Joint-optimal Probing and Scheduling
in Wireless Systems

Prasanna Chaporkar and Alexandre Proutiere

Abstract— Consider a wireless system where a sender can transmit data to various users with independent and varying channel conditions. To maximize its long-term transmission rate, the sender should always transmit to the user with the best channel. To discover which user has the best channel, it has to spend time to probe channels, and this reduces the time available for effective transmission. This paper aims at identifying optimal joint probing and scheduling strategies. These strategies realize the best trade-off between the channel state acquisition and effective transmission. We first provide general structural properties of optimal strategies, and then exactly characterize these strategies in particular but relevant cases. Finally we propose extensions of this problem, e.g., to impose fairness among the users, we investigate how to maximize system utility rather than throughput.

I. INTRODUCTION

A. Probing and Transmission Problems in Wireless Systems

In rate adaptive wireless systems, the sender can adapt the transmission rate to the receiver depending on the channel conditions between them. When the sender can transmit either to different users or to a single receiver but on various channels, it has to decide to which user it should transmit or which channel it should use. Before transmitting, it can spend some time to discover the conditions of the various users or channels. To maximize its transmission throughput, the sender should adopt an efficient probing/scheduling strategy. Probing many users/channels reduces the time remaining for effective transmission but can increase the rate during transmission, whereas probing very few users/channels reduces the probing procedure duration, but decreases the information available to the sender and thereby reduces the rate of transmission. We illustrate this probing/scheduling problem in two examples of wireless systems.

a) Problem 1. A broadcast channel with CSI acquisition: Consider a wireless network with a single sender and several receivers. The sender can transmit to one receiver at a time, and hence, has to decide which receiver to schedule at each time slot. The objective of the sender could be to maximize the system long-term throughput or when accounting for fairness issues, a certain *utility* of the system. Unlike 3G cellular systems where a dedicated channel per receiver is open to convey the channel state information (CSI), we consider a scenario where the CSI and the data are transmitted on the same channel. At the beginning of each time slot, the sender broadcasts a pilot signal used by each receiver i to estimate its channel state c_i. The sender may decide to acquire the channel state of receiver i, and to this aim, it sends a query to this receiver. For each receiver, this acquisition procedure takes a proportion β of the slot. The sender should then carefully design its probing and scheduling strategy.

b) Problem 2. Opportunistic spectrum access in multi-channel wireless networks: In future wireless systems, a user will be able to access a large number of channels. To maximize its throughput, a user should use a channel with good conditions. Here again, a user has to probe channels to look for the best possible and this has a cost, as it reduces the remaining time for effective transmission. The problem is clearly identical to problem 1.

B. System Model and Problem Formulation

For illustrative purposes, we choose to use a terminology related to Problem 1 (even though the formalism for Problem 2 is identical). We consider a system of N users whose channel conditions vary over time. Time is slotted and the channel conditions of the various users are assumed to remain constant for the duration of one slot, i.e., the coherence time of the channels is larger than one slot; these conditions may change at the slot boundaries. In other words, we consider the block fading model. 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 is an increasing function. For example, R can represent Shannon limit: $R(c) = W \log_2 \left(1 + \frac{Pc}{N_0}\right)$, where W is the channel bandwidth and P, N_0 are the transmission and the noise powers, respectively. Here the channel state c represents the fraction of transmission power received by the user considered. We assume that the channel states are independent across users, but the distributions of the channel state of each user may be different. For a given user, the channel states are independent and identically distributed (i.i.d.) across time slots with c.d.f. $F_i(\cdot)$ ($F_i(a) = Pr[c_i(t) \le a]$). In the following, we denote by C_i a generic r.v. with c.d.f. F_i. We assume that F_i, for every user i, is known at the transmitter.

At the beginning of each time slot, the sender can decide to probe some channels, to transmit to one of the probed users, or to transmit to a user that has not been probed[1]. We assume

[1]Note here that transmitting to an un-probed user 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.

that probing the channel state of a user takes a proportion β of the slot duration. Hence, in a given slot, when the transmitter decides to transmit to a user 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 users in that slot, and $|\mathcal{P}|$ is the cardinality of this set. We denote by $\overline{\mathcal{P}}$ the set $\{1, \ldots, N\} \setminus \mathcal{P}$.

