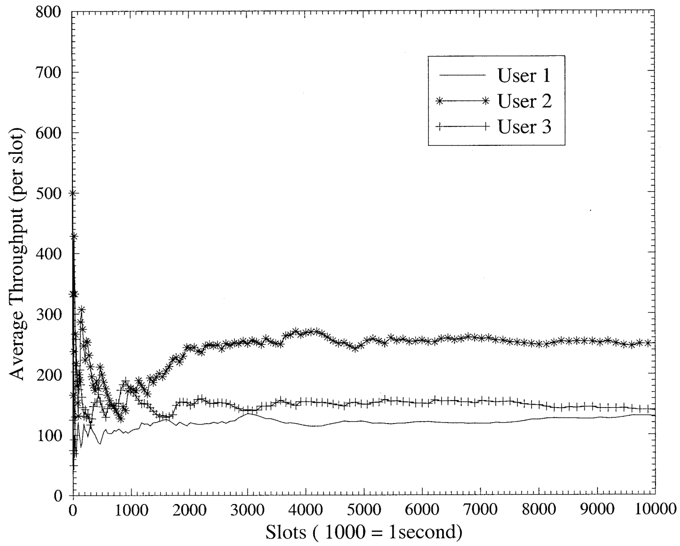

Fig. 9. Empirical average throughput for three users over 10 000 slots (move-to-average algorithm with $\delta_k = k^{-3/2}$).

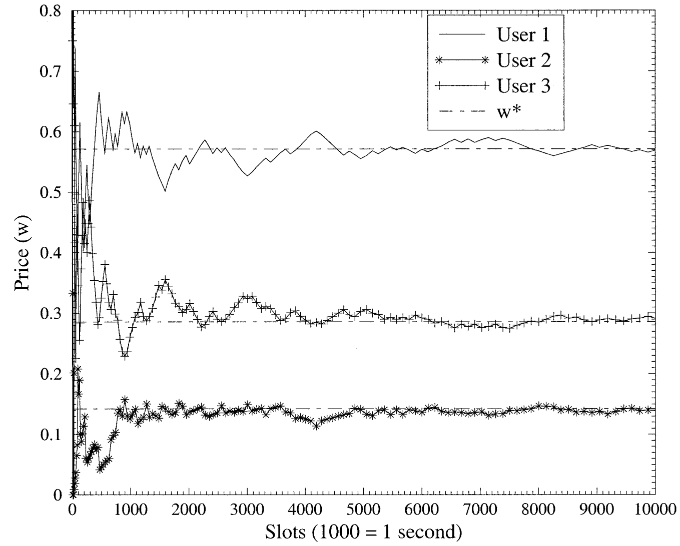

Fig. 11. Price trajectories for three users over 10 000 slots (move-to-average algorithm with $\delta_k = k^{-3/2}$).

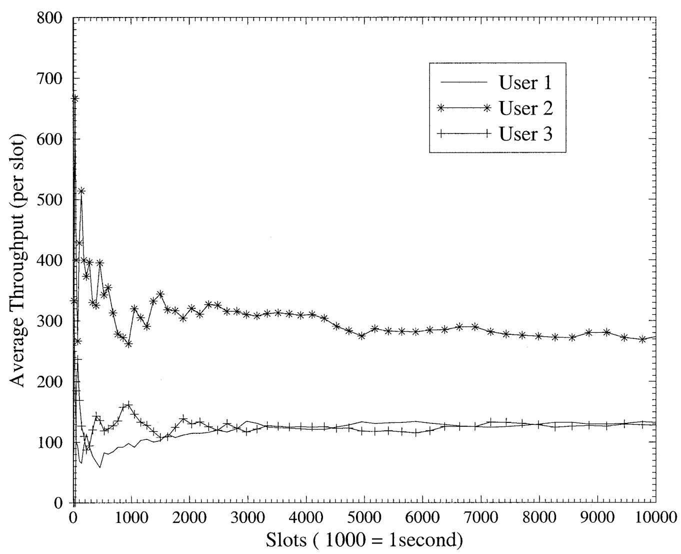

Fig. 10. Empirical average throughput for three users over 10 000 slots (update-extreme algorithm with $\delta_k = k^{-2}$).

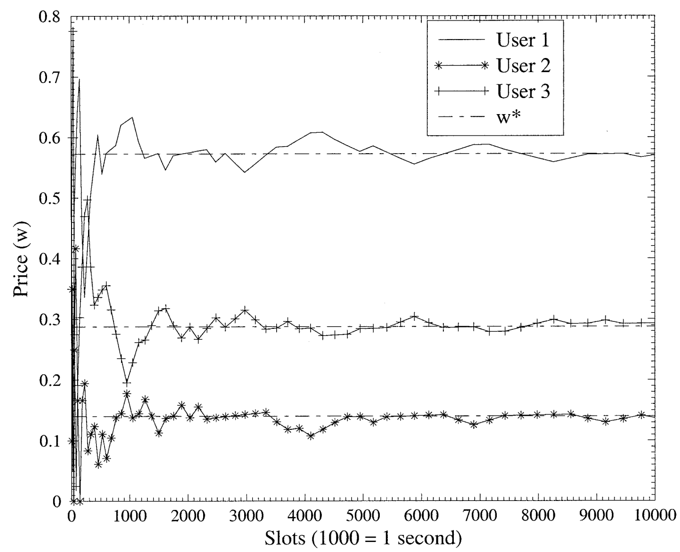

Fig. 12. Price trajectories for three users over 10 000 slots (update-extreme algorithm with $\delta_k = k^{-2}$).

