Joint Channel Probing and Proportional Fair Scheduling in Wireless Networks

Hui Zhou, Pingyi Fan, Dongning Guo

Abstract

The design of a scheduling scheme is crucial for the efficiency and user-fairness of wireless networks. Assuming that the quality of all user channels is available to a central controller, a simple scheme which maximizes the utility function defined as the sum logarithm throughput of all users has been shown to guarantee proportional fairness. However, to acquire the channel quality information may consume substantial amount of resources. In this work, it is assumed that probing the quality of each user's channel takes a fraction of the coherence time, so that the amount of time for data transmission is reduced. The multiuser diversity gain does not always increase as the number of users increases. In case the statistics of the channel quality is available to the controller, the problem of sequential channel probing for user scheduling is formulated as an optimal stopping time problem. A joint channel probing and proportional fair scheduling scheme is developed. This scheme is extended to the case where the channel statistics are not available to the controller, in which case a joint learning, probing and scheduling scheme is designed by studying a generalized bandit problem. Numerical results demonstrate that the proposed scheduling schemes can provide significant gain over existing schemes.

I. Introduction

Efficient and fair scheduling is important for wireless systems with limited resources and heterogeneous user conditions. A large class of resource allocation schemes with fairness considerations are obtained by maximizing some utility functions of the throughput [1]. In particular, proportional fairness is achieved when the utility is the sum of the logarithm of the users'

H. Zhou and P. Fan are with the Department of Electronic Engineering, Tsinghua University, Beijing, 100084, China (e-mail: h-zhou03@mails.thu.edu.cn; fpy@tsinghua.edu.cn)

D. Guo is with the Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL 60208, U.S.A. (e-mail: dguo@northwestern.edu)

throughput. In existing third generation wireless systems, like EV-DO and HSDPA, proportional fair (PF) scheduling scheme is employed at the base station to schedule downlink traffic to mobile users. The PF scheme strikes a good balance between throughput efficiency and fairness by exploiting the multiuser diversity [2] and the game-theoretic equilibrium [3]. Analysis and applications on PF scheduling have been extensively explored from various aspects due to its favorable performance and low implementation complexity. For example, there have been studies of the convergence and optimality [4], stability [5], throughput [6] and capacity region [7] of PF scheduling.

Most previous work on PF scheduling assume that the instantaneous channel quality information (CQI) of all users is known to the scheduler at no cost. In practice, however, acquiring the CQI often consumes a significant amount of resources in terms of time, bandwidth and power. It is important to understand the impact of the cost when the number of users is large, because the cost may scale linearly with the user population. The goal of this work is to answer the following two questions: 1) to what extent will the CQI acquisition affect the scheduling? and 2) how to probe and schedule the users to achieve the best performance with proportional fairness?

There have been related works on the impact of the channel uncertainty on the communication systems. The loss of throughput caused by poor estimates of channel quality is quantified in [8]. Joint channel probing and user scheduling has also been addressed recently. Several schemes with the objective of maximizing the system throughput have been designed in [9]–[12]. And the authors of [13]–[15] propose schemes for stabilizing the queues and characterize the network throughput region. In contrast to the preceding works, the goal of this paper is to design a proportional fair scheduling scheme which takes into account the cost of channel probing. Our previous work [16] has shown the scheme and its performance roughly. In this paper, we not only present the derivation of the scheme with rigorous arguments, but also show its asymptotic behavior and the optimality with theoretical rigor. In addition, the scheme is extended to a more generalized scenario. The organization and main contributions of this work are as follows:

- Section II describes the network model.

- In Section III, we assume the prior distribution of CQI is known to the scheduler, and formulate the problem of sequentially probing user channels to make scheduling decision as a stopping time problem. A simple scheme based on maximizing the sum logarithm throughput of all users is shown to guarantee proportional fairness and convergence. The scheduling gain of the scheme is determined analytically. Further reduction of computational complexity is also discussed.
- In Section IV, the statistics of the CQI is assumed not to be available to the scheduler. The problem is formulated as a generalized bandit problem, and a joint learning, probing and scheduling scheme is proposed.
- In Section V, significant advantages of the proposed schemes are demonstrated using numerical experiments. In typical scenarios where the statistics of the CQI are not available, the joint learning, probing and scheduling scheme achieves almost the same performance as that in the case where the statistics are known.

II. The Network Model

Consider a wireless system with one controller and K users with time-varying channel quality, such as in the downlink of a cellular system. Let time be divided into unit-length slots and only one user can be served in each slot. As in most related work (e.g., [4] and [6]), the transmit power is assumed to be fixed so that dynamic power allocation is not considered. Thus the achievable rate is only determined by the instantaneous channel quality. Moreover, we assume saturated traffic for all users.

Assume slow fading, where the duration of a slot is much shorter than the channel coherence time, so that the channel quality remains constant during each slot. We make the following *homogeneous rate assumption* that the rate of each user normalized by its mean value follows the same distribution:

(A1) Let $X_1, \ldots, X_K$ be independent identically distributed (i.i.d.) non-negative random variables with unit mean value. Let $r_1, \ldots, r_K \geq 0$ be constants. Let $R_k = r_k X_k$ for $k = 1, \ldots, K$.

The achievable rates $\{R_k(n)|k = 1, \dots, K; n = 1, 2, \dots\}$ are independent. For every user k, the rates over the time slots, $R_k(1), R_k(2), \dots$, are i.i.d. following the same distribution as that of R_k. Clearly, $\mathbb{E}R_k(n) = r_k$.

The instantaneous achievable rates of all users are not known *a priori*. During each slot n, to obtain the achievable rate $R_k(n)$ requires the scheduler to probe the channel of user k using a fraction β of the slot. Let $I_k(n)$ be an indicator of the event that user k is scheduled for transmission in slot n. Let $J(n)$ denote the number of probed users in slot n. The amount of data transmitted to or by user k during slot n is $B_k(n) = (1 - J(n)\beta)R_k(n)I_k(n)$, which is nonzero for only one user during each slot. The throughput of user k averaged over n slots is thus

$$T_k(n) = \frac{1}{n}\sum_{j=1}^{n} B_k(j). \tag{1}$$

III. Joint Probing and Scheduling with Known Channel Statistics

In this section, we consider the case where the statistics of $\boldsymbol{R} = [R_1, \dots, R_K]$ is known to the scheduler and design a proportional fair scheme.

A. The Algorithm

Consider first a scheme which maximizes the utility defined as the sum logarithm throughput:

$$u(\boldsymbol{T}(n)) = \sum_{k=1}^{K} \ln T_k(n)\ . \tag{2}$$

Note that by (1),

$$T_k(n) = \frac{n-1}{n}T_k(n-1) + \frac{1}{n}B_k(n). \tag{3}$$

So that the increase of the utility function after the n-th slot is

$$\begin{aligned}
& u(\boldsymbol{T}(n)) - u(\boldsymbol{T}(n-1)) \\
&= \sum_{k=1}^{K} (\ln T_k(n) - \ln T_k(n-1)) \\
&= \sum_{k=1}^{K} \ln\left(\frac{n-1}{n} + \frac{1}{n}\frac{B_k(n)}{T_k(n-1)}\right) \\
&= \sum_{k=1}^{K} \ln\left(\frac{n-1}{n} + \frac{1-\beta J(n)}{n} s_k(n) I_k(n)\right),
\end{aligned} \tag{4}$$

where the throughput-normalized rate is

$$s_k(n) = \frac{R_k(n)}{T_k(n-1)}. \tag{5}$$

Since the indicator $I_k(n)$ is zero for all but one user k in each slot, one can see that to greedily maximize the utility increment at time slot n, we should schedule the user with the maximum $s_k(n)$, which is the classical PF scheduling algorithm.