The problem is to design a joint probing and transmission strategy that maximizes the long-term average system throughput[2]. Such a strategy is called optimal. Since the system is i.i.d. across time slots, maximizing the long-term system throughput is equivalent to maximizing the average throughput in each slot. For notational simplicity, we consider any given slot and drop time t from the notation.

Assume that the channels of users in set $\mathcal{P}$ has been already probed, and that the state of the best of the probed channels is u (we say that the system is in state $(\mathcal{P}, u)$). Then a strategy π decides one of the following possible actions:

(i) transmit to the user with the best probed channel,
(ii) transmit to a user that has not been probed,
(iii) probe one more user from the set $\overline{\mathcal{P}}$.

In cases (i) and (ii), we say that we *retire*. In cases (ii) and (iii), the strategy has also to define to which user to transmit and which user to probe, respectively. We denote by $T^\pi(\mathcal{P}, u)$ the average throughput achieved by strategy π, starting from system state $(\mathcal{P}, u)$. Also denote by $T^\star(\mathcal{P}, u)$ the average throughput of an optimal strategy starting from system state $(\mathcal{P}, u)$. We are looking for an optimal strategy $\pi^\star$ in the sense that $T^{\pi^\star}(\emptyset, 0) = T^\star(\emptyset, 0)$.

Now, using examples, we demonstrate the challenges involved in designing the optimal policy π^*.

Example 1: Consider a broadcast channel with two receivers experiencing i.i.d. fading. 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 β denote the fraction of slot duration required to probe a receiver and acquire its CSI. In these settings, we compare two probing and transmission strategies π_1 and π_2. Under π_1, the transmitter probes both the receivers, and then transmits to one with a higher rate. Ties are broken arbitrarily. Under π_2, the transmitter probes one receiver at random, and transmits to it at the maximum rate possible. Thus, policy π_1 spends 2β units of time per slot to acquire CSI, and transmits at the expected rate of 7/4 in the remainder of the slot; while π_2 spends only β units of time per slot to acquire CSI, but transmits at a smaller expected rate of 3/2 in the remainder of the slot. Thus, the expected rate of transmission (throughput) under π_1 is $7(1 - 2\beta)/4$, while that under π_2 is $3(1 - \beta)/2$. Note that if $\beta \leq 1/8$, then π_1 has a higher throughput than

[2]As explained in Section V, we could also aim at maximizing system *utility*, where the choice of a utility function depends our fairness objective.

that under π_2. But, when $\beta > 1/8$, the throughput under π_2 is higher than that under π_1. ∎

The above example demonstrates that the probing and transmission strategy should be designed by taking into account the cost for probing, which is time here. Since, the channel state changes after coherence time, the CSI obtained can only be used in the same coherence time duration. 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 us to discover receivers with high channel gains and thereby achieve high throughput. Moreover, in Example 1, policy π_1 can be easily modified to provide better throughput in the following way. If at the first probe, π_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. Note that with this modification, π_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 challenge in designing an optimal probing and transmission strategy is that of deciding the sequence in which receivers should be probed. In Example 1, we have consider 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 independent across receivers, but do not have the same distribution, 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 these settings, 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 optimal 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 π^* 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 π_1 that probes R_1 and transmits at appropriate rate, and with policy π_2 that probes R_1 first. If the achievable rate is k, then π_2 transmits to R_1, otherwise probes R_2 and transmits to it if the achievable rate is $2k$, else transmits to R_1 at rate 2. Note that the throughput of π_1 is $(1 - \beta)\left[\frac{2(k-1)}{k} + 1\right]$, while that of π_2 is $(1 - \beta) + (1 - 2\beta)\frac{3k^2-4k+1}{k^2}$. It is easy to verify

that throughput of π_2 is always smaller than that of $\pi^\star$, while the throughput of π_1 is smaller than that of $\pi^\star$ when $\beta < \frac{2k^2}{8k^2-7k+2}$. Thus, sampling 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 sequence 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.

C. Related Work and Our Contributions

The problem of identifying optimal joint probing and transmission strategies has been addressed in the literature only recently [9], [11], [5], [6], [3]. It falls into the broad class of stochastic control problems [2]. However, as explained in [7], 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 [10], [12], acquiring the state of an arm (or of a channel here) before using it is not allowed. Optimal sampling order of random variables have been investigated in many contexts, see e.g. [1], [8]; 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 [4], 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 [11], 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 [2].

