# Optimal Joint Probing and Transmission Strategy for Maximizing Throughput in Wireless Systems

Prasanna Chaporkar and Alexandre Proutiere

***Abstract*— In broadcast fading channel, channel variations can be exploited through what is referred to as *multi-user diversity* and *opportunistic scheduling* for improving system performance. To achieve the gains promised by this kind of diversity, the transmitter has to accurately track the channel variations of the various receivers, which consumes resources (time, energy, bandwidth), and thus reduces the resources remaining for effective data transmissions. The transmitter may decide not to acquire or *probe* the channel conditions of certain receivers, either because these receivers are presumably experiencing severe fading, or because the transmitter wishes to spare resources for data transmissions. It may also decide to transmit to a receiver without probing its channel; in such cases, the transmitter *guesses* the channel state, which often results in a reduction of the transmission rate compared to when the transmitter knows the channel state. Ultimately, the transmitter has to decide to which receiver it should transmit. In this paper, we identifying the joint probing and transmission strategies realizing the optimal trade-off between the channel state acquisition and the effective data transmission. The objective is to maximize the system throughput. Finally, we propose several extensions of the proposed strategy, including a scheme to maximize the system utility and a scheme to ensure the system stability.**

***Index Terms*— Limited information MAC, stochastic control, generalized optimal stopping time problem.**

## I. Introduction

FADING variations between the transmitter and the receiver have traditionally been considered to have an adverse impact on the performance of the communication in wireless systems. It is well known that the capacity of the AWGN point-to-point channel with ergodic fading is less than that of the unfaded AWGN channel even when the channel side information (CSI) is available at the transmitter and the receiver [1], [2]. Recently however, Knopp and Humblet [3] have shown that fading could be exploited to increase the throughput in broadcast systems, where a single transmitter has to send data to several receivers with various and independent fading conditions. There, the throughput improvement is achieved by always transmitting to receivers with relatively favorable channel conditions. When the number of receivers is sufficiently large, there is always, and with high probability, a receiver whose channel conditions are better than in average. This principle is often termed *multi-user diversity*, and transmitting to users with relatively favorable conditions is referred to as *opportunistic scheduling*.

Manuscript received November 2, 2007; revised April 15, 2008. Work by P. Chaporkar is partly sponsored by Bharati Centre for Communications. Received Nov 1, 2007, revised April 15, 2008.

P. Chaporkar is with Indian Institute of Technology Bombay, Mumbai, India. A. Proutiere is with Microsoft Research, Cambridge, UK. (e-mails: chaporkar@ee.iitb.ac.in, alexandre.proutiere@microsoft.com).

Digital Object Identifier 10.1109/JSAC.2008.081020.

This promise of throughput gain via multi-user diversity instigated significant research efforts in developing theoretical and practical opportunistic scheduling schemes. Such schemes have been designed for various performance objectives. For example, Lyapunov-based opportunistic scheduling schemes that can provide the required throughput to each of the receivers when doing so is at all possible have been proposed in [4], [5]; schemes that provide delay differentiation are proposed in [6], [7], [8]; schemes that minimize the maximum mean queueing delay are developed in [9]; and more recently, opportunistic schedulers that maximize/minimize certain utility while providing the required throughput to each of the receivers have been proposed in e.g. [10], [11], [12], [13].

In all the work mentioned above, the basic underlying assumption is that the CSI between the transmitter and each of the receivers are known at both ends. But, the CSI is not automatically available, instead it has to be acquired; and this acquisition consumes resources like time, bandwidth and power, e.g. in CDMA/HDR cellular systems [10], a dedicated channel for each receiver is maintained for communicating the CSI to the base station. Moreover, the resource consumption is proportional to the number of receivers in the system. Thus, one has to carefully evaluate the trade-off between resource consumption in acquiring CSI and the performance improvement by opportunistically using the acquired information. In this paper, our aim is to evaluate this trade-off and propose an optimal CSI acquisition and transmission policy to maximize the rate at which the transmitter can send data to the receivers, i.e. to maximize the system throughput.

We motivate the problem with the following example.

*Example 1:* Consider a broadcast channel with two receivers, e.g., the down-link of a cellular network or of a wireless LAN. Time is slotted, and a slot is of unit duration. Here, the slot is assumed to represent the coherence time of the channels. We further assume that the receivers experience independent and identically distributed (i.i.d.) fading in each slot. Specifically, let the maximum rate of transmission in any slot be 1 with probability 1/2 and 2 w.p. 1/2 independently for each of the receivers given that CSI is known to both, the transmitter and the receiver. Also, let $\beta$ denote the fraction of slot duration required to probe a receiver and acquire its

CSI. In this setting, we compare two probing and transmission strategies $\pi_1$ and $\pi_2$. Under $\pi_1$, the transmitter probes both receivers, and then transmits to the receiver with the best channel state. Ties are broken arbitrarily. Under $\pi_2$, the transmitter probes one receiver at random, and transmits to it at the maximum rate possible. Policy $\pi_1$ spends $2\beta$ units of time per slot to acquire the channel states, and transmits at an expected rate of 7/4 in the remainder of the slot; while $\pi_2$ spends only $\beta$ units of time per slot to acquire the CSI, but transmits at a smaller expected rate of 3/2 in the remainder of the slot. Thus, the expected throughput under $\pi_1$ is $7(1-2\beta)/4$, while that under $\pi_2$ is $3(1-\beta)/2$. Note that if $\beta \leq 1/8$, then $\pi_1$ has a higher throughput than that of $\pi_2$. But, when $\beta > 1/8$, the throughput under $\pi_2$ is higher. ■

The above example demonstrates that the probing and transmission strategy should be designed by taking into account the cost for probing, which is time in the example. Since, the channel states change at each slot, whose duration is the coherence time of the channels, the CSI obtained in a given slot can not be used in subsequent slots. Thus, if the time required for probing a receiver consumes a significant portion of the coherence time, then probing only a small number of receivers may provide the optimal throughput. On the other hand, if the time required for probing is a small fraction of the coherence time, then probing a larger number of receivers may be optimal as it allows to discover receivers with high channel gains and thereby to achieve a high throughput.

In Example 1, policy $\pi_1$ can be trivially modified to provide better throughput in the following way. If at the first probe, $\pi_1$ finds a receiver to which transmission at rate 2 is possible, then it does not probe the second receiver as no further improvement in the transmission rate is possible. With this modification, $\pi_1$ achieves throughput of $\frac{7}{4}\left(1-\frac{10}{7}\beta\right)$ instead of $\frac{7}{4}(1-2\beta)$. This shows that the decision to probe further should depend on the channel states observed in the previous probes.

Another problem in designing an optimal probing and transmission strategy is that of deciding the order in which the receivers should be probed. In Example 1, we have considered i.i.d. channel states, and hence probing sequence does not matter. But, in the following example we demonstrate that when the channel states are not i.i.d. across receivers, then the sequence in which receivers are probed has a significant bearing of the achievable throughput.

*Example 2:* Consider the same settings as in Example 1, except that the channel gains are not i.i.d. across receivers. Specifically, in each slot, let the maximum rate to receiver $R_1$ be 2 w.p. $(k-1)/k$ and $k$ w.p. $1/k$, and for receiver $R_2$ let it be 1 w.p. $(2k-1)/2k$ and $2k$ w.p. $1/2k$. Now, the expected transmission rates to $R_1$ and $R_2$ are $\frac{2(k-1)}{k}+1$ and $\frac{2k-1}{2k}+1$, respectively. Thus, for $k > 3/2$, the expected rate to $R_1$ is strictly greater than that to $R_2$. Fix $k > 3/2$. In this setting, one would intuitively expect that probing $R_1$ first should be optimal as it provides a higher expected rate, but we show that if $\beta < \frac{2k^2}{8k^2-7k+2}$, then probing $R_2$ first provides a better throughput. Specifically, we show that the optimal policy $\pi^\star$ is, in every slot, to probe $R_2$ first. If the achievable rate is $2k$ then transmit to $R_2$, otherwise probe $R_1$ and transmit to it at the appropriate rate. The expected throughput of $\pi^\star$ is $(1-\beta)+(1-2\beta)\frac{6k^2-7k+2}{2k^2}$. To show that $\pi^\star$ achieves the highest throughput, it suffices to compare it with policy $\pi_1$ that probes $R_1$ and transmits at appropriate rate, and with policy $\pi_2$ that probes $R_1$ first. If the achievable rate is $k$, then $\pi_2$ transmits to $R_1$, otherwise it probes $R_2$ and transmits to it if the achievable rate is $2k$, else it transmits to $R_1$ at rate 2. Note that the throughput of $\pi_1$ is $(1-\beta)\left[\frac{2(k-1)}{k}+1\right]$, while that of $\pi_2$ is $(1-\beta)+(1-2\beta)\frac{3k^2-4k+1}{k^2}$. It is easy to verify that throughput of $\pi_2$ is always smaller than that of $\pi^\star$, while the throughput of $\pi_1$ is smaller than that of $\pi^\star$ when $\beta < \frac{2k^2}{8k^2-7k+2}$. Thus, probing $R_2$ provides the optimal throughput for $\beta < \frac{2k^2}{8k^2-7k+2}$. ■

The above example demonstrates that the system throughput depends on the order in which the receivers are probed. The example also demonstrates that heuristics like probing receivers in the order of their expected rates may not be optimal.