However, due to the assumption that the instantaneous rates $R_k(n)$ are unknown a priori, we can only probe the users rates and obtain $s_k(n)$ one by one in each slot. We formulate the following optimal stopping time problem [18]. Note that the scheduling decision made in one slot has no impact on future realization of the rates, it suffices to consider one arbitrary slot and omit the time index n. For the scheduler, the joint probing and scheduling problem at the beginning of the time slot is defined by two objects:

(i) The independent throughput-normalized rates $s_1, \ldots, s_K$.

(ii) A sequence of positive-valued reward functions $y_1, \ldots, y_K$, where if j channels have been probed to reveal their throughput-normalized instantaneous rates $t_1, \ldots, t_j$, the reward of terminating the probing phase and schedule the best user found so far is

$$y_j(t_1, \ldots, t_j) = (1 - j\beta) \max(t_1, \ldots, t_j). \tag{6}$$

The theory of optimal stopping is concerned with determining the stopping time J to maximize the expected reward $\mathbb{E}[y_J]$. The maximum number of probings in every slot is $J_{max} =$

$\min(K, \lfloor 1/\beta \rfloor)$. Compared with the classical optimal stopping problem, the formulation above is more general in the sense that the probing order of s_k is not deterministic. Hence the joint probing and scheduling scheme basically includes two tasks in each slot: to determine the order in which users are probed, and to select one user as the destination at a proper (stopping) time. Recalling the objective of maximizing the expected y_j, the user with the largest $\mathbb{E}[s_k(n)]$ should be probed first, and then the second largest and so on. From Assumption (A1), we know $\bar{s}_k(n) \triangleq \mathbb{E}[s_k(n)] = r_k/T_k(n-1)$. Hence the probing order is $\pi(n) = (k_1, \cdots, k_K)$ such that $\bar{s}_{k_1}(n) \geq \cdots \geq \bar{s}_{k_K}(n)$. Now that the probing order has been determined, the decision on when to stop can be addressed by investigating the structural property of the problem.

Theorem 1: Under the homogeneous rate assumption (A1), the joint probing and scheduling problem is a monotone stopping problem [18, Chapter 5], which means that, if $\mathcal{E}_j$ denotes the event

$$\left\{ y_j(s_{k_1}, \cdots, s_{k_j}) \geq \mathbb{E}[y_{j+1}(s_{k_1}, \cdots, s_{k_{j+1}}) | s_{k_1}, \cdots, s_{k_j}] \right\}, \tag{7}$$

then $\mathcal{E}_j \subseteq \mathcal{E}_{j+1}$ for $0 \leq j \leq J_{max} - 1$.

Proof: See appendix A.

Now the problem has been proved to be monotone, then from the [18, Theorem 1, Chapter 5], the one-state look-ahead rule is optimal. The one-stage look-ahead rule is the one that stops if the reward for stopping at current stage is at least as large as the expected reward of continuing one stage and then stop. Mathematically, the rule is described by the stopping time. Let w_j denote the largest value of the observed throughput-normalized rate after probing j users and $a \vee b \triangleq \max(a, b)$, the optimal stopping time is

$$J^* = \min \left\{ j \geq 0 : (1 - j\beta) w_j \geq (1 - (j+1)\beta) \mathbb{E} \left[w_j \vee \frac{R_{k_{j+1}}}{T_{k_{j+1}}(n-1)} \middle| w_j \right] \right\}, \tag{8}$$

which solves the stopping problem almost surely in each slot. Precisely, the optimal PF joint probing and scheduling (JPS-PF) scheme is described as Algorithm 1.

B. On the Optimality of Algorithm 1

To present the optimality of Algorithm 1, we need to show the convergence property.

Algorithm 1: JPS-PF

Initialization: $T_k(0) \leftarrow 1$ for $k = 1, \cdots, K$;
for $n = 1, 2, \cdots$ **do**
  $\bar{s}_k(n) \leftarrow r_k / T_k(n-1)$. Sort the throughput-normalized mean rate $\bar{s}_k(n)(k = 1, \cdots, K)$ in the descending order: $\bar{s}_{k_1}(n) \geq \cdots \geq \bar{s}_{k_K}(n)$;
  $j \leftarrow 0$, $w \leftarrow 0$;
  do
    $j \leftarrow j + 1$;
    Probe user k_j and get the rate $R_{k_j}(n)$;
    $w \leftarrow w \vee R_{k_j}(n) / T_{k_j}(n-1)$;
  while $(1 - j\beta)w < (1 - (j+1)\beta)\mathbb{E}\left[w \vee \frac{R_{k_{j+1}}}{T_{k_{j+1}}(n-1)}\right]$;
  Transmit to user k_j. Update $\boldsymbol{T}(n)$;
end

Theorem 2: Assume (A1). Then for any initial condition, the throughput sequence $\boldsymbol{T}(n)$ generated under Algorithm 1 converges almost surely to the limit point $\boldsymbol{T}^*$ of the ordinary differential equation $\dot{\boldsymbol{T}}(t) = \boldsymbol{h}(\boldsymbol{T}(t))$, where $\boldsymbol{h}(\boldsymbol{T}) = -\boldsymbol{T} + \mathbb{E}[\boldsymbol{B}(n) | \boldsymbol{T}(n-1) = \boldsymbol{T}]$. Moreover, all users' steady-state throughput are proportional to their mean rate with an identical ratio κ,

$$\frac{T_1^*}{r_1} = \frac{T_2^*}{r_2} = \cdots = \frac{T_K^*}{r_K} = \kappa. \tag{9}$$

Proof: Let $\boldsymbol{M}(n) = \boldsymbol{B}(n) - \mathbb{E}[\boldsymbol{B}(n) | \boldsymbol{T}(n-1)]$. By (3), the update of users' throughput can be organized in the form of stochastic approximation iteration [19, Eqn. 2.1.1]:

$$\boldsymbol{T}(n) = \boldsymbol{T}(n-1) + a(n)[\boldsymbol{h}(\boldsymbol{T}(n-1)) + \boldsymbol{M}(n)],$$

where $a(n) = 1/n$. The equation above is a standard stochastic approximation expression. It is easy to verify that $\boldsymbol{h}(\cdot)$ is Lipshitz, the stepsize satisfies $\sum_n a(n) = \infty, \sum_n a(n)^2 < \infty$ and $\boldsymbol{T}(n)$ is bounded. Furthermore, it is easy to verify that $\mathbb{E}[\boldsymbol{M}(n) | \boldsymbol{M}(1), \cdots, \boldsymbol{M}(n-1)] = 0$, so $\boldsymbol{M}(n)$ is a martingale difference sequence. Now the throughput update under the proposed scheme satisfies the assumptions (A1)-(A4) in [19, Section 2.1], then applying Theorem 2 in [19, Section 2.1] directly, the convergence conclusion holds.

Now the convergence of the throughput sequence has been obtained. The remainder of the proof is by contradiction. Suppose (9) does not hold at steady state and that $T_1^*/r_1 < T_2^*/r_2$ without loss of generality. Consider the throughput path starting at slot n_0 which is at steady state. At this time, $\bar{s}_l = r_l/T_l^*(l = 1, 2)$ and $\bar{s}_1 > \bar{s}_2$. Thus user 1 is probed first in each slot. From assumption (A1) we know that s_1 and s_2 are of the same type of distribution, but s_1 has a larger mean value. Thus user 1 is selected for transmission more often than user 2, which would further imply $T_1(n_0 + n_1)/r_1 > T_2(n_0 + n_1)/r_2$ after a sufficiently large number (n_1) of slots, which contradicts the steady state assumption with $T_1^*/r_1 < T_2^*/r_2$. ■

Note that the constant proportionality factor κ is a bridge connecting the steady-state throughput and the mean-rate. After obtaining κ, it is straightforward to evaluate the throughput and utility. On the other hand, due to the fact that κ is a constant, we have the following corollary from the proof of Theorem 2.