In this paper, we provide the detailed analysis of the stochastic control problem: we first give general structural properties of the optimal strategy. We then exactly characterize this strategy in specific but relevant cases where the distributions of the various channels are *ordered* (in a sense that will be defined later). We illustrate our findings with numerical results and finally propose several interesting generalizations of the model; for example, to introduce some fairness constraints among the various receivers, we propose a strategy maximizing the system utility rather than its total throughput. The detailed analysis of the latter strategy is left for future work.

II. 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 $(\mathcal{P}, u)$, 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 probe further. The latter question is a much more challenging issue than deriving the basic structural properties of an optimal strategy, and it will be addressed in the next section.

A. Results

Assume that the system state is $(\mathcal{P}_k, u)$, where the subscript k indicates that $|\mathcal{P}_k| = k$ (k users have been probed already).

(i) If we decide to transmit to a probed user, we should transmit to the user with the best channel state, i.e., u. In that case, the throughput would be: $T_{\mathrm{tr}}(\mathcal{P}_k, u) = (1 - k\beta)R(u)$.

(ii) If we decide to transmit to an un-probed user $i \in \overline{\mathcal{P}}_k$, then the throughput would be[3]: $T_{\mathrm{g}(i)}(\mathcal{P}_k, u) = (1 - k\beta)\mathbb{E}[R(C_i)]$. Of course, it is optimal to transmit to the user with the highest expected channel state, in which case the throughput becomes: $T_{\mathrm{g}}(\mathcal{P}_k, u) = (1 - k\beta)\max_{i\in\overline{\mathcal{P}}_k}\mathbb{E}[R(C_i)]$.

(iii) Finally, if we decide to probe another user $i \in \overline{\mathcal{P}}_k$, then the optimal expected throughput we can achieve, given that we make this choice and that the channel state of the newly probed channel is c_i, would be[4]: $T^\star(\mathcal{P}_k \cup \{i\}, u \vee c_i)$.

Now it can be easily seen that the maximal expected throughput $T^\star(\mathcal{P}_k, u)$ satisfies the following recursion equation (for all k, $\mathcal{P}_k$ and u):

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

where $\mathbb{E}_i[\cdot]$ is the expectation taken with respect to F_i. Thus, in each state $(\mathcal{P}_k, u)$, $\pi^\star$ chooses the control decision corresponding to the term that achieves the maximum in (1), e.g., if $T_{\mathrm{tr}}(\mathcal{P}_k, u)$ achieves the maximum then the optimal decision is to transmit.

Note that when the 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.

Define: $T_{\mathrm{pr}(i),\mathrm{tr}}(\mathcal{P}_k, u) = (1 - (k+1)\beta)\mathbb{E}_i[R(u \vee C_i)]$, for $i \in \overline{\mathcal{P}}_k$, and $T_{\mathrm{pr},\mathrm{tr}}(\mathcal{P}_k, u) = (1 - (k+1)\beta)\max_{i\in\overline{\mathcal{P}}_k}\mathbb{E}_i[R(u \vee C_i)]$. $T_{\mathrm{pr}(i),\mathrm{tr}}(\mathcal{P}_k, u)$ 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 in a large number of states, it suffices to consider $T_{\mathrm{pr},\mathrm{tr}}(\mathcal{P}_k, u)$, rather than $T^\star(\mathcal{P}_k, u)$ in (1). Since unlike $T^\star(\mathcal{P}_k, u)$, $T_{\mathrm{pr},\mathrm{tr}}(\mathcal{P}_k, u)$ can be computed with complexity $O(N)$, this considerably reduces the complexity of computing an optimal strategy.

---

[3]Subscript g indicates "guess".
[4]We use the notation $a \vee b = \max(a, b)$.

Theorem 1: Let $(\mathcal{P}_k, u)$ be the system state.