By construction, the forcing scheme realizes the target throughput ratios perfectly, in the sense that wp 1

$$\frac{y_i(N)}{y_j(N)} \to \frac{\alpha_i}{\alpha_j}, \quad \text{as } N \to \infty \tag{20}$$

for all pairs of users $i, j = 1, \ldots, M$.

The downside of the forcing scheme, of course, is that it generally achieves lower throughput in absolute terms, as it does not take advantage of the variations in feasible rates.

Under independent identically distributed (i.i.d.) assumptions, the throughput obtained under the forcing scheme may in fact be computed in closed form as follows. The decision as to whether or not the nth slot is assigned to user i is determined entirely by the normalized cumulative throughputs, which only depend on the feasible rates in previous slots. Under i.i.d. assumptions, the feasible rate for user i in the nth slot is independent of the feasible rates in previous slots. Hence, the decision variable $X_i(n)$ is independent of the feasible rate $R_i(n)$, so that

$$\mathbb{E}[Y_i(n)] = \mathbb{E}[X_i(n)]\mathbb{E}[R_i(n)] = \mathbb{E}[X_i(n)]\mathbb{E}[R_i]$$

and, thus

$$\mathbb{E}[y_i(N)] = \mathbb{E}\left[\sum_{n=1}^{N} \frac{Y_i(n)}{N}\right] = p_i(N)\mathbb{E}[R_i] \tag{21}$$

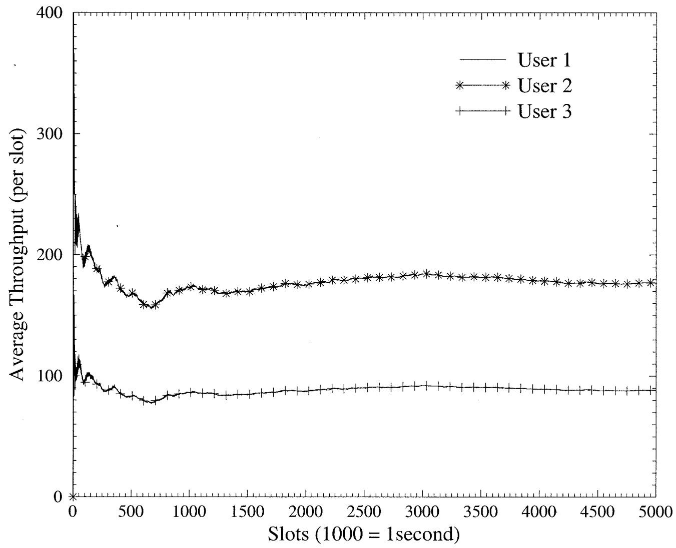

Fig. 13. Empirical average throughput for three users over 5000 slots (forcing algorithm).

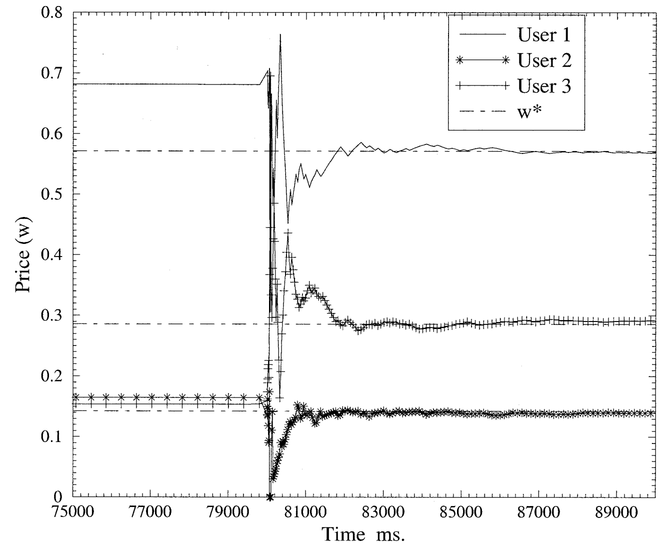

Fig. 14. Price adjustment to allow for data burst for user 3 (move-to-average algorithm).

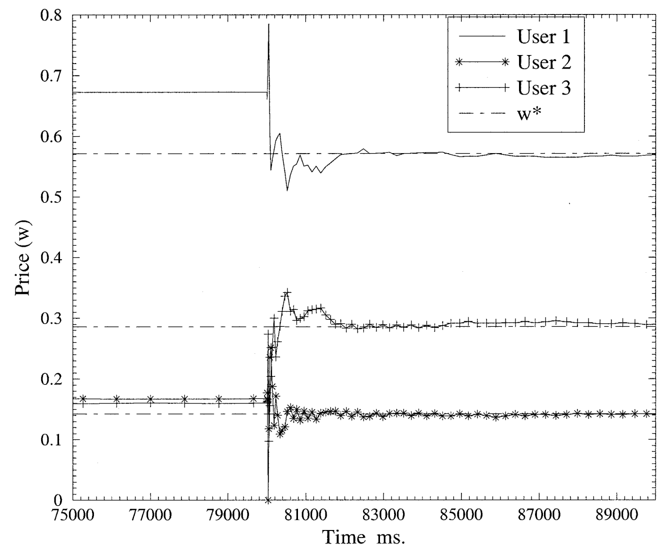

Fig. 15. Price adjustment to allow for data burst for user 3 (update-extreme algorithm).

with $p_i(N) := \mathbb{E}[\sum_{n=1}^{N} X_i(n)/N]$ denoting the expected fraction of slots assigned to user i out of the first N slots. Combining (20) and (21), we conclude

$$\frac{p_i(N)}{p_j(N)} \rightarrow \frac{\frac{\alpha_i}{\mathbb{E}[R_i]}}{\frac{\alpha_j}{\mathbb{E}[R_j]}}, \quad \text{as } N \rightarrow \infty$$

for all pairs of users $i, j = 1, \ldots, M$. Using the identity $\sum_{j=1}^{M} p_j(N) = 1$, we obtain $p_i(N) \rightarrow K\alpha_i/\mathbb{E}[R_i]$ and $y_i(N) \rightarrow \alpha_i K$ as $N \rightarrow \infty$ with $K^{-1} = \sum_{j=1}^{M} \alpha_j/\mathbb{E}[R_j]$.