Corollary 1: Under Algorithm 1, the probability that each user is selected as the destination is identical as $1/K$.

Algorithm 1 is asymptotically optimal in the following sense:

Theorem 3: Assume (A1). Then $\boldsymbol{T}^*$ maximizes the PF utility $u(\cdot)$ over the rate region generated by all joint probing and scheduling schemes.

Proof: Let $\mathcal{S}$ denote the set composed of all the feasible schemes Γ under the assumption that only one user can be selected in one slot. The developed scheme in this paper is denoted as Γ^*. We have shown in the derivation of Algorithm 1 that Γ^* is optimal for solving the monotone stopping problem in each slot, that is, it maximizes $B_k(n)/T_k(n-1)$ in slot n almost surely. Due to the constraint that only one user can be scheduled in one slot, we can see that the developed scheme Γ^* satisfies

$$\Gamma^* \in \arg\max_{\Gamma \in \mathcal{S}} \sum_{k=1}^{K} \frac{B_k^{(\Gamma)}(n)}{T_k(n-1)}, \tag{10}$$

where $B_k^{(\Gamma)}(n)$ is the number of bits transmitted to user k in slot n under the scheme Γ. Recalling the definition of the utility function in (2), it can be found that

$$\sum_{k=1}^{K} \frac{B_k^{(\Gamma)}(n)}{T_k(n-1)} = \nabla u(\boldsymbol{T}(n-1)) \cdot \boldsymbol{B}^{(\Gamma)}(n), \tag{11}$$

which means that the scheme chooses a decision maximizing the scalar product of $\boldsymbol{B}^{(\Gamma)}(n)$ and the gradient $\nabla u(\boldsymbol{T}(n-1))$.

The *gradient scheduling* algorithm developed by Stolyar [17] is that, at time n the controller chooses a decision $\Gamma(n) \in \arg\max_{\Gamma} \nabla u(\boldsymbol{T}(n-1)) \cdot \boldsymbol{B}^{(\Gamma)}(n)$. Let $\tilde{\boldsymbol{T}}$ denote the solution to the problem

$$\begin{aligned} \max \quad & u(\boldsymbol{T}) \\ s.t. \quad & \boldsymbol{T} \in \mathcal{V}, \end{aligned}$$

where $\mathcal{V}$ is the system rate region, i.e., the set of all feasible long-term service rate vectors. Then the [17, Theorem 2] shows that the expected average service rates under the gradient scheduling algorithm converges in probability to $\tilde{\boldsymbol{T}}$.

By (10) and (11), one can see that the joint probing and scheduling algorithm in this paper belongs to the gradient scheduling algorithm. From the convergence of Algorithm 1, we know $\boldsymbol{T}^* = \tilde{\boldsymbol{T}}$. Then the achieved throughput $\boldsymbol{T}^*$ maximizes the PF utility function asymptotically. ■

C. A Static Threshold Criteria

Note that in Algorithm 1, after each probe, the scheduler needs to evaluate the expectation in (8) which depends on the channel realizations. Further reduction in the computational complexity is possible by simply comparing the highest normalized rate against a sequence of deterministic thresholds, in lieu of computing (8). Consider the steady-state case where users' throughput is exactly $\boldsymbol{T}^*$. Note that by Theorem 2,

$$\frac{R_{k_{j+1}}}{T_{k_{j+1}}(n-1)} = \frac{R_{k_{j+1}}}{T^*_{k_{j+1}}},$$

which is identically distributed as X_1/κ. For $0 \leq j \leq J_{max} - 1$, the inequality of w_j in (8) reduces to

$$(1 - j\beta)w_j \geq (1 - (j+1)\beta)\mathbb{E}[\max(w_j, \kappa^{-1}X_1)|w_j]. \tag{12}$$

It turns out that (12) can be reduced to comparing κw_j with a static threshold v_j, which can be determined as follows. Let $F_X(\cdot)$ denote the cumulative distribution function (CDF) of X_k. Then

$$\mathbb{E}\left[\max\left(w_j, \frac{X_1}{\kappa}\right)\middle| w_j\right] = w_j + \int_{\kappa w_j}^{\infty} \left(\frac{x}{\kappa} - w_j\right) dF_X(x). \tag{13}$$

So that (12) can be rewritten as

$$(1 - j\beta)w_j \geq (1 - (j+1)\beta)\left[w_j + \int_{\kappa w_j}^{\infty} \left(\frac{x}{\kappa} - w_j\right) dF_X(x)\right], \tag{14}$$

or, equivalently,

$$\kappa w_j \geq g_j(\kappa w_j), \tag{15}$$

where

$$g_j(v) = \left[\beta^{-1} - (j+1)\right] \int_{v}^{\infty} (x - v) dF_X(x). \tag{16}$$

It is not hard to check that: (i) $g_j(v) > 0$ for $v \geq 0$; (ii) $g_j(v)$ is a strictly decreasing function of v; (iii) $\lim_{v\to\infty} g_j(v) = 0$. Then inequality (15) is equivalent to $\kappa w_j \geq v_j$, where v_j is the cross point of function $f(v) = v$ and $g_j(v)$. Also, we have $g_j(v) > g_{j+1}(v)$. Then it is easy to verify that $v_{j+1} < v_j$. The solution to (15) is illustrated in Fig. 2.

By observing the structure of (16), it is worth pointing out that the cross point v_j is only determined by j, β and the CDF $F_X(\cdot)$, i.e., the unit mean valued random variable X_j. And the value of v_j is independent of the number of users K, the mean rates of all users r_k as well as the achieved throughput to mean-rate ratio κ. Hence if the transmitter knows the distribution $F_X(\cdot)$, it can compute v_j in advance.

Now inequality (12) can be expressed as $w_j \geq \frac{1}{\kappa}v_j$ for $0 \leq j \leq J_{max} - 1$, which is also equivalent to the inequality in (8) in the steady-state case. Thus the decision on whether to keep probing or to start transmitting is decided by a static threshold criteria. For completeness, let

$v_{J_{max}} = 0$ in order to make sure the probing can always be terminated in each slot. We get the following static threshold based probing criteria, which can replace the line 9 in Algorithm 1.

Criteria 1: After probing j users, if the current value of the largest normalized rate $w_j \geq \frac{1}{\kappa} v_j$, then the transmitter transmits to the user with the largest normalized rate; otherwise it probes the $(j+1)$st user.

In practice, the scheduler can calculate v_j in advance but κ is unavailable at the beginning. One way to estimate κ is to start the joint probing and scheduling using the dynamic criteria in line 9 of Algorithm 1. After a period of time, the throughput approaches to its steady-state value. Then the throughput to mean-rate ratio κ is obtained and the static threshold criteria can be used thereafter. Alternatively, κ can be determined theoretically as discussed in the next subsection.

D. The Scheduling Gain

In this section we analyze the performance of the proposed scheme theoretically. We define the *scheduling gain* as the ratio of the achieved throughput to that using round robin scheduling without probing, which reflects how much multiuser diversity benefits can be exploited. The scheduling gain of the proposed joint probing and scheduling scheme is $\frac{T_k^*}{K^{-1} r_k} = \kappa K$. For a random variable X, let us denote the truncation of X over $[a, b]$ as $[X]_a^b$. Note that $\mathbb{E}[X | a \leq X \leq b] = \mathbb{E}[X]_a^b$.

Theorem 4: Under the homogeneous rate assumption (A1), the scheduling gain of Algorithm 1 is

$$\kappa K = \sum_{j=1}^{J_{max}} \left[(F_X(v_{j-1}))^{j-1} - (F_X(v_j))^j \right] (1 - j\beta) \mathbb{E} \left\{ \left[\max \left([X_1]_0^{v_{j-1}}, \cdots, [X_{j-1}]_0^{v_{j-1}}, X_j \right) \right]_{v_j}^{\infty} \right\},$$

where v_j is the solution of $v = g_j(v)$.