(a) Assume that $T_{\mathrm{pr,tr}}(\mathcal{P}_k, u) \geq \max\{T_{\mathrm{tr}}(\mathcal{P}_k, u), T_{\mathrm{g}}(\mathcal{P}_k, u)\}$. The optimal decision is to probe an additional user from $\overline{\mathcal{P}}_k$. Moreover, if $T_{\mathrm{tr}}(\mathcal{P}_k, u) \geq T_{\mathrm{g}}(\mathcal{P}_k, u)$, then after probing, we also have in the new system state, say $(\mathcal{P}_{k+1}, u')$, $T_{\mathrm{tr}}(\mathcal{P}_{k+1}, u') \geq T_{\mathrm{g}}(\mathcal{P}_{k+1}, u')$. In words, in this case, an optimal strategy will never transmit to an un-probed channel.

(b) Assume that $T_{\mathrm{tr}}(\mathcal{P}_k, u) \geq \max\{T_{\mathrm{g}}(\mathcal{P}_k, u), T_{\mathrm{pr,tr}}(\mathcal{P}_k, u)\}$. The optimal decision is to transmit to the user i in $\mathcal{P}_k$ such that $c_i = u$.

(c) Assume that $T_{\mathrm{g}}(\mathcal{P}_k, u) > T_{\mathrm{tr}}(\mathcal{P}_k, u) \geq T_{\mathrm{pr,tr}}(\mathcal{P}_k, u)$. The optimal decision is to transmit to the un-probed user maximizing $T_{\mathrm{g}(i)}(\mathcal{P}_k, u)$.

The above theorem shows that the optimal control decision can be obtained using only the one-step-look-ahead throughput $T_{\mathrm{pr,tr}}(\mathcal{P}_k, u)$ in most of the states $(\mathcal{P}_k, u)$. Indeed, one of the consequences of the theorem is that when guessing is not allowed, i.e., the sender can transmit to one of the probed users only, then the optimal control decision is characterized by using the one-step-look-ahead throughput in all the states. Thus, Theorem 1 plays an important role in reducing the computational complexity of determining the optimal policy.

When guessing is allowed, Theorem 1 does not characterize the optimal control decision is states $(\mathcal{P}_k, u)$ only when $T_{\mathrm{g}}(\mathcal{P}_k, u) > T_{\mathrm{pr,tr}}(\mathcal{P}_k, u) > T_{\mathrm{tr}}(\mathcal{P}_k, u)$. Using the following example, we demonstrate that in these states the optimal decision can not be characterized using one-step-look-ahead throughput alone. Specifically, we show that to guess and transmit may not be the optimal decision even when $T_{\mathrm{g}}(\mathcal{P}_k, u) > \max\{T_{\mathrm{pr,tr}}(\mathcal{P}_k, u), T_{\mathrm{tr}}(\mathcal{P}_k, u)\}$.

Example 3: Let us consider a broadcast channel with three users. For simplicity, let the system state $(\mathcal{P}, u)$ denote the set of probed users $\mathcal{P}$ and the rate u to the probed user with the best channel. Note that here u denotes the transmission rate and not the channel gain as assumed before. This change can be made without loss of generality as $R(\cdot)$ is monotone increasing. Let the system state be $(\{1\}, 0)$ and let the distributions for the remaining two users be as follows: $C_2 = 2k$ w.p. $1/k$ and 0 otherwise; $C_3 = k$ w.p. $1/k$ and 0 otherwise, for some constant $k > 1$. Now, note that $T_{\mathrm{tr}}(\{1\}, 0) = 0$, $T_{\mathrm{g}}(\{1\}, 0) = 2(1 - \beta)$ (achieved by transmitting to user 2), and $T_{\mathrm{pr,tr}}(\{1\}, 0) = 2(1 - 2\beta)$ (achieved by probing user 2). Now, we show that guessing and transmitting to user 2 is not the optimal decision in the state $(\{1\}, 0)$. To see this consider a policy π_1 that probes user 3, and if the system state becomes $(\{1, 3\}, k)$ then transmits to user 3, else guesses and transmits to user 2 in system state $(\{1, 3\}, 0)$. Note that the expected throughput of this policy is $(1 - 2\beta)\left[1 + 2\left(1 - \frac{1}{k}\right)\right]$. Thus, when $\frac{k}{2} > \frac{1 - 2\beta}{1 - 4\beta}$, π_1 achieves higher expected throughput than the policy which guesses and transmits to user 2 in state $(\{1\}, 0)$. Specifically, when $\beta = 1/20$, k has to be greater than 3 for π_1 to achieve the higher expected throughput.