We repeated the experiment of the previous subsection for the forcing scheme. The empirical average throughputs are reproduced in Fig. 13 for a period of 5000 slots. The achieved throughputs are approximately 90 bits per slot for both users 1 and 3 and 180 bits per slot for user 2. The results show how tightly the target throughput ratios are maintained under the forcing scheme. In absolute terms, however, the throughput for all users is about 30% smaller than for the revenue-based algorithms.

F. Tracking Capability

We now examine how well the algorithms are able to track sudden changes in the target throughput ratios or channel conditions.

In the first experiment, the throughput target for user 3 is initially set to some low value. After 80 s, the throughput target is suddenly incremented to allow for the transmission of a data burst for user 3.

The resulting price trajectories are plotted in Figs. 14 and 15. The optimal price values for the new throughput ratios are also indicated as dashed straight lines. The results show that, after a few oscillations, the prices quickly settle down to the new optimal values.

In the final set of experiments, the control is "cycled" approximately every 5 s. To test the tracking capability, the mean received SNR of user 3 is lowered at a rate of 5 dB/s for 5 s. This is expected to lead to a rapid change in w^*. The change in SNR is initiated after 15 s of simulation time and stopped 5 s later.

The results for the move-to-average algorithm are depicted in Figs. 16 and 17. Similar results for the update-extreme algorithm are displayed in Figs. 18 and 19.

In the first from each of these two pairs of graphs, the size of the price adjustment varies according to $\delta_k = k^{-2}$; in the second, it varies according to $\delta_k = k^{-3/2}$. Thus, it is expected that the control will converge more slowly in the former case and that the results confirm this. Indeed, with $\delta_k = k^{-2}$, convergence to the new price occurs only after about 25 s. In the latter case, the correct price is approached shortly after 20 s, but there are stronger fluctuations around the optimal price.

A more subtle observation is that in the interval where the power is being changed, the price adjustment remains fairly large, which is an advantage conferred by the reset conditions that we used. Standard control algorithms such as Robbins-Monro, in contrast, prescribe such adjustments in advance, see [13] and [15]. It should be stressed that no attempt

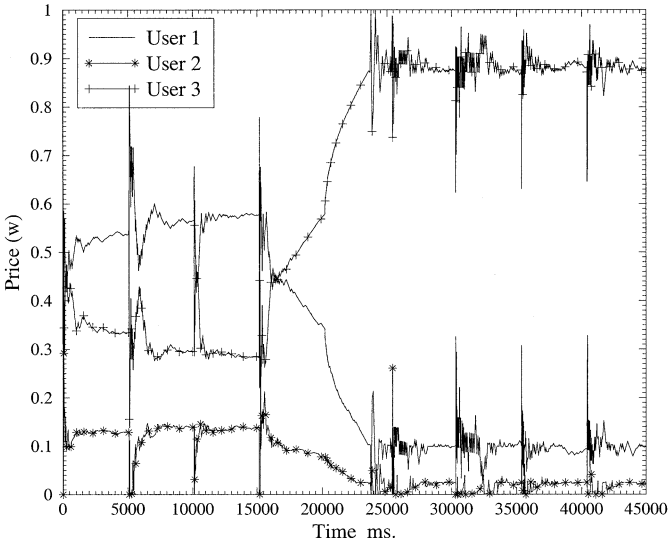

Fig. 16. Cycled control: lowered SNR, user 3 (move-to-average algorithm with $\delta_k = k^{-2}$).

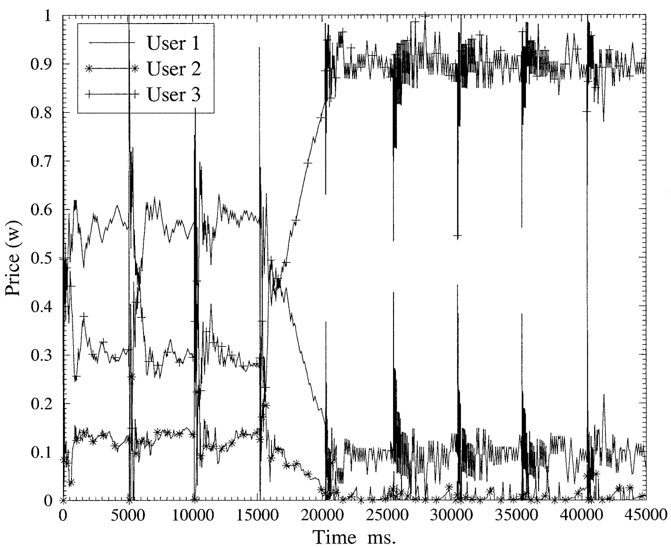

Fig. 17. Cycled control: lowered SNR, user 3 (move-to-average algorithm with $\delta_k = k^{-3/2}$).