Recall that J^* is the optimal stopping time, that is, the number of users probed before a user is scheduled. We prove Theorem 4 using the following supporting lemma.

Lemma 1: Using Algorithm 1, the steady-state probability of the event that j users are probed until transmission is given by

$$p_j = (F_X(v_{j-1}))^{j-1} - (F_X(v_j))^j, \; 1 \leq j \leqslant J_{max}. \tag{17}$$

Proof: At steady state, all users' throughput-normalized mean rates r_k/T_k^* are essentially identical. Let $q_j = \Pr\{J^* \geq j\}$, i.e., the probability that at least j users are probed before transmission. Then $q_1 = 1$. And from Criteria 1, we have for $j \geq 2$,

$$
\begin{aligned}
q_j &= \Pr\{\max(X_1, \cdots, X_{j-1}) < v_{j-1}\} \\
&= \Pr\{X_1 < v_{j-1}\} \cdots \Pr\{X_{j-1} < v_{j-1}\} \\
&= (F_X(v_{j-1}))^{j-1}.
\end{aligned}
$$

Like v_j, q_j is also completely determined by the rate distribution. Clearly, $p_j = q_j - q_{j+1}$ for $j \leq J_{max} - 1$ and $p_{J_{max}} = q_{J_{max}}$. ■

Proof of Theorem 4: Consider a specific user k. In the steady state, $\dot{\boldsymbol{T}}(t) = 0$. Then from Theorem 2, user k's throughput is given by $T_k^* = \mathbb{E}[B_k(n)|\boldsymbol{T}^*]$. Throughout, let K^* denote index of the user that is selected as destination. Then event $\{K^* = k\}$, i.e., user k is selected as destination, can be decomposed into J_{max} exclusive sub events: $\{K^* = k\} = \bigcup_{j=1,\cdots,J_{max}} \{K^* =$

$k, J^* = j\}$. Then we have

$$
\begin{aligned}
T_k^* =& \mathbb{E}[B_k(n)|\boldsymbol{T}^*] = \mathbb{E}[(1 - J^*\beta)R_k I_k] \\
=& \Pr\{K^* = k\}\mathbb{E}[(1 - J^*\beta)R_k|K^* = k] \\
\overset{(a)}{=}& \frac{1}{K}\mathbb{E}[(1 - J^*\beta)R_k|K^* = k] \\
\overset{(b)}{=}& \frac{1}{K}\sum_{j=1}^{J_{max}} \Pr\{J^* = j\}\mathbb{E}[(1 - j\beta)R_k|K^* = k, J^* = j] \\
=& \frac{T_k^*}{K}\sum_{j=1}^{J_{max}} p_j(1 - j\beta)\mathbb{E}\left[\frac{R_k}{T_k^*}\middle| K^* = k, J^* = j\right] \\
\overset{(c)}{=}& \frac{T_k^*}{K}\sum_{j=1}^{J_{max}} p_j(1 - j\beta)\mathbb{E}\left\{\left[\max\left(\left[\frac{R_1}{T_1^*}\right]_0^{\frac{v_{j-1}}{\kappa}}, \cdots, \left[\frac{R_{j-1}}{T_{j-1}^*}\right]_0^{\frac{v_{j-1}}{\kappa}}, \frac{R_j}{T_j^*}\right)\right]_{\frac{v_j}{\kappa}}^{\infty}\right\} \\
\overset{(d)}{=}& \frac{T_k^*}{K}\sum_{j=1}^{J_{max}} p_j(1 - j\beta)\mathbb{E}\left\{\left[\max\left(\left[\frac{X_1}{\kappa}\right]_0^{\frac{v_{j-1}}{\kappa}}, \cdots, \left[\frac{X_{j-1}}{\kappa}\right]_0^{\frac{v_{j-1}}{\kappa}}, \frac{X_j}{\kappa}\right)\right]_{\frac{v_j}{\kappa}}^{\infty}\right\} \\
\overset{(e)}{=}& \frac{T_k^*}{\kappa K}\sum_{j=1}^{J_{max}} p_j(1 - j\beta)\mathbb{E}\left\{\left[\max\left([X_1]_0^{v_{j-1}}, \cdots, [X_{j-1}]_0^{v_{j-1}}, X_j\right)\right]_{v_j}^{\infty}\right\},
\end{aligned}
$$

where (a) follows from Corollary 1, (b) from the law of total probability, (c) from the static threshold criteria, that is, $\{K^* = k, J^* = j\}$ means that: i) user k has the largest throughput-normalized rate among the first j users; ii) the first $j - 1$ users' throughput-normalized rates are smaller than $\kappa^{-1}v_{j-1}$ and iii) the largest value of the first j users' throughput-normalized rates is larger than $\kappa^{-1}v_j$, (d) from $R_k = r_k X_k$ and (9), and (e) from the distribution of X_j. By replacing p_j with (17) and removing T_k^* from both sides, the conclusion of Theorem 4 holds. ■

IV. Joint Learning, Probing and Scheduling

Consider the case where the scheduler does not know *a priori* the statistics of the quality of the downlink channels, and thus has to rely on the history of the probed CQI to decide on the user probing order and user selection. Under this assumption, the problem of maximizing the PF utility function is a generalization of the classical multiarmed bandit problem [20]. The problem is a generalization because in the classical bandit problem, the decision maker has to decide which of K random process to observe in a sequential of trials so as to maximize the reward, where the 'observing' operation is equivalent to the 'utilizing' operation. However, in our model, in each slot, the scheduler may probe (observe) more than one channels (random processes) and then choose only one for transmission (utilization). The observation does not always lead to a utilization.

At the beginning of slot n, i.e., the end of slot $n-1$, let $M_k(n-1)$ denote the number of time slots in which the channel to user k has been probed, and $\mathcal{R}_k(n-1) = \{R_k^{(1)}, \cdots, R_k^{(M_k(n-1))}\}$ record all the probed samples of the channel rate of user k. Clearly, the cardinality $|\mathcal{R}_k(n-1)| = M_k(n-1)$. The scheduler keeps updating the K sets $[\mathcal{R}_1(n), \cdots, \mathcal{R}_K(n)]$ from slot to slot. Also, the scheduler knows the throughput $\boldsymbol{T}(n-1)$ till the previous slot. The objective is still to find a scheme that solves the stopping problem in each slot. As analyzed in Section III-A, there still exists the same two tasks to find the optimal scheme: determining the user probing order and selecting one user for transmission. Hence the problem formulation and scheme design is similar to those in Section III-A. The only difference is that the scheduler just has the sampled values of all channels' rates instead of the explicit knowledge of the distribution of $R_k, (k = 1, \cdots, K)$, which means that we cannot calculate the expectations related to R_k directly. Alternatively, we can only evaluate the empirical average using the acquired samples of R_k, which readily leads to the index-based policy solution in the framework of bandit problem.

The index policy, consisting of choosing at any time the stochastic process with the currently highest index, is the solution to a class of bandit problems. Here to find the optimal scheme, we adopt the similar methodology as in the development of the index-based policy by Agrawal in [21]. For the decision on the user probing order, we use the current average reward, i.e., the throughput-normalized average rate as the index. For the decision on when to start transmission, we adopt the actually served bits in current slot, i.e., the product of $1 - j\beta$ and the conditional throughput-normalized-average rate. For the convenience of presenting the algorithm, we define the following two empirical averages

$$\tilde{s}_k(n) \triangleq \frac{1}{M_k(n-1)} \sum_{m=1}^{M_k(n-1)} \frac{R_k^{(m)}}{T_k(n-1)}, \tag{18}$$

$$\tilde{e}_k(n,w) \triangleq \frac{1}{M_k(n-1)} \sum_{m=1}^{M_k(n-1)} \left[w \vee \frac{R_k^{(m)}}{T_k(n-1)} \right]. \tag{19}$$

The $\tilde{s}_k(n)$ is used to replace the $\bar{s}_k(n)$ in Algorithm 1 and the $\tilde{e}_k(n,w)$ is for $\mathbb{E}\left[w \vee \frac{R_k}{T_k(n-1)}\right]$ in Algorithm 1. Then a joint PF learning, probing and scheduling (JLPS-PF) algorithm is described in Algorithm 2.