Until now we have considered a case where the transmitter transmits only to a probed receiver. But, in practice the transmitter may decide to transmit to a receiver by guessing, instead of probing, its CSI. Guessing saves the time required to probe the receiver, and this time may be used for the actual data transmission. But, the rate of transmission to an un-probed receiver can not be greater than the rate of this receiver averaged over its fading conditions. Clearly, the opportunity to guess leads to a more complicated possible trade-off: A probing and transmission strategy has to adaptively decide whether to (i) transmit to a probed receiver, (ii) probe a new receiver, or (iii) guess the channel state of an un-probed receiver and transmit to this receiver. The decision will depend upon the fraction of the channel coherence time required to probe a receiver, the receivers' channel gain distributions, and the channel gain values observed for the probed receivers. *Our aim is to obtain a probing and transmission strategy maximizing the expected system throughput.*

Though we have motivated the problem for broadcast fading channel, similar problem exists in other wireless systems of practical interest. We present the following example.

*Opportunistic spectrum access in multi-channel wireless networks:* In future wireless systems, a receiver will be able to access a large number of channels. For example, in IEEE 802.11 based wireless LANs, three orthogonal channels are available for communication. Also, in other systems such as cognitive radio systems, the transmitter has to choose the frequency band to communicate so as to avoid interference with licensed or unlicensed receivers. Here, to maximize its throughput, the transmitter should use a channel with high SINR. Thus, again, the transmitter has to probe channels to look for the best possible and this has a cost, as it reduces the time remaining for the effective transmission. The problem is clearly identical to the first problem mentioned above.

In this paper, we formulate the problem of designing an optimal probing and transmission strategy as a stochastic control problem (Section II), and explain how this problem is different from the classical stochastic control problems. We derive structural properties of optimal probing and transmission strategies, and in some special cases of practical importance, we completely characterize these strategies (Section III).

We then illustrate our theoretical findings through numerical results (Section IV), and conclude the paper (Section IV) by proposing several extensions of the work.

## II. Problem Formulation and Related Work

### A. System Model and Problem Formulation

We consider a system with $N$ receivers whose channel conditions vary over time. Time is slotted, and the channel conditions of the various receivers are assumed to remain constant for the duration of one slot, i.e., the coherence time of the channels are larger than one slot; these conditions may change at the slot boundaries. In other words, we consider the block fading model [2]. Denote by $c_i(t)$ the channel state of user $i$ during slot $t$. Now the transmission rate at which a user whose channel is in state $c \in \mathbb{R}^+$ can receive is denoted by $R(c)$ where $R(\cdot)$ is an increasing function. For example, $R$ can represent Shannon limit: $R(c) = W \log_2\left(1 + \frac{Pc}{WN_0}\right)$, where $W$ is the channel bandwidth and $P, N_0$ are the transmission power and the noise power spectral density, respectively. Here the channel state $c$ represents the fraction of transmission power received by the user. We assume that the channel states are independent across receivers, but the distributions of the channel state of each receiver may be different. For a given receiver, the channel states are i.i.d. across slots with cumulative distribution function (c.d.f.) $F_i(\cdot)$ ($F_i(a) = Pr[c_i(t) \leq a]$). In the following, we denote by $C_i$ a generic random variable (r.v.) with c.d.f. $F_i$. We assume that $F_i$ is known to the transmitter and the receiver.

At the beginning of each time slot, the sender can decide to probe some channels, to transmit to one of the probed receivers, or to transmit to a receiver that has not been probed[1]. We assume that probing the channel state of a receiver takes a proportion $\beta$ of the slot duration. Hence, in a given slot, when the transmitter decides to transmit to a receiver whose channel state is $c$, where $c$ can be either known or unknown, the throughput during this slot is:

$$T = (1 - \beta|\mathcal{P}|)R(c),$$

where $\mathcal{P}$ denotes the set of probed receivers in that slot, and $|\mathcal{P}|$ is the cardinality of this set. We denote by $\overline{\mathcal{P}}$ the set $\{1, \ldots, N\} \setminus \mathcal{P}$.

*Definition 1 (Probing Strategy):* A probing strategy is an algorithm that given the set $\mathcal{P}$ of probed receivers and the channel gains for the receivers in $\mathcal{P}$, decides whether to probe a receiver in $\overline{\mathcal{P}}$; and if the decision is to probe, then it also decides which receiver should be probed.

*Definition 2 (Transmission Strategy):* A transmission strategy is a rule that identifies a receiver to which the transmitter should transmit given the set $\mathcal{P}$ of probed receivers and the channel gains for the receivers in $\mathcal{P}$.

We note that a transmission policy need not always transmit to a receiver in $\mathcal{P}$, but it may also decide to guess the channel gain for a receiver in $\overline{\mathcal{P}}$ and transmit to it.

*Definition 3 (Joint Probing and Transmission Strategy):* A joint probing and transmission strategy $\pi$ is an algorithm that given the set $\mathcal{P}$ of probed receivers and the channel gains for the receivers in $\mathcal{P}$, decides among one of the following three actions: (i) transmit to some receiver $i \in \mathcal{P}$, (ii) transmit to some receiver $i \in \overline{\mathcal{P}}$, (iii) probe some receiver $i \in \overline{\mathcal{P}}$.

*Definition 4 (System Throughput):* Let $B^\pi(t)$ denote the number of bits transmitted in slot $t$ under policy $\pi$. Then, the system throughput under $\pi$ is

$$T^\pi \stackrel{\text{def}}{=} \liminf_{t\to\infty} \frac{\sum_{s=1}^{t} B^\pi(s)}{t}.$$

Our problem is to design a joint probing and transmission strategy that maximizes the system throughput. Such a strategy is said to be optimal. Since the system is i.i.d. across slots, by the strong law of large numbers, maximizing the system throughput is equivalent to maximizing the expected throughput in each slot. For notational simplicity, we consider any given slot and drop time $t$ from the notation.

We formulate the problem of maximizing the expected throughput in a slot as a stochastic control problem. The formulation is as follows. Assume that the receivers in set $\mathcal{P}_k$ has been already probed, where the subscript $k$ indicates that $|\mathcal{P}_k| = k$ ($k$ users have been probed already). Denote by $u$ the largest channel gain among the receivers in $\mathcal{P}_k$; we say that the system is in state $(\mathcal{P}_k, u)$. Note that since the system throughput is maximized by transmitting at the highest possible rate, we only need to maintain the maximum observed channel gain $u$. Then a strategy $\pi$ has the following possible control actions in state $(\mathcal{P}, u)$):

1) Transmit to the receiver with the best probed channel. In that case, the throughput will be: $T_{\text{tr}}(\mathcal{P}_k, u) = (1 - k\beta)R(u)$.
2) Transmit to a receiver that has not been probed. In that case, the throughput will be: $T_{\text{g}}(\mathcal{P}_k, u) = (1 - k\beta)\max_{i\in\overline{\mathcal{P}}_k} R_{g(i)}$, where $R_{g(i)}$ denotes the expected rate at which the transmitter can send data to user $i$ without knowing its current channel state. The exact value of $R_{g(i)}$ depends on the advanced coding and signaling schemes used in the system. For example, if the receiver knows the channel state, which is the case when the transmitter broadcasts pilot signal at the beginning of the slot, then $R_{g(i)}$ can be close to $\mathbb{E}[R(C_i)]$. To achieve the latter rate, the coding scheme should be able to reveal the ergodic nature of the channel. This is not always possible depending on the system considered. Hence, it will be relevant to study the special case where transmitting to an un-probed user is not allowed.
3) Probe one more receiver (say $i$) from the set $\overline{\mathcal{P}}$. In this case, the system state changes from $(\mathcal{P}_k, u)$ to $(\mathcal{P}_{k+1}, u \vee c_i)$ given that the channel state of the newly probed user is $c_i$, and where $\mathcal{P}_{k+1} = \mathcal{P}_k \cup \{i\}$. The operator $\vee$ is defined by $a \vee b = \max(a, b)$.