Next, we present the proof of Theorem 1.

B. Proof of Theorem 1

To prove Theorem 1, we use the following two lemmas, whose proofs are given in Appendices I and II.

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

$$\mathcal{D}_k = \{u : T_{\mathrm{tr}}(\mathcal{P}_k, u) \geq T_{\mathrm{pr,tr}}(\mathcal{P}_k, u)\}. \qquad (2)$$

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

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 $u_{\max}(\mathcal{P}_k) \geq u_{\max}(\mathcal{P}_{k+1})$.

Now, let us prove Theorem 1. The first statement in (a) holds since:

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

and in view of (1). For the second statement of (a): without loss of generality, let j denote the new probed user, and c_j the observed channel. Note that by assumption, we have:

$$R(u) \geq \max_{i \in \overline{\mathcal{P}}_k} \mathbb{E}[R(C_i)],$$

which implies that:

$$R(u \vee c_j) \geq \max_{i \in \overline{\mathcal{P}}_k} \mathbb{E}[R(C_i)] \geq \max_{i \in \overline{\mathcal{P}}_k, i \neq j} \mathbb{E}[R(C_i)].$$

The last inequality is equivalent to the desired result.

Now to prove (b) and (c), we assume that $T_{\mathrm{tr}}(\mathcal{P}_k, u) \geq T_{\mathrm{pr,tr}}(\mathcal{P}_k, u)$, and show that the optimal decision is to retire. More precisely, if $T_{\mathrm{g}}(\mathcal{P}_k, u) > T_{\mathrm{tr}}(\mathcal{P}_k, u)$, the optimal decision is to transmit to the best un-probed user, otherwise the optimal decision is to transmit to the best probed user.

An important remark is that starting from state $(\mathcal{P}_k, u)$, if the optimal strategy does not immediately transmit to an un-probed user, then it won't transmit to an un-probed user later (since when one probes, the set of un-probed users and the time remaining for transmission reduce). As a consequence, when computing for $T^\star(\mathcal{P}_k \cup \mathcal{P}, u')$ for any non-empty set $\mathcal{P}$ and any $u' \geq u$, we do not need to account for any of the terms T_{g} in the recursion (1).

Fix any 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 (2),

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

The last relation follows because after probing the last user, the optimal decision is to transmit as it is the only decision. Now, from (1), it follows that if $N - 1 > k$:

$$T^\star(\mathcal{P}_{N-1}, \vee_{i \in \mathcal{P}_{N-1}} c_i) = T_{\mathrm{tr}}(\mathcal{P}_{N-1}, \vee_{i \in \mathcal{P}_{N-1}} c_i). \qquad (3)$$

Note that (3) holds for any $\mathcal{P}_{N-1} \supset \mathcal{P}_k$ and for any values of the 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,

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

Then if $N-2 > k$, $T_{\mathrm{tr}}(\mathcal{P}_{N-2}, \vee_{i \in \mathcal{P}_{N-2}} c_i) = T^\star(\mathcal{P}_{N-2}, \vee_{i \in \mathcal{P}_{N-2}} c_i)$. By induction down to $k+1$, we show that $T_{\mathrm{tr}}(\mathcal{P}_{k+1}, \vee_{i \in \mathcal{P}_{k+1}} c_i) = T^\star(\mathcal{P}_{k+1}, \vee_{i \in \mathcal{P}_{k+1}} c_i)$, and finally deduce that:

$$
T^\star(\mathcal{P}_k, u) = \max\{T_{\mathrm{tr}}(\mathcal{P}_k, u), T_{\mathrm{g}}(\mathcal{P}_k, u), T_{\mathrm{pr,tr}}(\mathcal{P}_k, u)\}.
$$

This completes the proof. ∎

III. Optimal Probing Sequence

Theorem 1 provides useful guidelines to decide when to transmit to a probed or un-probed user, or when to probe a new user. This decision depends on the one-step-look-ahead expected throughput. Nevertheless, it is in general very difficult to know which user to probe next when probing is optimal. We provide an answer to this issue in practically relevant cases.

A. Stochastically Ordered Channels

We first introduce the notion of stochastically ordered channels.