Algorithm 2: JLPS-PF

Initialization: $n \leftarrow \lceil \beta K \rceil$. For $k = 1, \cdots, K$, $T_k(n) \leftarrow 1$. In the first n slots, sequentially probe each channel once, making sure that each one of the sets $\mathcal{R}_k(n), (k = 1, \cdots, K)$ is not empty. $M_k(n) \leftarrow 1$;

for $n = \lceil \beta K \rceil + 1, \lceil \beta K \rceil + 2, \cdots$ **do**

  $\tilde{s}_k(n) \leftarrow \frac{1}{M_k(n-1)} \sum_{m=1}^{M_k(n-1)} R_k^{(m)}/T_k(n-1)$. Sort $\tilde{s}_k(n)(k = 1, \cdots, K)$ in the descending order: $\tilde{s}_{k_1}(n) \geq \cdots \geq \tilde{s}_{k_K}(n)$;

  $j \leftarrow 0$, $w \leftarrow 0$;

  do

    $j \leftarrow j + 1$;

    Probe user k_j and get the rate $R_{k_j}(n)$;

    $w \leftarrow w \vee R_{k_j}(n)/T_{k_j}(n-1)$;

    $\tilde{e}_{k_{j+1}}(n,w) \leftarrow \frac{1}{M_{k_{j+1}}(n)} \sum_{m=1}^{M_{k_{j+1}}(n)} \left[w \vee \frac{R_{k_{j+1}}^{(m)}}{T_{k_{j+1}}(n-1)} \right]$;

    $\mathcal{R}_{k_j}(n) \leftarrow \mathcal{R}_{k_j}(n-1) \cup \{R_{k_j}(n)\}$, $M_{k_j}(n) \leftarrow M_{k_j}(n-1) + 1$;

  while $(1 - j\beta)w < (1 - (j+1)\beta)\tilde{e}_{k_{j+1}}(n,w)$;

  Transmit to user k_j. Update $\boldsymbol{T}(n)$;

  For $k = k_j + 1, \cdots, k_K$, $\mathcal{R}_k(n) \leftarrow \mathcal{R}_k(n-1)$, $M_k(n) \leftarrow M_k(n-1)$;

end

From the description of Algorithm 2, one may wonder such a phenomenon may exist that if one user is probed with relatively high values in the first few slots, then it will have low priority of being probed afterwards, resulting that the ensemble average of this channel is always higher than its statistical expectation. However, this does not happen thanks to the structure of the algorithm derived from the objective of maximizing the PF utility. As a matter of fact, if user k is probed and selected less frequently compared to other users, the achieved throughput $T_k(n)$ will become small, which will in return increase its priority of being probed and selected. In fact, the metric of throughput-normalized rate used in PF scheduling is a well-balanced rule that guarantees each user is sampled with sufficiently many times and identical frequencies. Hence after the Algorithm 2 runs a a sufficiently long time, the sampled data of each user's channel rate can characterize the statistics of $\boldsymbol{R}$ well. Then from the law of large number, the ensemble average converges to the statistical expectation. And the performance of Algorithm 2 is almost the same as that of Algorithm 1.

V. Numerical Results

In this section, we provide some numerical experiments illustrating the theoretical findings of the previous sections. Our objectives here are (i) to evaluate the performance of the developed schemes with and without channel statistics; (ii) to compare the developed scheme for achieving PF with some ideal and practical schemes and to quantify the impact of the cost of CQI on the scheduling. We consider the scenario where users' rates obey the exponential distributions with average equal to the user index. The exponential rate assumption is an appropriate approximation of the Shannon capacity under Rayleigh fading channels in low SNR regime.

A. Evaluation of the Proposed Algorithms

Consider $K = 20$ users and let the fraction of one probe be $\beta = 0.1$. Up to $J_{max} = 10$ users can be probed in each slot.

Fig. 3 presents a sample throughput trajectory of user 1 when scheduled with Algorithm 1, the static threshold criteria given in *criteria 1* and Algorithm 2. The simulation runs for $10,000$ slots in this experiment. The time axis is in logarithmic scale to highlight the transient behavior. We can see that the static threshold criteria works well. The variation of the throughput diminishes over time as more and more time slots are included in the averaging. It is worth noting that the low complexity of the static threshold criteria for solving the optimal stopping problem comes from the explicit knowledge of the channel statistics. If this information is not known, or if the distribution of the channel rate varies over time, we can only adopt the dynamic criteria given in Algorithm 1.

Fig. 4 illustrates the frequency of each user being scheduled in a relatively short period of 2000 slots. Each of the 20 user is selected as the destination for roughly 100 slots. That is, the scheme is fair to all users even within a small application time window.

Fig. 5 presents the probability that k users have been probed until transmission. The theoretical results are from Lemma 1. The figure shows that both the Algorithm 1 and Algorithm 2 coincide with the theoretical results. We observe from the figure that the probability decreases sharply as the probing step approaches J_{max}.

Fig. 6 plots the scheduling gain of the proposed algorithms versus the number of users in the system. The simulation runs for 20,000 slots. In fact the simulation result matches the analytical result of Theorem 4 quite well. Also, we note the scheduling gain remains about the same for more than 9 users. Because at this time, the cost of user probing is dominant and the scheme always tries to carry out the user probing till the end.

B. Comparison between the Proposed Scheme and Other Schemes

The fraction of slot for probing one user is still set $\beta = 0.1$. Here four schemes are considered: (a) the proposed joint probing and scheduling scheme; (b) Round robin scheduling; (c) Genie-aided PF (GA-PF) scheme where full CQI is available to the scheduler at the beginning of each slot; (d) Probe-all PF (PA-PF) scheme where the transmitter probes all users before scheduling. For both (c) and (d), the transmitter selects the user with the largest $R_k(n)/T_k(n-1)$ for transmission. From [22] we know that the scheduling gain of GA-PF is $\mathbb{E}\left[\max_{k=1,\cdots,K} X_k\right]$. Then that of PA-PF is $\max(1-K\beta,0)\mathbb{E}\left[\max_{k=1,\cdots,K} X_k\right]$.

Fig. 7 presents the scheduling gain of schemes (a)-(d) as a function of the number of users. We can see from Fig. 7 that when probing cost is taken into account, the scheduling gain does not always increase but approaches to a limit value as the number of users increases. This indicates that, by ignoring the cost of channel probing, the ideal genie-aided PF does not reflect the correct multiuser diversity characteristics. The comparison also shows the advantage of the proposed joint probing and scheduling scheme. For the probe-all PF scheme, it achieves higher gain than round robin when the user population is not very large compared with β^{-1}. However, when the number of user increases to some extent, the scheduling gain of probe-all algorithm vanishes. That is because almost all the period of one slot is used for user-probing instead of data transmission.

Fig. 8 displays the sum throughput of all schemes as the number of users increases. One can see that there exists a relative large gap between the ideal genie-aided PF curve and the proposed scheme. The gap quantifies the the extent to which the user probing decreases the system performance. For example, when the number of users is $K = 20$, the throughput of the joint probing and scheduling scheme only accounts for 55.64% of that of the genie-aided PF. And the throughput achieved by the joint scheme is the highest among all the non-ideal schemes (a), (b) and (d). The probe-all PF scheme performs similar to the joint probing and scheduling scheme when there are not many users ($K \leq 6$), but degrades fast and even vanishes when the number of users becomes large.