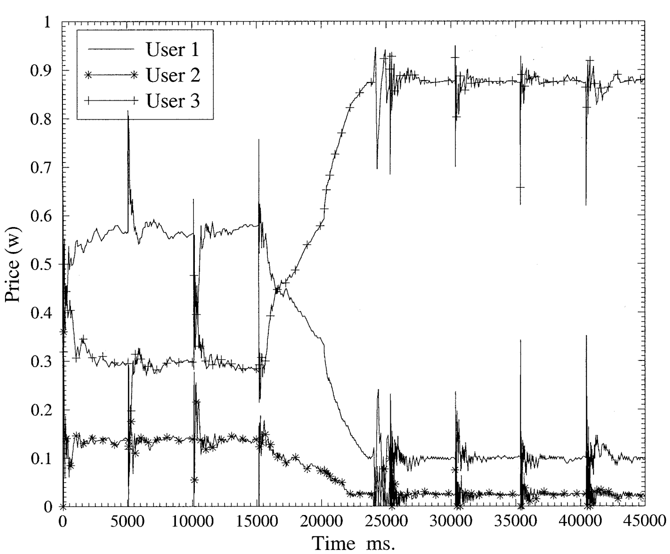

Fig. 18. Cycled control: lowered SNR, user 3 (update-extreme algorithm with $\delta_k = k^{-2}$).

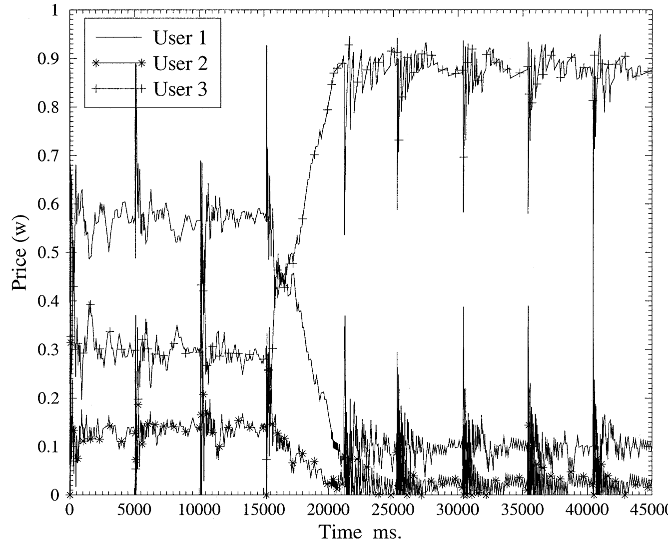

Fig. 19. Cycled control: lowered SNR, user 3 (move-to-average, with $\delta_k = k^{-1.5}$).

has been made here to design the sequences $L(n)$, δ_k, or the cycle interval in an optimal way. Also note that the control signal could be filtered to remove high-frequency components if necessary.

IX. Conclusion

We considered the problem of scheduling data users with varying channel conditions so as to obtain the optimal long-run throughputs for given target ratios. We have shown that the problem may be solved by selecting users for transmission according to an optimal-revenue vector w^*, which balances the expected throughputs. We presented a wide class of stochastic control algorithms that ensure almost-sure convergence to w^* and, thus, achieve the optimal long-run throughputs. The algorithms require only a convergent sequence of step sizes to be specified, in combination with an increasing sequence of sample sizes per price update.

Numerical experiments showed that the convergence to the optimal-revenue vector is, in practice, quite rapid (of the order of a few seconds), making the algorithms suitable for the IS-856 system. In addition, the results demonstrated that the algorithms have the ability to track changes in the channel conditions and throughput targets. Further experiments are required to determine which form of the algorithm is most adequate for implementation in the IS-856 system. The algorithms may also be enhanced by allowing the step sizes or the sample sizes to be adapted in response to nonstationary changes in the feasible rate declarations.

Since the control algorithms require only observations of the feasible rate, they may be used for admission-control purposes. This is reminiscent of *channel probing*, with the additional benefit that the prospective user need not be allocated any resources until the admission-control decision has been made.

In the present paper, we considered a scenario with only one user scheduled at a time and a single-rate sample per user per slot. These conditions, however, are actually not essential for the underlying optimality principle to apply. Revenue-based policies, which balance the throughputs, continue to be optimal in situations where several users may be scheduled at a time and various auxiliary decisions may be taken.

As an illustrative example, consider a throughput optimization problem for two adjacent base stations. Let R_m be the rate in a given slot for user m in cell 1 if both base stations transmit and let R'_m be the rate for user m if only base station 1 transmits. Let R_l and R'_l be defined similarly as the rate in a given slot for user l in cell 2. A revenue-based policy then selects the decision, which maximizes revenue over all feasible options as follows:

$$\text{Revenue} = \begin{cases} w_m^* R_m + w_l^* R_l, & m \in 1, l \in 2 \\ w_m^* R'_m, & m \in 1 \\ w_l^* R'_l, & l \in 2. \end{cases}$$

Observe that the decisions as to which users are scheduled and which base station transmits (1, 2, or both) are taken jointly. The revenue vector w^*, which balances the throughputs will be optimal and may be found by using the stochastic control algorithms as before. This approach may also be used in conjunction with antenna systems, for example.

APPENDIX I
LARGE-DEVIATIONS ASSUMPTION

In this appendix, we show that Large-Deviations Assumption 7.1 is satisfied for the feasible-rate process that we consider.

Given a price vector $w \in \mathcal{W}$, consider a closed neighborhood $N_\zeta(w)$ of w. Let $X_m(w')$ be a random variable representing the throughput per slot that user m receives under the price vector w' in stationarity. Then $X_m(w')$ may be formally represented as

$$X_m(w') = R_m I_{\{w'_m R_m = \max_{k=1,\ldots,M} w'_k R_k\}}$$

with $(R_1, \ldots, R_M)$ a random vector with distribution the joint stationary distribution of the feasible rates.