Definition 1: The channels of the N users are stochastically ordered if there exists a permutation σ of $\{1, \ldots, N\}$ such that for all $\sigma(i) \leq \sigma(j)$, $C_{\sigma(j)} \leq_{st} C_{\sigma(i)}$, where $\leq_{st}$ denotes the usual strong stochastic ordering[5].

Without loss of generality, when the channels are stochastically ordered, we assume that the permutation σ is $\sigma(i) = i$ for all i. Note also 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 $C_i = \mathbb{E}[C_i] Y_i$ where the random variables Y_i's are i.i.d. copies of a positive random variable Y, i.e., when the channels have similar distributions but different means. This is a quite usual fading model in wireless networks.

B. Optimal Probing Strategy

When the channels are stochastically ordered, i.e. $C_1 \geq_{st} \ldots \geq_{st} C_N$, we can prove that we should always probe the stochastically largest un-probed channel. So we should probe the users in increasing order. The result is formalized in the following theorem.

---

[5] $X \leq_{st} Y$ if and only if for all increasing function f such that $\mathbb{E}[f(Y)] < +\infty$, $\mathbb{E}[f(X)] \leq \mathbb{E}[f(Y)]$.

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

Before proving this result, note that it provides a full description of the optimal probing and transmission strategy when transmitting to an un-probed user is not possible. We summarize this policy in the following corollary.

Corollary 1: When transmitting to an un-probed channel is not possible and when the channels are stochastically ordered, the one-step-look-ahead strategy is optimal, when the system is in state $(\mathcal{P}_k, u)$:

(i) If $T_{\mathrm{pr,tr}}(\mathcal{P}_k, u) \geq T_{\mathrm{tr}}(\mathcal{P}_k, u)$, then we should probe the stochastically largest un-probed user,

(ii) otherwise, we should transmit to user $i \in \mathcal{P}_k$ such that $c_i = u$.

Note that the stochastic ordering of the channels play an important role in determining the optimal probing sequence. When the channels are not stochastically ordered then probing user that achieves $T_{\mathrm{pr,tr}}(\mathcal{P}_k, u)$ may not be optimal as demonstrated in Example 2.

C. Proof of Theorem 2

The result of Theorem 2 seems quite intuitive. However, as often in stochastic control problems, its proof is far from being straightforward. In the simpler case of linear probing cost, the authors of [5] could not come up with a simple proof of similar results. Here with non-linear cost, it is more complicated.

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 $N-k$. Denote by $(\mathcal{P}_k, u)$ the system state. Since $T_{\mathrm{pr,tr}}(\mathcal{P}_k, u) \geq T_{\mathrm{tr}}(\mathcal{P}_k, u)$, we have $u \leq u_{\max}(\mathcal{P}_k)$. Define for all i, j: $\alpha_i = u_{\max}(\mathcal{P}_k \cup \{i\})$, $\alpha_j = u_{\max}(\mathcal{P}_k \cup \{j\})$, and $\alpha = u_{\max}(\mathcal{P}_k \cup \{i, j\})$. Note that $\alpha_i \geq \alpha_j$. Finally we denote by j the stochastically greatest user in $\overline{\mathcal{P}}_k$, and let $i \in \overline{\mathcal{P}}_k$ with $i \neq j$. If $u \geq \alpha_i$, then after probing i or j, we should transmit. It is then optimal to probe j. From now on we assume that $u \leq \alpha_i$.

We compare the expected throughput 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 throughputs in scenarios (a) and (b) when we do not probe more users than i and j:

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

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

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

Lemma 3: For all $u \leq \alpha_j$, we have $G(u) = G(0)$.

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:

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

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

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

We now state the key lemma to conclude the proof.

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

Proof: We prove the result in the discrete setting[6] using a 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 $C_j \leq_{st} C_j^+$. ϵ 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 easily 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:

$$G^+(u)$$
$$\geq G(u) + o(\epsilon)$$
$$+ \epsilon \times 1_{l_0 \geq u} (R(l_0 + 1) - R(l_0))(a_{k+1} - a_{k+2} F_i(\alpha_i \vee l_0)) \quad (4)$$

[6] Proof for the continuous setting is similar.