VI. Conclusion

We have studied the problem of achieving proportional fairness in wireless systems when explicitly taking into account the channel probing cost. An optimal adaptive joint probing and scheduling scheme is presented, as well as a static threshold based criteria for determining whether to probe or to transmit. Using the steady-state analysis, we have evaluated the scheduling gain explicitly. Extension of the scheme to the case in which the scheduler has no knowledge of the channel rate distribution has been developed, which achieves almost the same performance of the algorithm obtained under known rate statistics assumption and outperforms other non-ideal PF schemes. In this work, we have focused on the well-studied proportional fairness rule. It is possible to extend the results to more general utilities, for example, the α fair utility [7]. The methodology presented in this paper can then be carried through to that case as well.

APPENDIX A
PROOF OF THEOREM 1

Proof: Let the largest throughput-normalized user rate after probing j users be denoted by

$$w_j = \max_{1\leq l\leq j} s_{k^{(l)}} \tag{20}$$

Then the current reward can be written as $y_j(s_{k_1}, \cdots, s_{k_j}) = (1-j\beta)w_j$ and the expected reward obtained from probing the next user is

$$\mathbb{E}[y_{j+1}(s_{k_1}, \cdots, s_{k_{j+1}})|s_{k_1}, \cdots, s_{k_j}] = (1-(j+1)\beta)\mathbb{E}[w_j \vee s_{k_{j+1}}|w_j]. \tag{21}$$

Then the event $\mathcal{E}_j$ can be expressed as

$$\mathcal{E}_j = \{(1-j\beta)w_j \geq (1-(j+1)\beta)\mathbb{E}[w_j \vee s_{k_{j+1}}|w_j]\}. \tag{22}$$

We first show that there exists a threshold $w_j^{(th)}$ such that the event $\mathcal{E}_j$ can be represented as $\mathcal{E}_j = \{w_j \geq w_j^{(th)}\}$. To this end, let $f_j(w) = (1-j\beta)w - (1-(j+1)\beta)\mathbb{E}[w \vee s_{k_{j+1}}]$. Then $w \in \mathcal{E}_j \Leftrightarrow f_j(w) \geq 0$. It is easy to verify that $f_j(0) < 0$ and $f_j(\infty) > 0$. The function $f_j(w)$ can be reorganized as $f_j(w) = \beta\mathbb{E}[w \vee s_{k_{j+1}}] + (1-j\beta)\mathbb{E}[w - w \vee s_{k_{j+1}}]$. For any $w' > w > 0$,

$$f_j(w') - f_j(w) = \beta\mathbb{E}[w' \vee s_{k_{j+1}} - w \vee s_{k_{j+1}}] + (1-j\beta)\mathbb{E}[w' - w + w' \vee s_{k_{j+1}} - w \vee s_{k_{j+1}}].$$

Note that $w' \vee s_{k_{j+1}} \geq w \vee s_{k_{j+1}}$ and $w' - w \geq w' \vee s_{k_{j+1}} - w \vee s_{k_{j+1}}$. Thus $f_j(w') - f_j(w) \geq 0$, that is, $f_j(w)$ is a nondecreasing function. Summarizing the properties of $f_j(w)$, it can be seen that the solution to $f_j(w) \geq 0$ can be expressed as $w \geq w_j^{(th)}$.

We next show that $w_{j+1}^{(th)} \leq w_j^{(th)}$. For fixed w,

$$
\begin{aligned}
& f_{j+1}(w) - f_j(w) \\
=& (1-(j+1)\beta)w - (1-(j+2)\beta)\mathbb{E}_{s_{k_{j+2}}}[w \vee s_{k_{j+2}}] - (1-j\beta)w + (1-(j+1)\beta)\mathbb{E}[w \vee s_{k_{j+1}}] \\
=& \beta\mathbb{E}_{s_{k_{j+2}}}[w \vee s_{k_{j+2}} - w] + (1-(j+1)\beta)\{\mathbb{E}[w \vee s_{k_{j+1}}] - \mathbb{E}_{s_{k_{j+2}}}[w \vee s_{k_{j+2}}]\} \\
\geq& 0. \qquad (23)
\end{aligned}
$$

where the last '$\geq$' follows from the fact that $s_{k_{j+1}}$ and $s_{k_{j+2}}$ are of the same type of distribution and $\mathbb{E}s_{k_{j+1}} \geq \mathbb{E}s_{k_{j+2}}$. Note that $w_j^{(th)}$ is the zero point of the function $f_j(w)$. Hence $w_{j+1}^{(th)} \leq w_j^{(th)}$, as illustrated in Fig. 1.

Collecting the preceding results, we have $\mathcal{E}_j = \{w_j \geq w_j^{(th)}\} \subseteq \{w_{j+1} \geq w_j^{(th)}\} \subseteq \{w_{j+1} \geq w_{j+1}^{(th)}\} = \mathcal{E}_{j+1}$. ■

REFERENCES

[1] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. Netw.*, vol. 8, no. 5, pp. 556-567, Oct. 2000.

[2] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inf. Theory.*, vol. 48, no. 6, pp. 1277-1294, June 2002.

[3] F. P. Kelly, "Charging and rate control for elastic traffic", *Euro. Trans. Telecommun.*, vol 8, pp. 7-20, 1997.

[4] H. J. Kushner and P. A. Whiting, "Convergence of proportional-fair sharing algorithms under general conditions", *IEEE Trans. Wireless Commun.*, vol. 3, no. 4. pp. 1250-1259, July 2004.

[5] S. Borst and M. Jonckheere, "Flow-level stability of channel-aware scheduling algorithms", in *Proc. WiOpt 06*, 2006.

[6] J. G. Choi and S. Bahk, "Cell-throughput analysis of the proportional fair scheduler in the single-cell environment", *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 766 - 778, 2007.

[7] J. Liu, A. Proutiere, Y. Yi, M. Chiang and H. V. Poor, "Stability, fairness, and performance: A flow-level study on nonconvex and time-varying rate regions", *IEEE Trans. Inf. Theory.*, vol. 55, no. 8, pp. 3437 - 3456, 2009.

[8] C. W. Chan and N. Bambos, "Throughput loss in task scheduling due to server state uncertainty," *VALUETOOLS* 2009.

[9] S. Guha, K. Mungala, and S. Sarkar, "Jointly optimal transmission and probing strategies for multichannel wireless systems", In CISS 06, Mar 2006.

[10] N. Chang and M. Liu. "Optimal channel probing and transmission scheduling for opportunistic spectrum access," In *ACM MOBICOM*, 2007.

[11] P. Chaporkar and A. Proutiere, "Optimal joint probing and transmission strategy for maximizing throughput in wireless systems," *IEEE Journal on Selected Areas in Communications*, vol. 26, no.8, pp. 1546-1556, 2008.

[12] J. Chen, R. A. Berry, and M. L. Honig, "An adaptive limited feedback scheme for MIMO OFDM based on optimal stopping", *Proc. Allerton Conference*, 2008.

[13] A. Gopalan, C. Caramanis and S. Shakkotai, “On wireless scheduling with partial channel-state information,” in *Allerton Conference on Communication, Control, and Computing*, 2007.

[14] M. Ouyang and L. Ying, “On scheduling in multi-channel wireless downlink networks with limited feedback”, In *Proc. Allerton Conference*, 2009.

[15] P. Chaporkar, A. Proutiere, H. Asnani and A. Karandikar, “Scheduling with limited information in wireless systems”, in *ACM MobiHoc*, 2009.

[16] H. Zhou, P. Fan and D. Guo, “The Impact of Limited Information on Proportional Fair Scheduling In Wireless Networks”, accepted by IEEE GLOBECOM 2010.

[17] A. L. Stolyar, “On the asymptotic optimality of the gradient scheduling algorithm for multiuser throughput allocation”, *Opreations Research*, vol. 53, no. 1, January-February 2005, pp. 12-25.