Now define random variables

$$Y_m\big(N_\zeta(w)\big) = R_m I_{\{\forall w' \in N_\zeta(w): w'_m R_m = \max_{k=1,\ldots,M} w'_k R_k\}}.$$

Thus, $Y_m\big(N_\zeta(w)\big)$ represents the rate that user m would receive in the case it were selected only if it has the maximum rate-reward product under *all prices* $w' \in N_\zeta(w)$. Evidently, $X_m(w') \geq Y_m\big(N_\zeta(w)\big)$ for all $w' \in N_\zeta(w)$.

Similarly, define random variables

$$Z_m\big(N_\zeta(w)\big) = R_m I_{\{\exists w' \in N_\zeta(w): w'_m R_m = \max_{k=1,\ldots,M} w'_k R_k\}}.$$

Thus, $Z_m\big(N_\zeta(w)\big)$ represents the rate that user m would receive in the case it were assigned the slot if it has the maximum rate-reward product under *some price* $w' \in N_\zeta(w)$. Obviously, $X_m(w') \leq Z_m\big(N_\zeta(w)\big)$ for all $w' \in N_\zeta(w)$.

Denote by $\mathbb{E}_\pi[Y_m\big(N_\zeta(w)\big)]$ and $\mathbb{E}_\pi[Z_m\big(N_\zeta(w)\big)]$ the respective expectations under the stationary distribution $\pi(s)$, $s \in \mathcal{S}$, of the Markov chain governing the feasible-rate process. By dominated convergence

$$\mathbb{E}_\pi[Y_m\big(N_\zeta(w)\big)] \leq \Xi_m(w) \leq \mathbb{E}_\pi[Z_m\big(N_\zeta(w)\big)]$$

for all $m = 1, \ldots, M$, with $\mathbb{E}_\pi[Y_m\big(N_\zeta(w)\big)] \uparrow \Xi_m(w)$ and $\mathbb{E}_\pi[Z_m\big(N_\zeta(w)\big)] \downarrow \Xi_m(w)$ as $\zeta \downarrow 0$.

Let $X_m^n(w')$, $Y_m^n\big(N_\zeta(w)\big)$, and $Z_m^n\big(N_\zeta(w)\big)$ be the throughput per slot obtained by user m in a sample period of length n under the above three rules. For any $w' \in N_\zeta(w)$, sample path wise, $Y_m^n\big(N_\zeta(w)\big) \leq X_m^n(w') \leq Z_m^n\big(N_\zeta(w)\big)$, so in particular

$$\mathbb{P}\{|\, X_m^n(w') - \Xi_m(w)\, | > \xi\} \leq \mathbb{P}\{Y_m(N_\zeta(w)) < \Xi_m(w) - \xi\} + \mathbb{P}\{Z_m\big(N_\zeta(w)\big) > \Xi_m(w) + \xi\}. \quad (22)$$

Denote by

$$\varphi_m^n(nt) := \log \mathbb{E}_\pi[\, e^{tZ_m^n\left(N_\zeta(w)\right)}]$$

the log-moment generating function of $Z_m^n\big(N_\zeta(w)\big)$. Define

$$\varphi_m(t) := \lim_{n\to\infty} \frac{1}{n} \varphi_m^n(nt).$$

We have that

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\{Z_m^n\big(N_\zeta(w)\big) > nx\} = I_m(x)$$

with

$$I_m(x) := \sup_t \{tx - \varphi_m(t)\}.$$

We now compute $I_m(x)$. For any $s \in \mathcal{S}$, denote

$$\phi_m(t,s) = \mathbb{E}_\pi[\, e^{tZ_m\left(N_\zeta(w)\right)} \mid s]$$

as the log-moment generating function of $Z_m\big(N_\zeta(w)\big)$, conditional on the state s of the Markov chain governing the feasible-rate process, and define the $\mathcal{S} \times \mathcal{S}$-matrix

$$\Pi_m(t) = \{Q(s_1,s_2)\phi_m(t,s_1)\}_{s_1,s_2}$$

with Q the $\mathcal{S} \times \mathcal{S}$ transition matrix of the Markov chain.

It may be then shown that

$$\varphi_m^n(nt) = \log \sum_{s\in\mathcal{S}} (\Pi_m(t))^{n-1} \pi(s) \phi_m(t,s)$$

(see Dembo and Zeitouni [8]).

Hence

$$\varphi_m(t) = \log \rho_1\big(\Pi_m(t)\big)$$

with ρ_1 the Perron–Frobenius eigenvalue of the matrix $\Pi_m(t)$ so that

$$I_m(x) = \sup_t \{tx - \log \rho_1\big(\Pi_m(t)\big)\}.$$

It remains to be shown that $I_m(x) > 0$ for $x > \mathbb{E}_\pi[Z_m\big(N_\zeta(w)\big)]$.

Since $\Pi_m(t)$ is a compact family of nonnegative matrices, we have

$$\big(\Pi_m(t')\big)^n \rho_1^{-n} - l(t') r^T(t') \to 0$$

component wise and uniformly for all $t' \in [0,t]$, with $l(t')$ and $r(t')$ the left and right Perron eigenvectors, normalized such that $\sum_{s\in\mathcal{S}} l_s(t') = 1$ and $\sum_{s\in\mathcal{S}} r_s(t') = 1$ (see Seneta [17, Theorem 3.6]).

Thus, $\varphi_m(t)$ may be uniformly approximated by $(1/n)\varphi_m^n(nt)$: for any given $t > 0$, $\psi > 0$, there exists an n such that

$$\left|\, \frac{1}{n} \varphi_m^n(nt') - \varphi_m(t')\, \right| < \psi$$

for all $t' \in [0,t]$.