Note that the difference between $G(u)$ and $G^+(u)$ may come from the variation of F_j, which may imply a modification of α_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 (4), we conclude that $G^+(u) \geq 0$. ∎

IV. 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, whose computation is based on the results of the previous section; (b) the optimal strategy when guessing is not allowed; (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.

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 β, 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 β 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.

V. EXTENSIONS: IMPOSING FAIRNESS AMONG RECEIVERS

So far, we have proposed a strategy maximizing the long-term system throughput. To impose fairness among users, we may aim at maximizing utility instead. Let U be a concave non-decreasing function, and denote by $t_i^\pi(t)$ the throughput received by user i under strategy π in time slot t. The long-term throughput of user i is then $\phi_i^\pi = \lim_{t=1}^M \frac{1}{M} t_i^\pi(t)$. Now the objective is to maximize $\sum_{i=1}^N U(\phi_i^\pi)$. Denote by $T_i^\pi(t)$ the expected (with respect to the channel state distributions) throughput received by user i. We propose the following
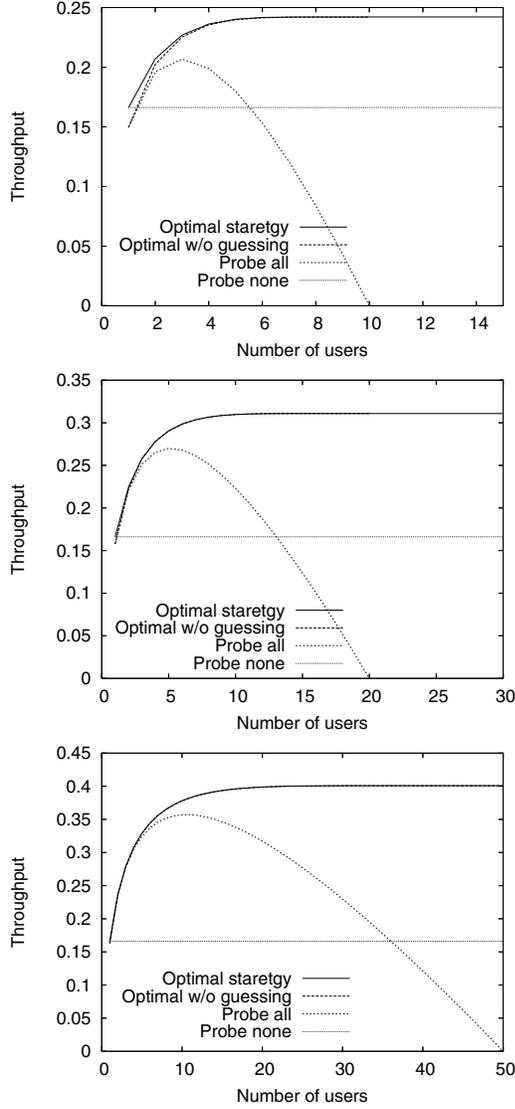
Fig. 1. Average throughput of the various strategies as the number of users N increases - Exponential channels with different means - $\beta = 0.1$ (upper figure), 0.05 (middle figure), 0.02 (lower figure).

gradient algorithm: each time slot t, do

$$\max_{\pi} \sum_{i=1}^{N} T_i^{\pi}(t) \times U'(\bar{t}_i(t)), \qquad (5)$$

$$\bar{t}_i(t+1) = (1-\eta)\bar{t}_i(t) + \eta t_i(t). \qquad (6)$$

Note that solving (5) is equivalent to maximizing a weighted sum of expected throughputs each time slot. This can be done as in Sections II, III. We let the analysis of the proposed algorithm, and in particular its convergence towards the optimum when η tends to 0, for future work. Note that the optimization problem considered is convex (to ensure the convexity, we can make use of strategies with probabilistic decisions).
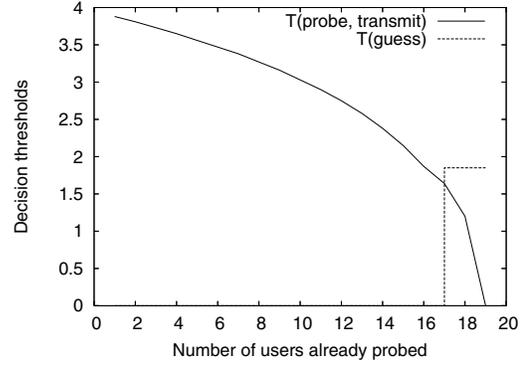

Fig. 2. Decision thresholds as a function of the set of already probed users - $\beta = 0.05$, $N = 20$ users.