[18] T. Ferguson, “Optimal stopping and applications,” [Online]. Available: http://www.math.ucla.edu/~tom/Stopping/Contents.html, 2006.

[19] V. S. Borkar, *Stochastic Approximation: A Dynamical Systems Viewpoint.* Cambridge University Press, 2008.

[20] D. A. Berry and B. Fristedt, *Bandit Problems: Sequential Allocation of Experiments*, London: Chapman and Hall, 1985.

[21] R. Agrawal, “Sample mean based index policies with O(log n) regret for the multi-armed bandit problem,” *Advances in Applied Probability*, Vol. 27, No. 4, pp. 1054-1078, 1995.

[22] S. Borst, “User level aware performance of channel-aware scheduling algorithms in wireless data networks,” in *Proc. Infocom*, 2002.

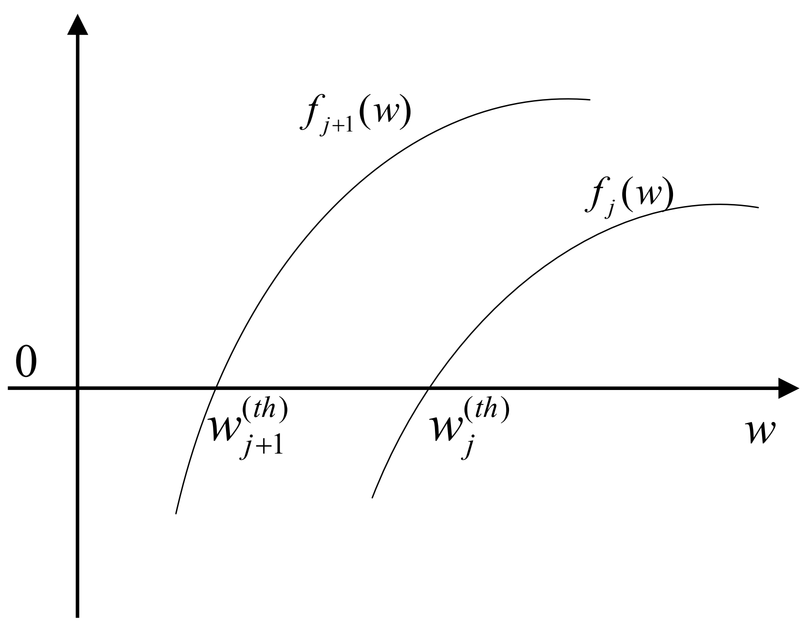

Fig. 1: Illustration of the property of function $f_j(w)$.

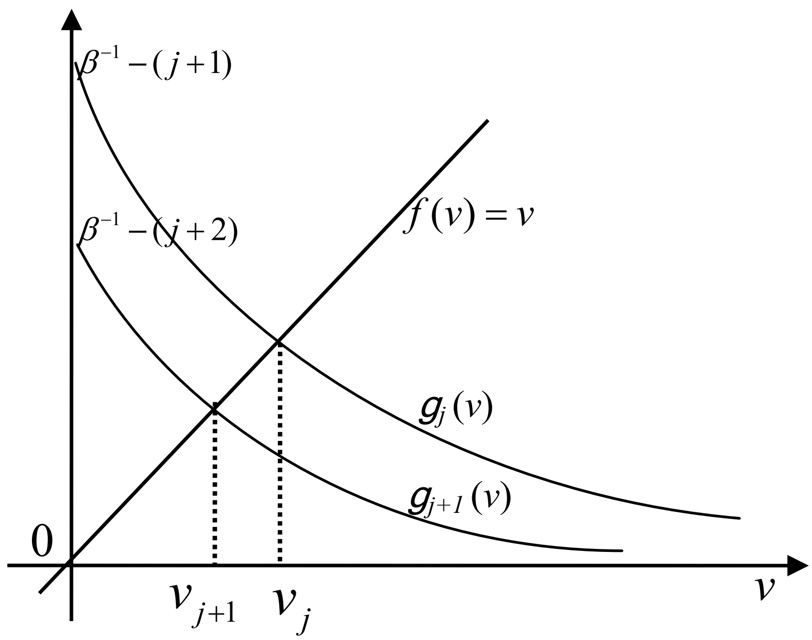

Fig. 2: Illustration of the solution to inequality (15).

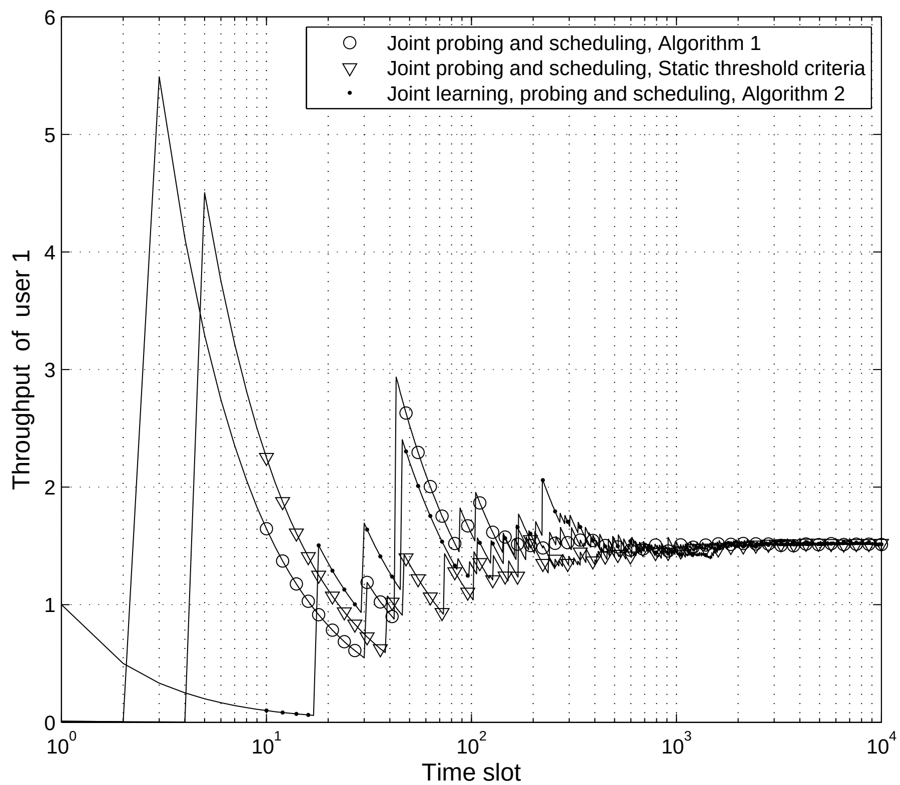

Fig. 3: The throughput trajectory of user 1 when scheduled with Algorithm 1, the static threshold criteria and Algorithm 2 respectively. $N_{slot} = 10,000, K = 20, \beta = 0.1$.

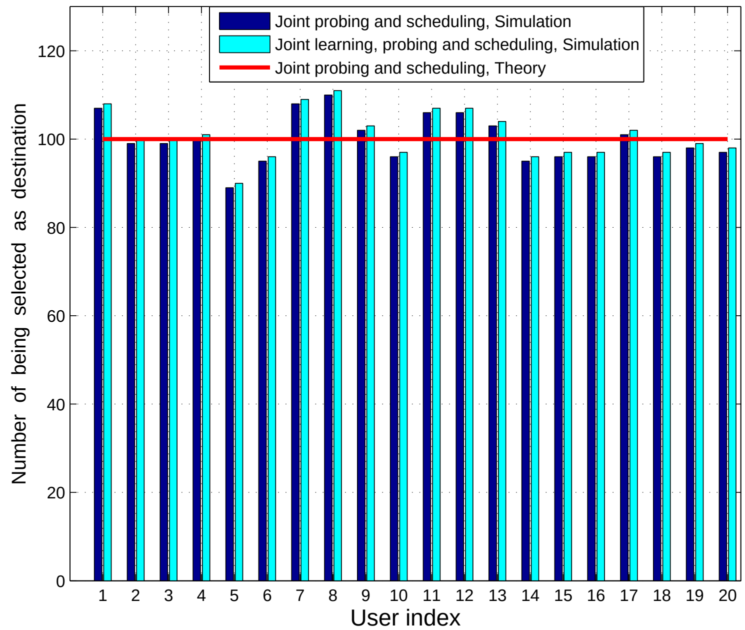

Fig. 4: The number of slots in which each user is selected as the destination. $N_{slot} = 2000, K = 20, \beta = 0.1$.