Hence

$$\sup_t\{tx-\log\rho_1(\Pi_m(t))\}\geq\max_{0\leq t'\leq t}\{t'x-\log\rho_1(\Pi_m(t'))\}$$
$$\geq\max_{0\leq t'\leq t}\{t'x-\frac{1}{n}\varphi_m^n(nt')-\psi\}=\max_{0\leq t'\leq t}\gamma_n(t',\psi).$$

However, since all moments exist, $\varphi_m^n(t')$ may be expanded to third order around 0, using Taylor's theorem as follows:

$$\frac{1}{n}\varphi_m^n(nt')=t'(\mathbb{E}_\pi Z_m(N_\zeta(w))-\varpi_n)+\frac{t'^2}{2}\sigma_n^2+\rho_n\frac{t'^3}{6}$$

with $\lim_{n\to\infty}\varpi_n=0$, $\liminf_{n\to\infty}\sigma_n^2>0$, and $\rho_n\leq K<\infty$.

For $x=\mathbb{E}_\pi[Z_m(N_\zeta(w))]+\epsilon$, we may take $t^*=(\epsilon-\varpi_n)/\sigma_n^2$. If n is sufficiently large and ϵ sufficiently small, then $\gamma_n(t^*,\psi)>0$ and, hence, $I(x)>0$ for $x>\mathbb{E}_\pi[Z_m(N_\zeta(w))]$ due to monotonicity in x. It follows that there exist numbers $C_m^\epsilon(N_\zeta(w),Z)>0, D_m^\epsilon(N_\zeta(w),Z)>0$, such that

$$\mathbb{P}\{Z_m(N_\zeta(w))>\mathbb{E}_\pi[Z_m(N_\zeta(w))]+\epsilon\}\leq C_m^\epsilon(N_\zeta(w),Z)\,e^{-nD_m^\epsilon(N_\zeta(w),Z)}. \quad (23)$$

Similarly,

$$\mathbb{P}\{Y_m(N_\zeta(w))<\mathbb{E}_\pi[Y_m(N_\zeta(w))]-\epsilon\}\leq C_m^\epsilon(N_\zeta(w),Y)\,e^{-nD_m^\epsilon(N_\zeta(w),Y)}. \quad (24)$$

Now take $\epsilon=\xi/2$ and $\zeta>0$ small enough so that $\mathbb{E}_\pi[Y_m(N_\zeta(w))]\geq\Xi_m(w)-\xi/2$, $\mathbb{E}_\pi[Z_m(N_\zeta(w))]\leq\Xi_m(w)+\xi/2$, $C_m^\xi(w):=2\Big(C_m^{\xi/2}(N_\zeta(w),Y)+C_m^{\xi/2}(N_\zeta(w),Z)>0\Big)$ and $D_m^\xi(w):=\min\Big\{D_m^{\xi/2}(N_\zeta(w),Y),D_m^{\xi/2}(N_\zeta(w),Z)\Big\}>0$.

Combining (22), (23), and (24), we then obtain

$$\mathbb{P}\{|\,X_m^n(w')-\Xi_m(w)\,|>\xi\}\leq C_m^\xi(w)\,e^{-D_m^\xi(w)n}$$

as required.

Appendix II
Function $T(\cdot)$

In this appendix, we prove that Assumption 7.3 is satisfied for suitable functions $T(\cdot)$ under certain assumptions on the feasible-rate process

For any subset $S\subseteq\mathcal{R}$, denote by $\mu(S)$ the Lebesgue measure of S and denote by $\pi(S)$ the stationary probability that the feasible rate vector is in S. We assume that there are fixed constants K_1, K_2 such that $K_1\mu(S)\leq\pi(S)\leq K_2\mu(S)$ for all $S\subseteq\mathcal{R}$.

We will prove that Assumption 7.3 is satisfied provided $K_1>0, K_2<\infty$. It may then be shown that there exist $\eta>0,\theta<\infty$ such that if $w'=w+\delta v(w)$, with $v(w)$ as in (8) and(9), then for all $i\in\Omega^-$

$$\Xi_i(w')-\Xi_i(w)\in\delta[\eta,\theta]$$

and for all $j\in\Omega^+$

$$\Xi_j(w')-\Xi_j(w)\in-\delta[\eta,\theta]$$

(see also Lemma 7.1).

Now consider a price vector $w\notin\Gamma_\epsilon$. By definition, if a price direction is ξ-right, then all the ξ below-average users will belong to Ω^- and all the ξ-above-average users will belong to Ω^+. Thus, if $i\in\Omega^-$, then $\Xi_i(w)\leq\Xi_{\rm ave}(w)+\xi$ and if $j\in\Omega^+$, then $\Xi_j(w)\geq\Xi_{\rm ave}(w)-\xi$.

As mentioned earlier, we consider two alternative choices for the function $T(\cdot)$. The first one is

$$T(w)=\Xi_{\max}(w)-\Xi_{\min}(w).$$

Define $\xi:=\min\{\Xi_{\max}(w)-\Xi_{\rm ave}(w),\ \Xi_{\rm ave}(w)-\Xi_{\min}(w)\}/2\geq\epsilon/2M>0$. Then

$$\begin{aligned}\Xi_{\min}(w')&=\min_{m=1,\dots,M}\Xi_m(w')\\&=\min\{\min_{i\in\Omega^-}\Xi_i(w'),\min_{j\in\Omega^+}\Xi_j(w')\}\\&\geq\min\{\min_{i\in\Omega^-}\Xi_i(w)+\delta\eta,\min_{j\in\Omega^+}\Xi_j(w)-\delta\theta\}\\&\geq\min\{\Xi_{\min}(w)+\delta\eta,\Xi_{\rm ave}(w)-\xi-\delta\theta\}\\&\geq\Xi_{\min}(w)+\min\{\delta\eta,\xi-\delta\theta\}\\&\geq\Xi_{\min}(w)+\min\left\{\delta\eta,\frac{\epsilon}{2M-\delta\theta}\right\}\\&\geq\Xi_{\min}(w)+\delta\eta\end{aligned}$$

for $\delta<\epsilon/2M(\eta+\theta)$.

