Optimal Channel Probing and Transmission Scheduling for Opportunistic Spectrum Access

Nicholas B. Chang, *Student Member, IEEE*, and Mingyan Liu, *Member, IEEE*

Abstract—In this study, we consider optimal opportunistic spectrum access (OSA) policies for a transmitter in a multichannel wireless system, where a channel can be in one of multiple states. In such systems, the transmitter typically does not have complete information on the channel states, but can learn by probing individual channels at the expense of certain resources, e.g., energy and time. The main goal is to derive optimal strategies for determining which channels to probe, in what sequence, and which channel to use for transmission. We consider two problems within this context and show that they are equivalent to different data maximization and throughput maximization problems. For both problems, we derive key structural properties of the corresponding optimal strategy. In particular, we show that it has a threshold structure and can be described by an index policy. We further show that the optimal strategy for the first problem can only take one of three structural forms. Using these results, we first present a dynamic program that computes the optimal strategy within a finite number of steps, even when the state space is uncountably infinite. We then present and examine a more efficient, but suboptimal, two-step look-ahead strategy for each problem. These strategies are shown to be optimal for a number of cases of practical interest. We examine their performance via numerical studies.

Index Terms—Channel probing, cognitive radio, dynamic programming, opportunistic spectrum access (OSA), optimal stopping, scheduling, stochastic optimization.

I. INTRODUCTION

EFFECTIVE transmission over wireless channels is a key component of wireless communication. To achieve this, one must address a number of issues specific to the wireless environment. One such challenge is the time-varying nature of the wireless channel due to multipath fading caused by factors such as mobility, interference, and environmental objects. The

resulting unreliability must be accounted for when designing robust transmission strategies. Recent works such as [1] and [2] have studied opportunistic transmission when channel conditions are better to exploit channel fluctuations over time.

At the same time, many wireless systems also provide transmitters with multiple channels to use for transmission. As mentioned in [3], a channel can be thought of as a frequency in a frequency division multiple access (FDMA) network, subcarrier in an orthogonal frequency division multiple access (OFDM) network, a code in a code division multiple access (CDMA) network, or as an antenna or its polarization state in multiple-input–multiple-output (MIMO) systems. In addition, software-defined radio (SDR) [4] and cognitive radio networks [5] may provide users with multiple channels (e.g., tunable frequency bands and modulation techniques) by means of a programmable hardware that is controlled by software. The transmitter, for example, could be a secondary user seeking spectrum opportunities in a network whose N channels have been licensed to a set of primary users [5].

In these systems, the transmitter is generally supplied with more channels than needed for a single transmission. Thus, the transmitter could possibly utilize the time-varying nature of the channels by opportunistically selecting the best one to use for transmission [6], [7]. This may be viewed as an exploitation of spatial channel fluctuations (i.e., across different channels) and is akin to the idea of multiuser diversity [2].

In order to utilize such channel diversity, it is desirable for the transmitter and/or receiver to periodically obtain information on channel quality. One distributed method of accomplishing this is to allow nodes to exchange control packets. For example, recent works such as [6] and [8] have proposed enhancing the multi-rate capabilities of the IEEE 802.11 RTS/CTS handshake mechanism to obtain channel information. In particular, [8] proposes the Receiver Based Auto Rate (RBAR) protocol in which the receivers use physical-layer analysis of received RTS packets to find out the maximum possible transmission rate that achieves less than a specific bit error rate. The receiver then controls the sender's transmission rate by piggybacking this information into the CTS packet. In cognitive radio systems, channel probing may be accomplished by using a spectrum sensor at the physical layer (see, for example, [5]), whereby at the beginning of each time slot, the spectrum sensor detects whether a channel is available. This detection may be imperfect, and energy/hardware constraints might limit the number of channels sensed in a given slot.

In all these scenarios, channel probing can help the transmitter obtain useful information and therefore make better decisions about which channel to use for transmission. On the other

hand, channel measurement and estimation consume valuable resources; the exchange of control packets or spectrum sensing consumes energy and decreases the amount of time available to send actual data. Thus, channel probing must be done efficiently to balance the tradeoff between the two.