In cases (i) and (ii), we say that we *retire*. In cases (ii) and (iii), the strategy has also to define to which receiver to transmit and which receiver to probe, respectively. We denote by $T^\pi(\mathcal{P}_k, u)$ the average throughput achieved by strategy $\pi$,

[1]Note here that transmitting to an un-probed receiver requires advanced adaptive coding schemes, and often, it is not possible. That is why the case where the sender has to probe a channel before using it is quite relevant.

starting from system state $(\mathcal{P}_k, u)$. Also denote by $T^\star(\mathcal{P}_k, u)$ the average throughput of an optimal strategy starting from system state $(\mathcal{P}_k, u)$. Now, $T^\star(\mathcal{P}_k, u)$ is given by the following Bellman's equation for all $k$, $\mathcal{P}_k$ and $u$.

$$\begin{aligned} T^\star(\mathcal{P}_k, u) &= \max\{T_{\rm tr}(\mathcal{P}_k, u), T_{\rm g}(\mathcal{P}_k, u), \\ &\quad \max_{i \in \overline{\mathcal{P}}_k} \{\mathbb{E}_i\left[T^\star(\mathcal{P}_k \cup \{i\}, u \vee C_i)\right]\}\}, \end{aligned} \quad (1)$$

where $\mathbb{E}_i[\cdot]$ is the expectation taken w.r.t. $F_i$. Thus, in each state $(\mathcal{P}_k, u)$, the optimal control decision corresponds to the term that achieves the maximum in (1), e.g., if $T_{\rm tr}(\mathcal{P}_k, u)$ achieves the maximum then the optimal decision is to transmit. Note that at the beginning of every slot, the state is $(\emptyset, 0)$, where $\emptyset$ denotes the empty set. Our aim is to obtain an optimal strategy $\pi^\star$ in the sense that $T^{\pi^\star}(\emptyset, 0) = T^\star(\emptyset, 0)$.

Before providing some properties of an optimal strategy $\pi^*$, we first explain how the problem considered here is different than all other stochastic control problems previously analyzed.

### B. Related work

The problem of identifying optimal joint probing and transmission strategies has been addressed in the literature recently only [14], [15], [16], [17], [18]. It falls into the broad class of stochastic control problems [19]. However, as explained in [20], it does not correspond to any of the existing classical control problems such as multi-armed bandits, optimal sampling order, or optimal stopping problems. In the various versions of the multi-armed bandit problems [21], [22], acquiring the state of an arm (or of a channel here) before using it is not allowed. Optimal sampling order of random variables has been investigated in many contexts, see e.g. [23], [24]; however, in all existing work, these variables can take 2 values only (On or Off channels here), and exploiting a variable that has not been probed is not allowed. Finally, in usual stopping time problems [25], one has to select between two possible actions, proceed further or stop; this can be applied to our problem only when all channels are equivalent [15], i.e., when they have the same statistical distribution. The latter assumption is never valid in practical scenarios. In any case, stopping time problems are very challenging and most of them are open [19].

We discuss now papers specifically related to optimal joint probing and transmission strategies. In most of these papers (see e.g., [15], [20]), a linear cost structure is used, which means that the *reward* or throughput can be written as: $T = R(c) - |\mathcal{P}|\beta$. This considerably simplifies the analysis, although characterizing an optimal strategy is an open problem, unless transmitting to an un-probed user is not allowed [20]. In practice, the cost structure is logarithmic, as in our model where $\log T = \log R(c) + \log(1 - \beta|\mathcal{P}|)$. To our knowledge, [18] is the only paper aiming at analyzing the latter model, but the results in [18] are very preliminary: only rough structural properties of the optimal probing and scheduling strategy are stated without proof. In the present work, we provide general structural properties of the optimal strategy, but also exactly characterize this strategy in specific but relevant cases.

## III. Structural Properties of the Optimal Strategy

In this section, we state some structural properties that an optimal probing and transmission strategy should have. Specifically, when the system is in some state $(\mathcal{P}_k, u)$, we will give conditions under which an optimal strategy should either transmit to one of the probed users, or transmit to an un-probed user, or probe another user. These conditions do not fully characterize an optimal strategy, as it remains to define which user to probe next if the strategy decides to further probe. The latter question is a much more challenging issue than deriving the basic structural properties of an optimal strategy. We address this question of obtaining optimal probing strategy in certain special cases of practical interest.

We start our exploration in a special case where guessing is not allowed, i.e., in the state $(\mathcal{P}_k, u)$, a joint probing and transmission strategy can either transmit to some $i \in \mathcal{P}_k$ or probe some $j \in \overline{\mathcal{P}}_k$. This special case is of practical interest as already explained in Section II. Additionally, studying this special case provides valuable insights in designing optimal policy when guessing is allowed, as we shall see in Section III-B.

### A. Optimal Strategy when Guessing is not Allowed

When guessing is not allowed, the Bellman's equation (1) reduces to

$$\begin{aligned} T^\star(\mathcal{P}_k, u) &= \max\{T_{\rm tr}(\mathcal{P}_k, u), \\ &\quad \max_{i \in \overline{\mathcal{P}}_k} \{\mathbb{E}_i\left[T^\star(\mathcal{P}_k \cup \{i\}, u \vee C_i)\right]\}\}. \end{aligned} \quad (2)$$

We note that when the number of possible channel states is finite for each user, it is indeed possible to solve (1), and thereby obtain an optimal strategy. But, the brute force computation has exponential (in terms of number of users) complexity as the quantity $T^\star(\mathcal{P}_k, u)$ has to be evaluated for every subset $\mathcal{P}_k$. So, deriving properties of optimal strategies is crucial, either to exactly characterize these strategies or to reduce their computational complexity.

Let $a_k \stackrel{\rm def}{=} (1 - k\beta)$. Define $T_{{\rm pr}(i),{\rm tr}}(\mathcal{P}_k, u) \stackrel{\rm def}{=} a_{k+1}\mathbb{E}_i[R(u \vee C_i)]$, for $i \in \overline{\mathcal{P}}_k$, and $T_{\rm pr,tr}(\mathcal{P}_k, u) \stackrel{\rm def}{=} a_{k+1}\max_{i \in \overline{\mathcal{P}}_k} \mathbb{E}_i[R(u \vee C_i)]$. The quantity $T_{{\rm pr}(i),{\rm tr}}(\mathcal{P}_k, u)$ is the expected throughput that can be achieved, starting from state $(\mathcal{P}_k, u)$, when we probe just one additional user $i \in \overline{\mathcal{P}}_k$ and then transmit to the best probed user. We will show that to obtain an optimal transmission strategy, it suffices to consider the one-step-look-ahead throughput $T_{\rm pr,tr}(\mathcal{P}_k, u)$, rather than $T^\star(\mathcal{P}_k, u)$ in (2). Since unlike $T^\star(\mathcal{P}_k, u)$, $T_{\rm pr,tr}(\mathcal{P}_k, u)$ can be computed with complexity $O(N)$, this considerably reduces the complexity of computing an optimal strategy. Let $\pi^\star_{\rm NG}$ denote the optimal policy when guessing is not allowed.

*Theorem 1:* Let $(\mathcal{P}_k, u)$ be the system state. Then, $\pi^\star_{\rm NG}$ transmits to the receiver with the best channel gain in $\mathcal{P}_k$ if and only if $T_{\rm tr}(\mathcal{P}_k, u) \geq T_{\rm pr,tr}(\mathcal{P}_k, u)$.

*Proof:* The proof is presented in Appendix I. ■

Theorem 1 states that $\pi^\star_{\rm NG}$ can determine when to probe by considering one-step-look-ahead throughput only. But, Theorem 1 does not determine which user to probe when $\pi^\star_{\rm NG}$ decides to probe an additional receiver. We believe that obtaining an optimal probing strategy is much more challenging question in general settings, and it remains open. As Example 2 demonstrates, probing a receiver $i \in \overline{\mathcal{P}}_k$ that maximizes $T_{{\rm pr}(i),{\rm tr}}(\mathcal{P}_k, u)$ may not be optimal. Recall that in the example, receiver 1 maximizes one-step-look-ahead throughput in state $(\emptyset, 0)$, but probing user 2 was optimal for certain values of $\beta$. Nevertheless we are able to determine an optimal probing strategy when the channels are stochastically ordered as defined below.

*Definition 5 (Stochastically Ordered Channels):* The channels of the $N$ users are stochastically ordered if there exists a permutation $\sigma$ of $\{1, \ldots, N\}$ such that for all $i, j$, if $\sigma(i) \leq \sigma(j)$, then $C_{\sigma(j)} \leq_{st} C_{\sigma(i)}$, where $X \leq_{st} Y$ if and only if for all increasing function $f$ such that $\mathbb{E}[f(Y)] < +\infty$, $\mathbb{E}[f(X)] \leq \mathbb{E}[f(Y)]$.

Without loss of generality, when the channels are stochastically ordered, we assume that the permutation $\sigma$ is $\sigma(i) = i$ for all $i$. Note that having a stochastic order on the channels is equivalent to having a similar order for the corresponding rates (i.e., $C_j \leq_{st} C_i$ iff $R(C_j) \leq_{st} R(C_i)$). An example of ordered channels is when one can write $C_i = \mathbb{E}[C_i]Y$ where $Y$ is a fixed r.v., i.e., when the channels have similar distributions but different means. This is a quite usual fading model in wireless networks. In these settings, we obtain an optimal probing strategy.