Similarly

$$\begin{aligned}\Xi_{\max}(w')&=\max_{m=1,\dots,M}\Xi_m(w')\\&=\max\{\max_{i\in\Omega^-}\Xi_i(w'),\max_{j\in\Omega^+}\Xi_j(w')\}\\&\leq\max\{\max_{i\in\Omega^-}\Xi_i(w)+\delta\theta,\max_{j\in\Omega^+}\Xi_j(w)-\delta\eta\}\\&\leq\max\{\Xi_{\rm ave}(w)+\xi+\delta\theta,\Xi_{\max}(w)-\delta\eta\}\\&\leq\Xi_{\max}(w)+\max\{-\xi+\delta\theta,-\delta\eta\}\\&\leq\Xi_{\max}(w)+\max\left\{-\frac{\epsilon}{2M}+\delta\theta,-\delta\eta\right\}\\&\leq\Xi_{\max}(w)-\delta\eta\end{aligned}$$

for $\delta<\epsilon/2M(\eta+\theta)$.

Thus

$$\begin{aligned}T(w')=&\Xi_{\max}(w')-\Xi_{\min}(w')\\\leq&\Xi_{\max}(w)-\Xi_{\min}(w)-2\eta\delta=T(w)-2\eta\delta\end{aligned}$$

for all $\delta\in(0,\delta^*)$ with $\delta^*=\epsilon/2M(\eta+\theta)$.

The second choice that we consider is the function

$$T(w)=\sum_{m=1}^M w_m\Xi_m(w).$$

Define $\xi:=\nu\epsilon/4(M-1)$. For convenience, relabel the users such that $\Omega^-:=\{1,\dots,K\}$, with $\Xi_1(w)\leq\cdots\leq\Xi_K(w)$, and $\Omega^+:=\{K+1,\dots,M\}$, with $\Xi_{K+1}(w)\leq\cdots\leq\Xi_M(w)$. Recall that if a price direction is ξ-right, then all ξ-below-average users belong to Ω^- and all ξ-above-average users belong to Ω^+, so that if $i\in\Omega^-$, $j\in\Omega^+$, then

$$\Xi_j(w)-\Xi_i(w)\geq-2\xi.$$

Denote $v_{\rm tot}=\sum_{i=1}^K v_i=\sum_{j=K+1}^M v_j$. It may be easily verified that there exist numbers $u_1,\dots,u_{M-1}\geq 0$, with $\sum_{k=1}^{M-1} u_k = v_{\rm tot}$, and integers $i(k) \in \{1,\ldots,K\}$, $j(k) \in \{K+1,\ldots,M\}$, such that

$$v_i = \sum_{k:i(k)=i} u_k$$

for all $i = 1,\ldots,K$

$$v_j = -\sum_{k:j(k)=j} u_k$$

for all $j = K+1,\ldots,M$.

Without loss of generality, we may assume that $i(1) = 1$, $j(1) = M$, $u_1 = \min\{v(1), v(M)\} \geq \nu > 0$ and that $u_k \leq \max\limits_{m=1,\ldots,M} v_m \leq 1$ for all $k = 1,\ldots,M-1$.

Thus

$$\begin{aligned}
&T(w') \\
&= \sum_{m=1}^{M} w'_m \Xi_m(w') \\
&= \sum_{i\in\Omega^-} w'_i \Xi_i(w') + \sum_{j\in\Omega^+} w'_j \Xi_j(w') \\
&= \sum_{i\in\Omega^-} \left(w_i + \delta \sum_{k:i(k)=i} u_k \right) \Xi_i(w') \\
&\quad + \sum_{i\in\Omega^-} \left(w_j - \delta \sum_{k:j(k)=j} u_k \right) \Xi_j(w') \\
&= \sum_{m=1}^{M} w_m \Xi_m(w') - \delta \sum_{k=1}^{M-1} u_k \left(\Xi_{i(k)}(w') - \Xi_{j(k)}(w') \right) \\
&\leq \sum_{m=1}^{M} w_m \Xi_m(w) - \delta \sum_{k=1}^{M-1} u_k \left(\Xi_{i(k)}(w) - \Xi_{j(k)}(w) - 2\theta\delta \right) \\
&= T(w) - \delta \Big[u_1 \left(\Xi_{i(1)}(w) - \Xi_{j(1)}(w) \right) \\
&\quad - \sum_{k=2}^{M-1} u_k \left(\Xi_{i(k)}(w) - \Xi_{j(k)}(w) \right) - 2(M-1)\,\theta\delta \Big] \\
&\leq T(w) - \delta \left[\nu\epsilon - 2(M-2)\xi - 2(M-1)\theta\delta \right] \\
&\leq T(w) - \delta \left[\nu\epsilon - 2(M-1) \left(\frac{\nu\epsilon}{4(M-1)} + \theta\delta \right) \right] \\
&\leq T(w) - \frac{\delta\nu\epsilon}{4}
\end{aligned}$$

for all $\delta \in (0, \delta^*)$ with $\delta^* = \nu\epsilon / 8\theta(M-1)$.