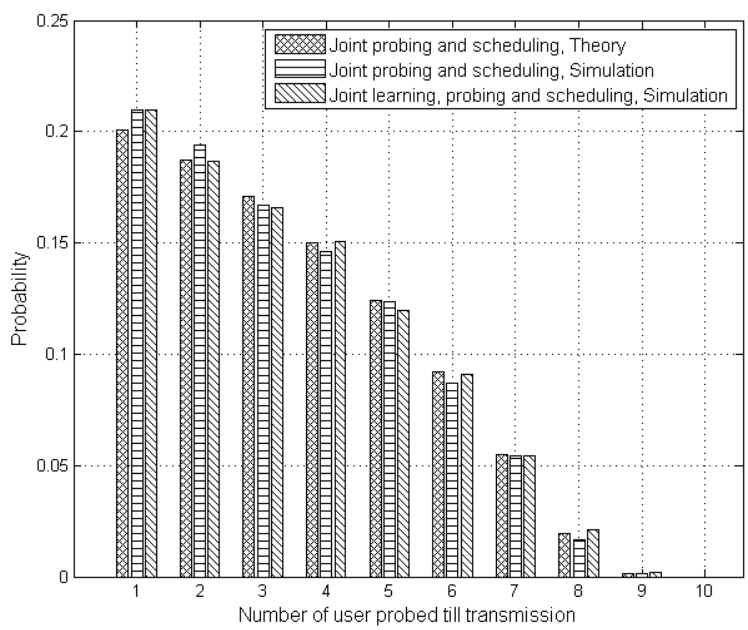

Fig. 5: The probability that k users have been probed until transmission. $K = 20, \beta = 0.1$.

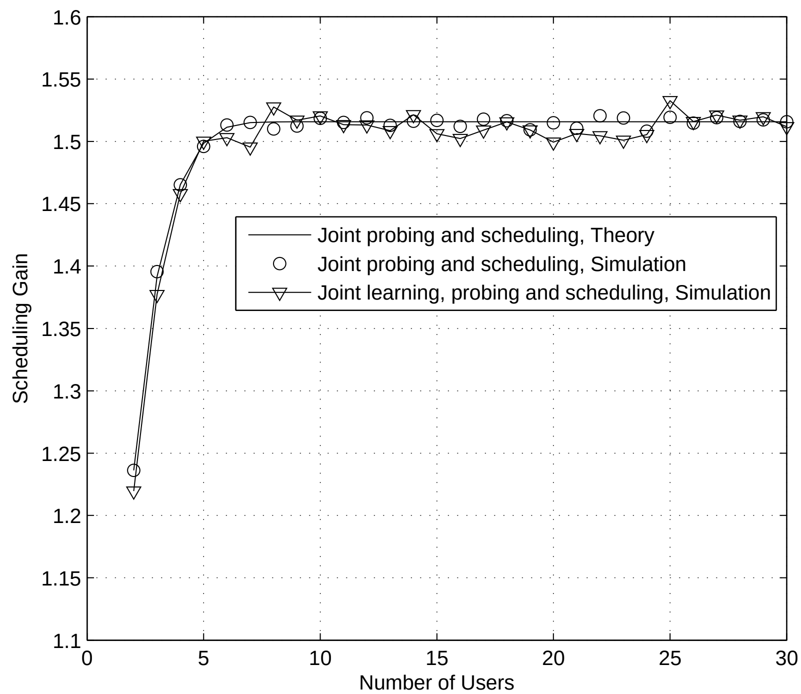

Fig. 6: The scheduling gain comparison between Algorithm 1, Algorithm 2 and theoretical results. $\beta = 0.1$.

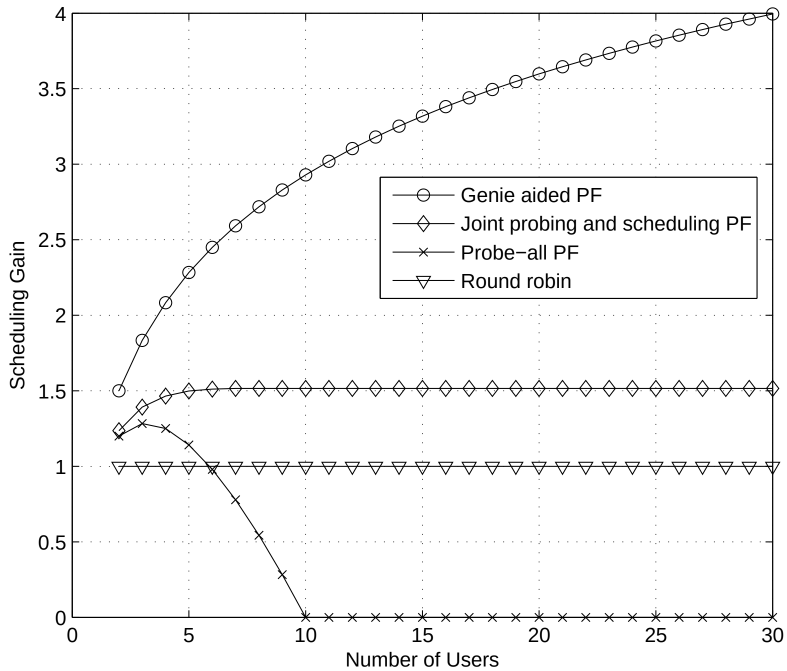

Fig. 7: Scheduling gain VS number of users. $\beta = 0.1$.

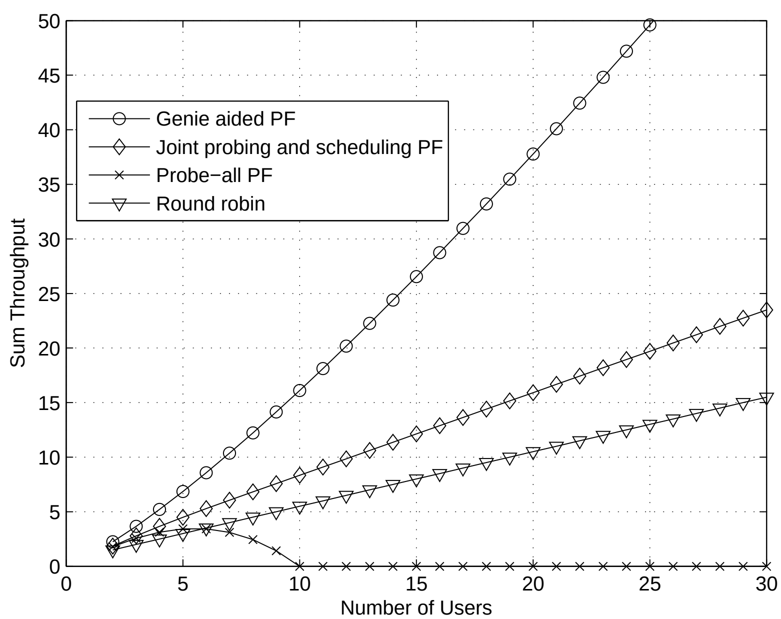

Fig. 8: Sum throughput VS number of users. $\beta = 0.1$.