*Theorem 2:* If the channels are stochastically ordered, then when $T_{\rm pr,tr}(\mathcal{P}_k, u) > T_{\rm tr}(\mathcal{P}_k, u)$, the optimal decision is to probe the user $j \in \overline{\mathcal{P}}_k$ such that for all $i \in \overline{\mathcal{P}}_k$, $C_i \leq_{st} C_j$.

*Proof:* The proof is presented in Appendix II. ■

Theorems 1 and 2 provide a full description of the optimal probing and transmission strategy $\pi^\star_{\rm NG}$ when the channels are stochastically ordered. Remark that even though the results of these theorems seem quite intuitive, as often in stochastic control problems, their proofs are far from being trivial; and as illustrated in the two examples presented in introduction, intuitive decisions may sometimes be sub-optimal. The optimal strategy is summarized in the following corollary.

*Corollary 1:* When the channels are stochastically ordered, the one-step-look-ahead strategy is optimal. The optimal one-step-look-ahead strategy is as follows: when the system is in state $(\mathcal{P}_k, u)$:
(i) If $T_{\rm pr,tr}(\mathcal{P}_k, u) \geq T_{\rm tr}(\mathcal{P}_k, u)$, then we should probe the stochastically largest un-probed user,
(ii) otherwise, we should transmit to the user $i \in \mathcal{P}_k$ such that $c_i = u$.

In the following subsection we obtain structural properties of the optimal policy $\pi^\star$ when guessing is allowed.

### B. Optimal Strategy when Guessing is Allowed

First, we show that in many states, the optimal policy when guessing is allowed (denoted by $\pi^\star$) takes the same decisions as the policy $\pi^\star_{\rm NG}$. Specifically, we have the following result.

*Theorem 3:* In every state $(\mathcal{P}_k, u)$ such that $T_{\rm tr}(\mathcal{P}_k, u) \geq T_{\rm g}(\mathcal{P}_k, u)$, $\pi^\star$ and $\pi^\star_{\rm NG}$ take identical decisions. Moreover, after probing a receiver in state $(\mathcal{P}_k, u)$, we also have in the new system state, say $(\mathcal{P}_{k+1}, u')$, $T_{\rm tr}(\mathcal{P}_{k+1}, u') \geq T_{\rm g}(\mathcal{P}_{k+1}, u')$.

*Proof:* The proof is presented in Appendix III. ■

Theorem 3 states that once we have probed a receiver with channel gain large enough to provide a greater throughput than that we would obtain by guessing and transmitting to any other un-probed receiver, the optimal policy does not need to consider guessing any further. In other words, the optimal decisions from this state is either to transmit to the probed receiver or to probe a new receiver, but never to guess and transmit to an un-probed receiver. Thus, after reaching a state $(\mathcal{P}_k, u)$ such that $T_{\rm tr}(\mathcal{P}_k, u) \geq T_{\rm g}(\mathcal{P}_k, u)$, the optimal policy $\pi^\star$ is as described in Corollary 1.

In view of Theorem 3, we have characterized $\pi^\star$ except in states $(\mathcal{P}_k, u)$ such that $T_{\rm tr}(\mathcal{P}_k, u) < T_{\rm g}(\mathcal{P}_k, u)$. Now, we provide the structural properties of $\pi^\star$ in these states. First, we introduce some notation. For a receiver $i$, let $u_{{\rm g}(i)}$ be a state such that $R(u_{{\rm g}(i)}) = R_{{\rm g}(i)}$ ($u_{{\rm g}(i)}$ denotes the channel state corresponding to the rate one would obtain by guessing and transmitting to the receiver $i$; without loss of generality,we assume that such state exists). Also, let $u_{{\rm g}(\mathcal{P}_k)} \overset{\rm def}{=} \max_{i \in \overline{\mathcal{P}}_k} \{u_{{\rm g}(i)}\}$.

*Theorem 4:* In every state $(\mathcal{P}_k, u)$ such that $T_{\rm tr}(\mathcal{P}_k, u) < T_{\rm g}(\mathcal{P}_k, u)$:

1) If $T_{\rm pr,tr}(\mathcal{P}_k, u) \geq T_{\rm g}(\mathcal{P}_k, u)$, then $\pi^\star$ probes some receiver $i \in \overline{\mathcal{P}}_k$.
2) If $T_{\rm pr,tr}(\mathcal{P}_k, u) < T_{\rm g}(\mathcal{P}_k, u)$, then $\pi^\star$ satisfies:
   a) If $T_{\rm pr,tr}(\mathcal{P}_k, u_{{\rm g}(\mathcal{P}_k)}) \leq T_{\rm tr}(\mathcal{P}_k, u_{{\rm g}(\mathcal{P}_k)})$, then $\pi^\star$ guesses and transmits to $j = \arg\max_{i \in \overline{\mathcal{P}}_k} \{u_{{\rm g}(i)}\}$.
   b) If there exists $j \in \overline{\mathcal{P}}_k$ such that $T_{{\rm pr}(j),{\rm tr}}(\mathcal{P}_k, u_{{\rm g}(\mathcal{P}_k \cup \{j\})}) \geq T_{\rm tr}(\mathcal{P}_k, u_{{\rm g}(\mathcal{P}_k)})$, then $\pi^\star$ probes some $i \in \overline{\mathcal{P}}_k$.

*Proof:* The proof is presented in Appendix III. ■

Theorems 3 and 4 provide quite detailed structure properties of the optimal strategy when guessing is allowed. Furthermore, we know which user to probe next when the system state $(\mathcal{P}_k, u)$ is such that $T_{\rm tr}(\mathcal{P}_k, u) \geq T_{\rm g}(\mathcal{P}_k, u)$. To fully characterize the optimal strategy, one need to identify which user to probe next when the system state is such that $T_{\rm tr}(\mathcal{P}_k, u) < T_{\rm g}(\mathcal{P}_k, u)$. This problem remains open even in the case where the optimal control problem has a linear cost structure as in [20]. Fortunately, in practical scenarios, as observed in the next section, states such that $T_{\rm tr}(\mathcal{P}_k, u) < T_{\rm g}(\mathcal{P}_k, u)$ appear rarely, which simplifies the characterization of the optimal strategy.

## IV. Extensions: Utility and Queues

### A. Maximizing System Utility

In the case of the broadcast channel, one may propose to impose fairness among receivers, i.e., to maximize a certain notion of system utility instead of the system throughput as we have done so far. Denote by $U(\cdot)$ a concave non-decreasing utility function, and denote by $B_i^\pi(t)$ the throughput received by user $i$ under strategy $\pi$ in slot $t$. The long-term throughput of user $i$ is then $T_i^\pi = \liminf_{t\to\infty} \frac{1}{t}\sum_{s=1}^t B_i^\pi(s)$. Now the objective is to maximize $\sum_{i=1}^N U(T_i^\pi)$. Finally, let $\hat{B}_i^\pi(t)$ be the expected (with respect to the channel state distributions) throughput received by user $i$. A natural candidate algorithm to maximize the system utility is the following gradient algorithm: each time slot $t$, choose joint probing and transmission