REFERENCES

[1] R. Agrawal, A. Bedekar, R. J. La, and V. Subramanian, "Class and channel condition based weighted proportional fair scheduler," in *Proc. ITC-17 Teletraffic Engineering in the Internet Era*, S. da Bahia, J. M. de Souza, N. L. S. da Fonseca, and E. A. de Souza e Silva, Eds., Amsterdam, The Netherlands, 2001, pp. 553–565.

[2] D. M. Andrews, K. Kumaran, K. Ramanan, A. L. Stolyar, R. Vijayakumar, and P. A. Whiting, "CDMA data QoS scheduling on the forward link with variable channel conditions," Report 10009626–000404–05TM, Bell Laboratories, Lucent Technologies, Murray Hill, NJ, 2000.

[3] D. M. Andrews, K. Kumaran, K. Ramanan, A. L. Stolyar, R. Vijayakumar, and P. A. Whiting, "Providing quality of service over a shared wireless link," *IEEE Commun. Mag.*, vol. 39, pp. 150–154, Feb. 2001.

[4] N. Bambos, S. C. Chen, and D. Mitra, "Channel probing for distributed access control in wireless communication networks," in *Proc. IEEE Globecom '95*, Singapore, Nov. 1995.

[5] A. Bedekar, S. C. Borst, K. Ramanan, P. A. Whiting, and M. Yeh, "Downlink scheduling in cdma data networks," in *Proc. IEEE Globecom '99*, Rio de Janeiro, Brazil, Dec. 1999, pp. 2653–2657.

[6] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushayana, and A. Viterbi, "CDMA/HDR: A bandwidth-efficient high-speed wireless data service for nomadic users," *IEEE Commun. Mag.*, vol. 38, pp. 70–77, July 2000.

[7] S. C. Borst and P. A. Whiting, "Dynamic rate control algorithms for HDR throughput optimization," in *Proc. IEEE Infocom 2001*, Anchorage, AK, May 2001, pp. 976–985.

[8] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*. Amsterdam, The Netherlands: Jones Barlett, 1992.

[9] W. C. Jakes, "Multipath interference," in *Microwave Mobile Communications*, W. C. Jakes, Ed. Piscataway, NJ: IEEE Press, 1974.

[10] A. Jalali, R. Padovani, and R. Pankaj, "Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system," in *Proc. 51th IEEE Vehicular Technology Conf.*, Tokyo, Japan, Spring 2000, pp. 1854–1858.

[11] N. Kahale and P. E. Wright, "Dynamic global packet routing in wireless networks," in *Proc. IEEE Infocom '97*, Kobe, Japan, Apr. 1997, pp. 1441–1421.

[12] X. Liu, E. K. P. Chong, and N. B. Shroff, "Opportunistic transmission scheduling with resource-sharing constraints in wireless networks," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 2053–2064, Oct. 2001.

[13] M. Metivier, *Semimartingales*, De Gruyter, Ed. Berlin, Germany: Walter de Gruyter, 1982.

[14] A. K. Parekh and R. G. Gallager, "A generalized processor sharing approach to flow control in integrated services networks: The single-node case," *IEEE/ACM Trans. Networking*, pp. 344–357, June 1993.

[15] L. Robbins and S. Monro, "A stochastic approximation method," *Ann. Math. Statist.*, vol. 22, pp. 400–407, 1951.

[16] A. Schrijver, *Theory of Linear and Integer Programming*. Chichester, U.K.: Wiley, 1986.

[17] E. Seneta, *Non-Negative Matrices*. New York: Wiley, 1973.

[18] S. Shakkottai and A. L. Stolyar, "Scheduling for multiple flows sharing a time-varying channel: The exponential rule," Bell Labs, Lucent Tech., Murray Hill, NJ, Rep. 10009626–010102–01TM, 2000.

[19] L. Tassiulas, "Linear complexity algorithms for maximum throughput in radio networks and input queued switches," in *Proc. IEEE Infocom '98*, San Francisco, CA, Mar. 1998, pp. 553–559.

Sem Borst received the M.Sc. degree in applied mathematics from the University of Twente, Enschede, The Netherlands, in 1990 and the Ph.D. degree from the University of Tilburg, Tilburg, The Netherlands, in 1994.

During the fall of 1994, he was a Visiting Scholar at the Statistical Laboratory of the University of Cambridge, Cambridge, U.K. In 1995, he joined the Mathematics of Networks and Systems Department, Bell Laboratories, Lucent Technologies, Murray Hill, NJ, as a Member of the technical staff. Since the fall of 1998, he has been a Senior Member of the Probability, Networks, and Algorithms Department of the Center for Mathematics and Computer Science (CWI), Amsterdam, The Netherlands. He also has a part-time appointment as a Professor of Stochastic Operations Research at Eindehoven University of Technology, Eindehoven, The Netherlands. His main research interests are in the performance evaluation of communication networks and computer systems.

Phil Whiting (M'94) received the M.Sc. degree in probability and statistics from the University of London, U.K., in 1983 and the Ph.D. degree in electronic engineering from the University of Strathclyde, Glasgow, U.K., in 1987. His postdoctoral reasearch was at the Statistical Laboratory of the University of Cambridge, Cambridge, U.K.

In 1993, he participated in the trial of Qualcomm CDMA by Australia Telecom. In 1997, he joined the Mathematics of Networks and Systems Department of Bell Laboratories, Lucent Technologies, Murray Hill, NJ, as a Member of the technical staff. His main research interests are in information theory and the performance evaluation of wireless networks.