APPENDIX I
PROOF OF LEMMA 1

Proof: Let $u \in \mathcal{D}_k$ and consider any $u' > u$. Note that the condition

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

implies that

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

Thus, $(1-k\beta)R(u)[1-F_i(u')]$

$$\geq (1-k\beta)\int_{u'}^\infty R(x)\mathbf{d}F_i(x) - \beta\mathbb{E}_i[R(u\vee C_i)],.$$

From the above inequality we can conclude the following.

$$
\begin{aligned}
(1&-k\beta)R(u')[1-F_i(u')] \\
&\geq (1-k\beta)\int_{u'}^\infty R(x)\mathbf{d}F_i(x) - \beta\mathbb{E}_i[R(u'\vee C_i)]. \;(7)
\end{aligned}
$$

Relation (7) holds for every $i \in \overline{\mathcal{A}}_k$, and the lemma is proved. Note that to obtain (7), we used the fact that $R(\cdot)$ and $F_i(\cdot)$ are monotonically non-decreasing in u. ∎

APPENDIX II
PROOF OF LEMMA 2

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,

$$
\begin{aligned}
(1&-(k+1)\beta)R(u) \\
&< (1-(k+2)\beta)\max_{i\in\overline{\mathcal{P}}_{k+1}}\{\mathbb{E}_i[R(u\vee C_i)]\}, \\
\Rightarrow (1&-k\beta)R(u) - \beta R(u) \\
&< (1-(k+1)\beta)\max_{i\in\overline{\mathcal{P}}_{k+1}}\{\mathbb{E}_i[R(u\vee C_i)]\} \\
&\quad -\beta\max_{i\in\overline{\mathcal{P}}_{k+1}}\{\mathbb{E}_i[R(u\vee C_i)]\},
\end{aligned}
$$

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

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

References

[1] S. Babu, R. Motwani, K. Munagala, I. Nishizawa, J. Widom. Adaptive ordering of pipelined stream filters. In proc. of *ACM Sigmod*, 2004.

[2] D. Bertsekas. Dynamic Programming and Optimal Control, 3rd edition. Athena Scientific, 2007.

[3] N. Chang, M. Liu. Optimal channel probing and transmission scheduling for opportunistic spectrum access. In proc. of *ACM MobiCom*, 2007.

[4] Y.S. Chow, H. Robbins, D. Siegmund. Great expectations: the theory of optimal stopping. *Houghton Mufflin Company*, 1971.

[5] S. Guha, K. Munagala, S. Sarkar. Jointly optimal transmission and probing strategies for multichannel wireless systems. In proc. of *CISS*, 2006.

[6] S. Guha, K. Munagala, S. Sarkar. Optimizing Transmission Rate in Wireless Channels using Adaptive Probes. Poster paper in *ACM Sigmetrics/Performance Conference*, 2006.

[7] S. Guha, K. Munagala, S. Sarkar. Performance guarantees through partial information based control in multi-channel wireless networks. Technical report, UPenn, 2007.

[8] M. Kodialam. The throughput of sequential testing. *Lectures notes in Computer Science*, 2081 pp 280-292, 2001.

[9] Z. Li, Y. Yang, J. Zhou, M. Takai, R. Bagrodia. Exploiting medium access diversity in rate adaptive wireless LANs. In proc. of *ACM Mobicom*, 2004.

[10] H. Robbins. Some aspects of the sequential design of experiments. *Bull. Amer. Math. Soc.*, 55 pp 527-535, 1952.

[11] A. Sabharwal, A. Khoshnevis, E. Knightly. Opportunistic spectral usage: Bounds and multi-band CSMA/CA protocol. *ACM/IEEE Trans. on Networking*, vol 15-3, 2007.

[12] P. Whittle. Restless bandits: Activity allocation in a changing world. In: *A celebration of Applied Probability, J. Gani (Ed), J. Appl. Probab. Spec.*, 25 pp 287-298, 1988.