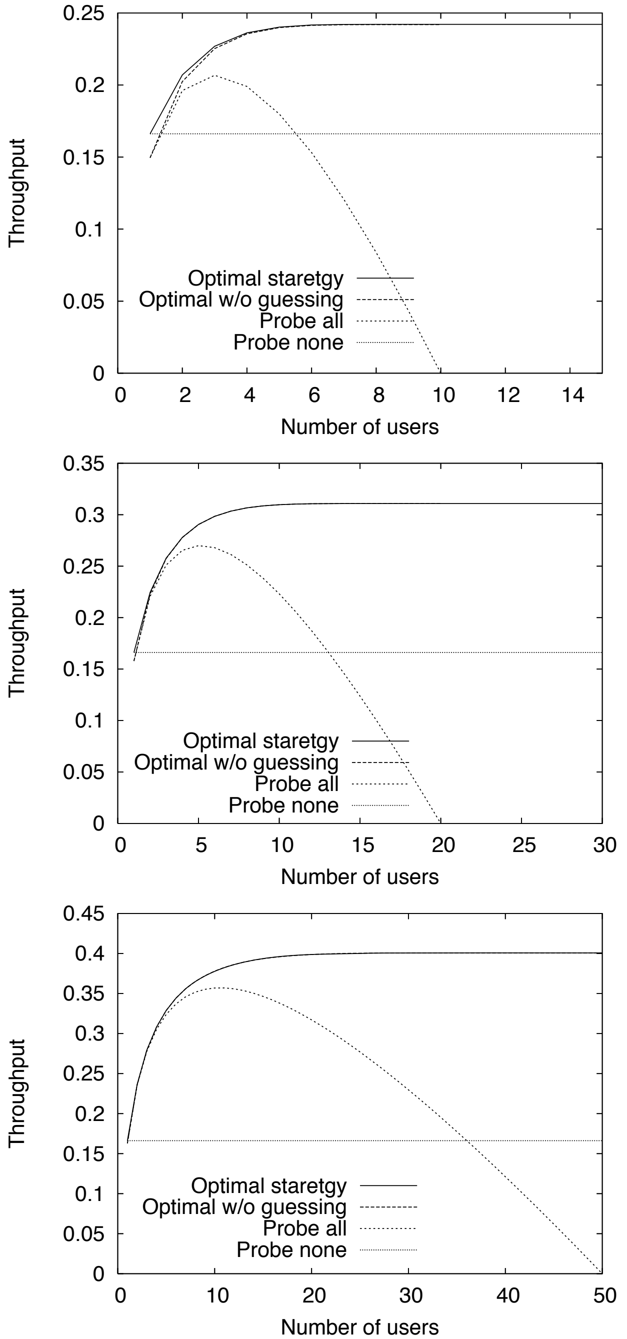


Fig. 1. Average throughput of the various strategies as the number of users $N$ increases - Exponential channels with different means - $\beta = 0.1$ (upper figure), 0.05 (middle figure), 0.02 (lower figure).

strategy as follows.

$$\max_{\pi} \sum_{i=1}^{N} \hat{B}_i^{\pi}(t) \times U'(\overline{t}_i(t)), \quad (3)$$
$$\overline{t}_i(t+1) = (1-\eta)\overline{t}_i(t) + \eta t_i(t).$$

Note that solving (3) is equivalent to maximizing a weighted sum of expected throughputs each time slot. This can be done as in Section III. The above algorithm is expected to maximize the system utility when the parameter $\eta$ tends to 0. We do not provide a detailed analysis here, and reserve it for future work.

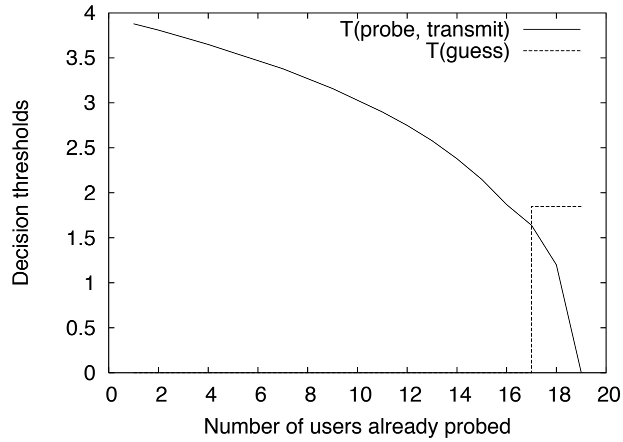


Fig. 2. Decision thresholds as a function of the set of already probed users - $\beta = 0.05$, $N = 20$ users.

### B. Queue Stability

Alternatively, one can also study the system with queues corresponding to each of the receivers in which the bits arriving from the higher layer are stored. The problem, in these settings, is to obtain a scheduling policy that stabilizes the system, i.e., provides the finite expected delay to each of the users. Let $Q_i(t)$ denote the queue length of user $i$ in slot $t$. It is well known that the max-weight policies stabilize the system if doing so is possible [26]. We can use a similar idea here and choose joint probing and transmission strategy that solves $\max_{\pi}\{\sum_{i=1}^{N} Q_i(t) \times \hat{B}_i^{\pi}(t)\}$ given the current queue lengths. Since, $Q_i(t)$ is known, finding the maximum is equivalent to maximizing the weighted sum of the throughputs, which can be done using results in Section III. Again, a detailed analysis will be provided in future work.

## V. Numerical Results

In this section, we give some numerical experiments illustrating the theoretical findings of the previous sections. We compare the following probing and transmission strategies: (a) the optimal strategy when guessing is allowed ($\pi^*_{\mathrm{NG}}$); (b) the optimal strategy when guessing is not allowed ($\pi^*$); (c) the strategy where all channels are probed before transmission; (d) the strategy where no channel is probed, i.e., where the transmission is made on the channel with the highest average state. The policy $\pi^*_{\mathrm{NG}}$ is obtained using brute force computations in which the results from the previous section have been utilized.

We consider an asymmetric fading scenario: the channel states of the various users are exponentially distributed but with different means. We further assume that these averages are ordered, i.e., the channels are stochastically ordered. The averages are linearly decreasing with the channel index $i$. For a given channel state, the corresponding rate follows Shannon formula ($P = 40$dBm, $N_0 = -100$dBm, $W = 1$). With a path loss exponent equal to -3.5, the user with the worst average channel is located roughly 2 times further from the transmitter than the user with the best channel.

In Figure 1, we present the average throughputs of strategies (a)-(d) when the number of users grows and for different values of $\beta$, the proportion of slot required to probe a channel. Note that the optimal strategies with or without guessing have very similar performance except when the probing cost $\beta$ is very large. In fact, in this example, it turns out that the optimal strategy transmits to a un-probed user very rarely, and only when a lot of users have been probed already and when the observed channel state is still low. This observation is confirmed in Figure 2: here we consider the case where $\beta = 0.05$ and $N = 20$ users. The curve T(probe,transmit) represents the value of the maximum channel state for which it is better to further probe than transmit to an already probed user. T(guess) shows the maximum channel state for which it is optimal to guess and transmit to a user that has not been probed.

## VI. Conclusion

In exploiting multi-user diversity, there is an inherent trade-off between the consumption of resources to probe the channel states of the receivers, and the throughput improvement obtained by opportunistic scheduling. We have shown that acquiring the CSI of all the receivers can in fact reduce the system throughput compared to that obtained when CSI of only a few receivers are acquired. We have proposed guidelines to the design of a joint probing and transmission strategy that maximizes the system throughput. Additionally, we have fully characterized the optimal strategy in some specific, but relevant, cases. We also mention how our framework can be used to provide fairness and to ensure queue stability.

## Appendix

## Appendix I
## Proof of Theorem 1

We prove the theorem using the following two supporting lemmas.

Fix the set $\mathcal{P}_k$ and define $\mathcal{D}_k$ as follows:

$$\mathcal{D}_k = \{u : T_{\text{tr}}(\mathcal{P}_k, u) \geq T_{\text{pr,tr}}(\mathcal{P}_k, u)\}. \quad (4)$$

*Lemma 1:* There exists $u_{\max}(\mathcal{P}_k)$ such that $\mathcal{D}_k = \{u : u \geq u_{\max}(\mathcal{P}_k)\}$.

*Proof:* Let $u \in \mathcal{D}_k$ and consider any $u' > u$.

$$(1-k\beta)R(u) \geq (1-(k+1)\beta) \max_{i \in \overline{\mathcal{P}}_k} \{\mathbb{E}_i\left[R(u \vee C_i)\right]\},$$

Now, it follows that

$$\begin{aligned}
&(1-k\beta)R(u)\\
&\geq (1-(k+1)\beta)\mathbb{E}_i\left[R(u \vee C_i)\right] \ \forall\ i \in \overline{\mathcal{P}}_k,\\
&= (1-(k+1)\beta)\left[R(u)F_i(u) + \int_u^\infty R(x)\mathbf{d}F_i(x)\right]\\
&\geq (1-(k+1)\beta)\left[R(u)F_i(u') + \int_{u'}^\infty R(x)\mathbf{d}F_i(x)\right].
\end{aligned}$$

From above, we can conclude the following.

$$\begin{aligned}
&(1-k\beta)R(u)[1-F_i(u')]\\
&\quad\geq \quad (1-k\beta)\int_{u'}^\infty R(x)\mathbf{d}F_i(x) - \beta\mathbb{E}_i\left[R(u \vee C_i)\right],\\
&\Rightarrow (1-k\beta)R(u')[1-F_i(u')]\\
&\quad\geq \quad (1-k\beta)\int_{u'}^\infty R(x)\mathbf{d}F_i(x) - \beta\mathbb{E}_i\left[R(u' \vee C_i)\right]. \quad (5)
\end{aligned}$$

(5) holds for every $i \in \overline{\mathcal{P}}_k$, and the lemma is proved. Note that to obtain (5), we used the fact that $R(\cdot)$ and $F_i(\cdot)$ are monotonically non-decreasing. ■

*Lemma 2:* Fix any sequence of sets of probed users such that $\mathcal{P}_{k+1} = \mathcal{P}_k \cup \{i\}$ for some $i \in \overline{\mathcal{P}}_k$ for $k \in \{0, \ldots, N-1\}$. We have: for all $k$, $\mathcal{D}_k \subseteq \mathcal{D}_{k+1}$, or equivalently $u_{\max}(\mathcal{P}_k) \geq u_{\max}(\mathcal{P}_{k+1})$.

*Proof:* The proof is by contradiction. Assume that there exist $u$ such that $u \in \mathcal{D}_k$, but $u \notin \mathcal{D}_{k+1}$. Thus,

$$\begin{aligned}
&(1-(k+1)\beta)R(u)\\
&\quad< \quad (1-(k+2)\beta) \max_{i \in \overline{\mathcal{P}}_{k+1}} \{\mathbb{E}_i\left[R(u \vee C_i)\right]\},\\
&\Rightarrow (1-k\beta)R(u) - \beta R(u)\\
&\quad< \quad (1-(k+1)\beta) \max_{i \in \overline{\mathcal{P}}_{k+1}} \{\mathbb{E}_i\left[R(u \vee C_i)\right]\}\\
&\qquad -\beta \max_{i \in \overline{\mathcal{P}}_{k+1}} \{\mathbb{E}_i\left[R(u \vee C_i)\right]\},\\
&\Rightarrow T_{\text{tr}}(\mathcal{P}_k, u) - \beta R(u)\\
&\quad< \quad (1-(k+1)\beta) \max_{i \in \overline{\mathcal{P}}_k} \{\mathbb{E}_i\left[R(u \vee C_i)\right]\}\\
&\qquad -\beta \max_{i \in \overline{\mathcal{P}}_{k+1}} \{\mathbb{E}_i\left[R(u \vee C_i)\right]\},\\
&\Rightarrow T_{\text{tr}}(\mathcal{P}_k, u) - \beta R(u)\\
&\quad< \quad T_{\text{pr,tr}}(\mathcal{P}_k, u) - \beta \max_{i \in \overline{\mathcal{P}}_{k+1}} \{\mathbb{E}_i\left[R(u \vee C_i)\right]\},\\
&\Rightarrow T_{\text{tr}}(\mathcal{P}_k, u) - T_{\text{pr,tr}}(\mathcal{P}_k, u)\\
&\quad< \quad \beta R(u) - \beta \max_{i \in \overline{\mathcal{P}}_{k+1}} \{\mathbb{E}_i\left[R(u \vee C_i)\right]\}\\
&\Rightarrow 0 < \left[R(u) - \max_{i \in \overline{\mathcal{P}}_{k+1}} \{\mathbb{E}_i\left[R(u \vee C_i)\right]\}\right].
\end{aligned}$$

Note that the last relation above provides the required contradiction as $u \leq (u \vee C_i)$. ■

Next we prove Theorem 1.

### A. Proof of Theorem 1

Fix arbitrary $\mathcal{P}_{N-1} \supset \mathcal{P}_k$, and let us assume that the users in $\mathcal{P}_{N-1}$ are probed. Then, the resulting system state is $(\mathcal{P}_{N-1}, \vee_{i \in \mathcal{P}_{N-1}} c_i)$. Note that $\vee_{i \in \mathcal{P}_{N-1}} c_i \geq u$ as $u = \vee_{i \in \mathcal{P}_k} c_i$ and $\mathcal{P}_k \subset \mathcal{P}_{N-1}$. Thus, by Lemma 1, $\vee_{i \in \mathcal{P}_{N-1}} c_i \in \mathcal{D}_k$, and by Lemma 2, $\vee_{i \in \mathcal{P}_{N-1}} c_i \in \mathcal{D}_{N-1}$. Thus, by (4),

$$\begin{aligned}
&T_{\text{tr}}(\mathcal{P}_{N-1}, \vee_{i \in \mathcal{P}_{N-1}} c_i)\\
&\quad\geq \quad T_{\text{pr,tr}}(\mathcal{P}_{N-1}, \vee_{i \in \mathcal{P}_{N-1}} c_i)\\
&\quad= \quad \max_{i \in \overline{\mathcal{P}}_{N-1}} \left\{\mathbb{E}[T_{\text{tr}}(\mathcal{P}_{N-1} \cup \{i\}, \vee_{i \in \mathcal{P}_{N-1}} c_i \vee C_i)]\right\}\\
&\quad= \quad \max_{i \in \overline{\mathcal{P}}_{N-1}} \left\{\mathbb{E}[T^\star(\mathcal{P}_{N-1} \cup \{i\}, \vee_{i \in \mathcal{P}_{N-1}} c_i \vee C_i)]\right\}.
\end{aligned}$$

The last relation follows because after probing the last user, the optimal decision is to transmit as it the only decision. Now, from (2), it follows that

$$T^\star(\mathcal{P}_{N-1}, \vee_{i \in \mathcal{P}_{N-1}} c_i) = T_{\text{tr}}(\mathcal{P}_{N-1}, \vee_{i \in \mathcal{P}_{N-1}} c_i). \quad (6)$$

Note that (6) holds for any $\mathcal{P}_{N-1} \supset \mathcal{P}_k$ and for any values of $c_i$'s for $i \in \mathcal{P}_{N-1} \setminus \mathcal{P}_k$.

Next consider any state $(\mathcal{P}_{N-2}, \vee_{i \in \mathcal{P}_{N-2}} c_i)$ that can appear after probing $N-2$ users starting from $(\mathcal{P}_k, u)$. As argued before, here also we can conclude that $\vee_{i\in\mathcal{P}_{N-2}}c_i \in \mathcal{D}_{N-2}$. Thus,

$$\begin{aligned} &T_{\mathrm{tr}}(\mathcal{P}_{N-2}, \vee_{i\in\mathcal{P}_{N-2}}c_i) \\ &\quad\geq\ T_{\mathrm{pr,tr}}(\mathcal{P}_{N-2}, \vee_{i\in\mathcal{P}_{N-2}}c_i) \\ &\quad=\ \max_{i\in\overline{\mathcal{P}}_{N-2}} \left\{\mathbb{E}[T_{\mathrm{tr}}(\mathcal{P}_{N-2}\cup\{i\}, \vee_{i\in\mathcal{P}_{N-2}}c_i \vee C_i)]\right\} \\ &\quad=\ \max_{i\in\overline{\mathcal{P}}_{N-2}} \left\{\mathbb{E}[T^\star(\mathcal{P}_{N-2}\cup\{i\}, \vee_{j\in\mathcal{P}_{N-2}}c_j \vee C_i)]\right\}. \end{aligned}$$

The last equality follows from (6) as (6) holds for any $\mathcal{P}_{N-1}\supset\mathcal{P}_k$ and for any values of $c_i$'s for $i\in\mathcal{P}_{N-1}\setminus\mathcal{P}_k$. But, with (2), this implies that

$$T^\star(\mathcal{P}_{N-2}, \vee_{i\in\mathcal{P}_{N-2}}c_i) = T_{\mathrm{tr}}(\mathcal{P}_{N-2}, \vee_{i\in\mathcal{P}_{N-2}}c_i),$$

for any $\mathcal{P}_{N-2}\supset\mathcal{P}_k$ and for any values of $c_i$'s for $i\in\mathcal{P}_{N-2}\setminus\mathcal{P}_k$.

Reproducing the above reasoning, we get the result by induction down to $k$. ■

## APPENDIX II
## PROOF OF THEOREM 2

*Proof:* We prove the result by induction on the number of un-probed users. When this number is equal to 1, the result holds since we can only probe this user. Now assume the result holds when the number of un-probed users is strictly smaller than $N-k$. Let us establish the result when the number of un-probed users is exactly equal to $N-k$. We use contradiction. Denote by $(\mathcal{P}_k, u)$ the system state, and let the optimal policy probe receiver $i$ instead of stochastically largest $j$ in $\overline{\mathcal{P}}_k$. For brevity define: $\alpha_i = u_{\max}(\mathcal{P}_k\cup\{i\})$, $\alpha_j = u_{\max}(\mathcal{P}_k\cup\{j\})$, and $\alpha = u_{\max}(\mathcal{P}_k\cup\{i,j\})$. Note that $\alpha_i\geq\alpha_j$.

First, note that if $u\geq\alpha_i$, then after probing $i$ or $j$, the optimal policy $\pi^*$ will transmit by Theorem 1. Thus, it is then optimal to probe $j$. From now on we assume that $u\leq\alpha_i$.

We compare the expected throughputs obtained starting from state $(\mathcal{P}_k, u)$ (a) when first probing $i$ and then $j$, and (b) when first probing $j$ and then $i$.

- In scenario (a), probing $i$ results in a channel state $x_i$. By induction, we know that the next user to probe should be $j$. Then if $x_i\geq\alpha_i$, we should not probe $j$ and transmit. If $x_i<\alpha_i$, we should probe $j$. Denote by $x_j$ the state of channel $j$. If $(u\vee x_i\vee x_j)\geq\alpha$, we should transmit; otherwise we should probe further.
- In scenario (b), we first probe $j$. If $x_j\geq\alpha_j$, we should transmit. Otherwise, we probe $i$. Then if $(u\vee x_i\vee x_j)\geq\alpha$, we transmit; otherwise we probe further.

We just need to compare the expected throughput in scenarios (a) and (b) in cases where we transmit after probing $i$ and/or $j$. This is simply due to the fact that if we have to probe further after $i$ and $j$, the systems (a) and (b) are identical. Denote by $T^{(a)}(u)$ and $T^{(b)}(u)$ the expected throughput in scenarios (a) and (b) when we do not probe more users than $i$ and $j$:

$$\begin{aligned} &T^{(a)}(u) \\ &\quad= a_{k+1}\int_{\alpha_i}^\infty dF_i(x)R(x) \\ &\quad\quad+a_{k+2}\int_0^{\alpha_i} dF_i(x)\int_0^\infty dF_j(y)1_{u\vee x\vee y\geq\alpha}R(u\vee x\vee y), \end{aligned}$$

$$\begin{aligned} &T^{(b)}(u) \\ &\quad= a_{k+1}\int_{\alpha_j}^\infty dF_i(x)R(x\vee u) \\ &\quad\quad+a_{k+2}\int_0^{\alpha_j} dF_j(x)\int_0^\infty dF_i(y)1_{u\vee x\vee y\geq\alpha}R(u\vee x\vee y), \end{aligned}$$

where $a_k=(1-k\beta)$. We want to prove that $G(u)=T^{(b)}(u)-T^{(a)}(u)\geq 0$. We prove this using the following two lemmas.

*Lemma 3:* For all $u\leq\alpha_j$, we have $G(u)=G(\alpha_j)$.

*Proof:* First note that when $u\leq\alpha$, then $T^{(a)}$ and $T^{(b)}$ are independent of $u$, and so is $G(u)$. Now assume that $\alpha\leq u\leq\alpha_j$. The first terms in $T^{(a)}$ and $T^{(b)}$ do not depend on $u$. Furthermore their second terms are respectively equal to:

$$\begin{aligned} &a_{k+2}\int_0^{\alpha_i} dF_i(x)\int_0^\infty dF_j(y)1_{x\vee y\geq\alpha}R(x\vee y) \\ &\quad-\int\int_{\Gamma(\alpha,u)} dF_i(x)dF_j(y)\left(R(x\vee y)-R(u)\right), \end{aligned}$$

and

$$\begin{aligned} &a_{k+2}\int_0^{\alpha_j} dF_j(x)\int_0^\infty dF_i(y)1_{x\vee y\geq\alpha}R(x\vee y) \\ &\quad-\int\int_{\Gamma(\alpha,u)} dF_i(x)dF_j(y)\left(R(x\vee y)-R(u)\right), \end{aligned}$$

where $\Gamma(\alpha,u)=\{(x,y):\alpha\leq x,y\leq u\}$. We deduce that indeed $G(u)$ is independent of $u$ when $u\leq\alpha_j$. Thus, $G(u)=G(\alpha_j)$. ■

*Lemma 4:* For all $u$ such that $\alpha_j\leq u\leq\alpha_i$, $G(u)\geq 0$.

*Proof:* Because of the space constraints, we prove the result in the discrete setting. The proof is using the perturbation approach. Without loss of generality, let $\mathbb{N}$ be the channel state space. Denote by $p_i(l)$ the probability that the channel of user $i$ is in state $l$. Observe that when $F_i=F_j$, the result holds. Now we assume the result is true for $F_j$ and show that increasing stochastically $F_j$ does not change this conclusion. We use $F_j^+$ defined by: for $\epsilon>0$, for a particular $l_0\in\mathbb{N}$ in the support of $F_j$, $p_j^+(l_0)=p_j(l_0)-\epsilon$, $p_j^+(l_0+1)=p_j(l_0+1)+\epsilon$ and for all $l\neq l_0, l_0+1$, $p_j^+(l)=p_j(l)$. If $C_j^+\sim F_j^+$, then $C_j\leq_{st}C_j^+$. $\epsilon$ is meant to be chosen as small as we wish. Note that using this kind of perturbations, we can start from $F_i$ and modify it to obtain $F_j$ (it can be proved by coupling arguments). Now it can be shown that the function $G^+(u)$ obtained with $F_j^+$ instead of $F_j$ is such that:

$$\begin{aligned} &G^+(u) \\ &\quad\geq\ G(u)+o(\epsilon) \\ &\quad\quad+\epsilon\times\mathbf{1}_{\{l_0\geq u\}}\times(R(l_0+1)-R(l_0)) \\ &\quad\quad\times(a_{k+1}-a_{k+2}F_i(\alpha_i\vee l_0))+o(\epsilon), \end{aligned} \quad (7)$$

where $\mathbf{1}_A$ is the indicator of event $A$. Note that the difference between $G(u)$ and $G^+(u)$ may come from the variation of $F_j$, which may imply a modification of $\alpha_i$. The latter modification holds only in the very specific cases where $a_{k+2}\mathbb{E}[R(\alpha_i\vee C_j)]=a_{k+1}R(\alpha_i)$, which simplifies the analysis.

From (7), we conclude that $G^+(u)\geq 0$. ■

Note that Lemma 4 shows that $G(u)\geq 0$ for every $u\in[\alpha_j,\alpha_i]$. Now, Lemma 3 shows that $G(u)=G(\alpha_j)$ for every $u\in[0,\alpha_j]$. Thus, $G(u)\geq 0$ for every $u\in[0,\alpha_i]$. This proves Theorem 2. ■

## APPENDIX III
## PROOFS FOR THE RESULTS IN SECTION III-B

### A. Proof of Theorem 3

*Proof:* First, note that as we probe more and more users the maximum rate at which one can transmit increases monotonically, while the maximum rate at which one can guess and transmit decreases monotonically. To see this, consider two system state $(\mathcal{P}_k, u)$ and $(\mathcal{P}_{k+1}, u')$, where $\mathcal{P}_k \subset \mathcal{P}_{k+1}$. Then, clearly $u \leq u'$. Thus, the rate at which one can transmit in state $(\mathcal{P}_k, u)$ (equals $R(u)$) is less than or equal to that in state $(\mathcal{P}_{k+1}, u')$ (equals $R(u')$). Now, the maximum rate at which one can guess and transmits in state $(\mathcal{P}_k, u)$ (equals $\max_{i \in \overline{\mathcal{P}}_k}\{R_{\mathrm{g}(i)}\}$) is less than or equal to that in state $(\mathcal{P}_{k+1}, u')$ (equals $\max_{i \in \overline{\mathcal{P}}_{k+1}}\{R_{\mathrm{g}(i)}\}$). Thus, from a state $(\mathcal{P}_k, u)$ satisfying $T_{\mathrm{tr}}(\mathcal{P}_k, u) \geq T_{\mathrm{g}}(\mathcal{P}_k, u)$, every subsequent state that can be reached by probing a new receiver (say $(\mathcal{P}_{k+1}, u')$) also satisfies $T_{\mathrm{tr}}(\mathcal{P}_{k+1}, u') \geq T_{\mathrm{g}}(\mathcal{P}_{k+1}, u')$. Thus, from any such state (1) reduces to (2). Since the decision process depends only on the current state and the future evolution, the optimal decision process from state $(\mathcal{P}_k, u)$ is exactly the same as that when guessing is not allowed. ■

### B. Proof of Theorem 4

*Proof:* Statement 1) of the theorem follows from (1) and the fact that

$$T_{\mathrm{pr,tr}}(\mathcal{P}_k, u) \leq \max_{i \in \overline{\mathcal{P}}_k} \left\{ \mathbb{E}_i \left[ T^\star(\mathcal{P}_k \cup \{i\}, u \vee C_i) \right] \right\}.$$

Now, we prove the statement 2-a) of the theorem. Let us assume that the system state is $(\mathcal{P}_k, u_{\mathrm{g}(\mathcal{P}_k)})$. Note that by definition of $u_{\mathrm{g}(\mathcal{P}_k)}$, $T_{\mathrm{tr}}(\mathcal{P}_k, u_{\mathrm{g}(\mathcal{P}_k)}) \geq T_{\mathrm{g}}(\mathcal{P}_k, u_{\mathrm{g}(\mathcal{P}_k)})$. Thus, Theorem 3 applies. Since, by assumption, $T_{\mathrm{pr,tr}}(\mathcal{P}_k, u_{\mathrm{g}(\mathcal{P}_k)}) \leq T_{\mathrm{tr}}(\mathcal{P}_k, u_{\mathrm{g}(\mathcal{P}_k)})$, the optimal decision in this state is to transmit to user with channel state $u_{\mathrm{g}(\mathcal{P}_k)}$. But, in actual system, the system state is $(\mathcal{P}_k, u)$ with $u < u_{\mathrm{g}(\mathcal{P}_k)}$ as $T_{\mathrm{tr}}(\mathcal{P}_k, u) < T_{\mathrm{g}}(\mathcal{P}_k, u)$. Now, note that $T^*(\mathcal{P}_k, u)$ is a monotonically increasing function of $u$. Thus, $T^*(\mathcal{P}_k, u) \leq T^*(\mathcal{P}_k, u_{\mathrm{g}(\mathcal{P}_k)}) = T_{\mathrm{tr}}(\mathcal{P}_k, u_{\mathrm{g}(\mathcal{P}_k)}) = T_{\mathrm{g}}(\mathcal{P}_k, u)$. But, $T_{\mathrm{g}}(\mathcal{P}_k, u) \leq T^*(\mathcal{P}_k, u)$ by (1). This concludes the proof.

Finally, we prove the statement 2-b) of the theorem. Proof is by contradiction. Assume that $\pi^*_{\mathrm{NG}}$ retires in state $(\mathcal{P}_k, u)$. Thus, by (1) and since $T_{\mathrm{tr}}(\mathcal{P}_k, u) < T_{\mathrm{g}}(\mathcal{P}_k, u)$, $T^*(\mathcal{P}_k, u) = T_{\mathrm{g}}(\mathcal{P}_k, u) = T_{\mathrm{tr}}(\mathcal{P}_k, u_{\mathrm{g}(\mathcal{P}_k)})$. Now, consider another strategy $\pi$ which, in state $(\mathcal{P}_k, u)$ probes user $j$ such that $T_{\mathrm{pr}(j),\mathrm{tr}}(\mathcal{P}_k, u_{\mathrm{g}(\mathcal{P}_k \cup \{j\})}) > T_{\mathrm{tr}}(\mathcal{P}_k, u_{\mathrm{g}(\mathcal{P}_k)})$, and then transmits to the user with the best channel if $(u \vee c_j) \geq u_{\mathrm{g}(\mathcal{P}_k \cup \{j\})}$ or guesses and transmit to receiver $j_1$ such that $R_{\mathrm{g}(j_1)} = R(u_{\mathrm{g}(\mathcal{P}_k \cup \{j\})})$. Note that $T^\pi(\mathcal{P}_k, u) = T_{\mathrm{pr}(j),\mathrm{tr}}(\mathcal{P}_k, u_{\mathrm{g}(\mathcal{P}_k \cup \{j\})}) > T^*(\mathcal{P}_k, u)$. This concludes the proof. ■

## REFERENCES


[1] A. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1986–1992, 1997.

[2] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. on Inform. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[3] R. Knopp and P. Humblet, "Information capacity and power control in single cell multiuser communications," in *IEEE International Conf. on Communications (ICC)*, Seattle, WA, Jun. 1995, pp. 331–335.

[4] L. Tassiulas and A. Ephremides, "Dynamic server allocation to parallel queues with randomly varying connectivity," *IEEE Trans. Infor. Theory*, vol. 39, no. 2, pp. 466–478, Mar 1993.

[5] M. J. Neely, E. Modiano, and C. E. Rohrs, "Dynamic power allocation and routing for time varying wireless networks," *IEEE Journal on Selected Areas in Communications, Special Issue on Wireless Ad-Hoc Networks*, vol. 23, no. 1, pp. 89–103, Jan 2005.

[6] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar, "Providing quality of service over a shared wireless link," *IEEE Communications Magazine*, vol. 39, no. 2, pp. 150–154, Feb 2001.

[7] M. Neely and E. Modiano, "Capacity and delay tradeoffs for ad-hoc mobile networks," *IEEE Trans. on Inform. Theory*, vol. 51, no. 6, pp. 1917–1937, Jun 2005.

[8] A. Stolyar and K. Ramanan, "Largest weighted delay first scheduling: Large deviations andoptimality," *Annals of Applied Probability*, vol. 11, no. 1, pp. 1–48, 2001.

[9] S. Shakkottai and A. Stolyar, "Scheduling for multiple flows sharing a time-varying channel: The exponential rule," *American Mathematical Society Translations, Series 2, A volume in memory of F. Karpelevich*, 2002.

[10] P. Bender, . Black, M. Grob, P. R, N. Sindhushyana, and S. Viterbi, "Cdma/hdr: a bandwidth efficient high speed wireless data servicefor nomadic users," *IEEE Communications Magazine*, vol. 38, no. 7, pp. 70–77, Jul 2000.

[11] S. Borst and P. Whiting, "Dynamic rate control algorithms for hdr throughput optimization," in *IEEE INFOCOM*, Anchorage, Alaska, Apr 2001.

[12] X. Liu, E. Chong, and N. Shroff, "A framework for opportunistic scheduling in wireless networks," *Computer Networks Journal (Elsevier)*, vol. 41, no. 4, pp. 451–474, 2003.

[13] A. Stolyar, "Greedy primal-dual algorithm for dynamic resource allocation in complex networks," *Queueing Systems*, vol. 54, no. 3, pp. 203–220, 2006.

[14] Z. Ji, Y. Yang, J. Zhou, M. Takai, and R. Bagrodia, "Exploiting medium access diversity in rate adaptive wireless lans," in *the 10th annual international conference on Mobile computing and networking (MOBICOM'04)*, Philadelphia, PA, 2004, pp. 345–359.

[15] A. Sabhrawal, A. Khoshnevis, and E. Knightly, "Opportunistic spectral usage: bounds and a multi-band csma/ca protocol," *IEEE/ACM Trans. on Networking (TON)*, vol. 5, no. 3, pp. 533–545, Jun 2007.

[16] S. Guha, K. Mungala, and S. Sarkar, "Jointly optimal transmission and probing strategies for multichannel wireless systems," in *Conference on Information Sciences and Systems (CISS)*, Mar 2006, (Invited Paper).

[17] ——, "Optimizing transmission rate in wireless channels using adaptive probes," in *ACM SIGMETRICS Performance Evaluation Review*, New York, NY, 2006, pp. 381–382, (Poster Paper).

[18] N. Chang and M. Liu, "Optimal channel probing and transmission scheduling for opportunistic spectrum access," in *the 13th annual ACM international conference on Mobile computing and networking (MOBICOM'07)*, Montral, Qubec, Canada, 2007, pp. 27–38.

[19] D. Bertsekas, *Dynamic Programming and Optimal Control*, 3rd ed. Athena Scientific, 2007.

[20] S. Guha, K. Mungala, and S. Sarkar, "Performance guarantees through partial information based control in multichannel wireless networks," Univ. of Pennsylvania, Philadelphia, PA, Tech. Rep., 2007.

[21] H. Robbins, "Some aspects of the sequential design of experiments," *Bull. Amer. Math. Soc.*, 1952.

[22] P. Whittle, "Restless bandits: Activity allocation in a changing world," *Journal of Applied Probability*, vol. 25, pp. 287–298, 1988, special Issue: A Celebration of Applied Probability.

[23] S. Babu, R. Motwani, K. Mungala, E. Nishizhawa, and J. Widom, "Adaptive ordering of pipelined stream filters," in *the 2004 ACM SIGMOD international conference on Management of data*, Paris, France, 2004, pp. 407–418.

[24] M. Kodialam, "The throughput of sequential testing," in *the 8th International IPCO Conference on Integer Programming and Combinatorial Optimization*. Springer-Verlag London, UK, 2001, pp. 280–292, lecture Notes In Computer Science, Vol. 2081.

[25] Y.-S. Chow, H. Robbins, and D. Siegmund, *Great Expectations: Theory of Optimal Stopping*. Houghton Mifflin Co, Jun 1972.

[26] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Trans. on Automatic Control*, vol. 37, no. 12, pp. 1936–1948, Dec 1992.

**Prasanna Chaporkar** received M.S. in Faculty of Engineering from Indian Institute of Science, Bangalore, India in 2000, and Ph.D. from University of Pennsylvania, Philadelphia, PA in 2006. He was a ERCIM post-doctoral fellow at ENS, Paris, France and NTNU, Trondheim, Norway. Currently, he is an Assistant Professor at Indian Institute of Technology, Mumbai, India. His research interests are in resource allocation in communication networks, stochastic control and queueing theory.

**Alexandre Proutiere** is a researcher in the Systems and Networking group at Microsoft Research, Cambridge, Great Britain. His research interests are in the design and the performance evaluation of computer networks, with a specific interest in resource allocation and control in wireless systems. Before joining Microsoft in June 2007, he was a senior expert researcher at France Telecom R&D and an assistant professor in the computer science department of Ecole Normale Superieure (Paris). He received a PhD in mathematics from Ecole Polytechnique (Palaiseau, France) in 2003, graduated in mathematics from Ecole Normale Superieure (Paris), and qualified as an engineer at Ecole Nationale Superieure des Telecommunications (Paris). He, with Thomas Bonald, is the recipient of the best paper award of ACM Sigmetrics /Performance 2004.