In this paper, we study optimal strategies for a joint channel probing and transmission problem. Specifically, we consider a transmitter with multiple channels of known state distributions. It can sequentially probe any channel with channel-dependent costs. The goal is to decide which channels to probe, in what order, when to stop, and upon stopping, which channel to use for transmission. Similar problems have been studied in [3], [6], [7], [9], and [10]. The commonality and differences between our study and previous work are highlighted within the context of our main contributions, summarized as follows.

First, we derive key properties of the optimal strategy for the problem outlined above and show that it has a threshold property and can only take on one of a few structural forms. In contrast to [3], [9], and [10], we do not restrict the channels to take a finite number of states; our work also applies to the case of (uncountably) infinite channel states. This generalization is useful if one uses the probability of successful transmission as channel state.

Second, we explicitly derive the optimal strategy for a number of special cases of practical interest. In [6] and [7], variants of the problem outlined above were studied. In particular, [7] analyzed a problem where channels can only be used immediately after probing (i.e., no recall of past channel probes) and unprobed channels cannot be used for transmission. Under these conditions, the problem reduces to an optimal stopping time problem for a given ordering of channels to be probed. In this paper, we allow both recall and transmitting in unprobed channels; the resulting problem is thus quite different from the optimal stopping time problem. [6] assumes independent Rayleigh fading channels and, because all channels are independent and identically distributed, does not focus on which channels should be probed and in what order. In contrast, we consider channels that are not necessarily statistically identical.

Finally, based on the key structural properties of the optimal strategies, we present an algorithm that computes the optimal strategy in a finite number of steps even when the channel has an uncountably infinite state space. We also propose computationally efficient strategies that, although potentially suboptimal, perform well for an arbitrary number of channels and arbitrary number of channel states (finite or infinite). To the best of our knowledge, these are the first channel probing algorithms for the combined scenario of an arbitrary number of channels, arbitrary channel distributions, statistically nonidentical channels, and possibly different probing costs.

The remainder of this paper is organized as follows. We formulate two channel probing problems in Section II and present important structural results on the optimal strategy in Section III. Three algorithms for the first problem are then presented in Section IV and are shown to be optimal for a number of special cases. The incorporation of additional regulatory constraints into the first problem is discussed in Section V. These results are then extended to the second problem in Section VI. Section VII provides numerical results, and Section VIII concludes the paper.

II. PROBLEM FORMULATION

We consider a wireless system consisting of N channels, indexed by the set $\Omega = \{1, 2, \cdots, N\}$, and a single transmitter who wants to send a message (to a receiver) using exactly one of the channels. (While there may be multiple transmitters and receivers present in the network, we limit our attention to a single transmitter–receiver pair in this paper.)

With each channel j, we associate a reward of transmission denoted by X_j, which is a random variable (discrete or continuous) with some distribution over some bounded interval $[0, M]$ where $M < \infty$. We call this the channel reward. The $\{X_j\}$ may represent either the probability of transmission success or the data rate of using channel j. The randomness of the transmission probability or data rate comes from the time-varying and uncertain nature of the wireless medium. It is assumed that the transmitter knows *a priori*[1] the distribution of X_j for all $j \in \Omega$, and by probing channel j, it finds out the exact realization[2] of X_j.

We assume $\{X_j\}_{j \in \Omega}$ are independent random variables, thus probing channel j does not provide any information about the state of any other channel in $\Omega - \{j\}$. If channels are correlated, then one can update the distributions of these random variables every time a channel is probed. However, this leads to a very different problem than the one presented below and is therefore not further considered in the present paper. Note that the interchannel independence assumption does not necessarily mean that the transmitter can only use one channel at a time. This is because we can think of each channel as a family of channels and probing simply determines the values of representative channels. For example, in an OFDM system, probing one OFDM tone j may reveal the value of all tones within a coherence bandwidth of j (the channel family). In this case, X_j could represent the reward of the best channel in the channel family (for single-channel access) or the collective reward of the entire family (for multichannel access).

Note that in reality, channel probes may only allow the transmitter to measure the received signal-to-noise ratio (SNR) [6], [7]. This measured SNR, however, essentially affects the probability of transmission success or data rate and translates into a measured valued of X_j. Thus, X_j can be thought of as an abstraction of the information obtained through probing. We will associate a cost c_j, where $c_j > 0$, with probing channel j.

The system proceeds as follows. The transmitter first decides whether to probe a channel in Ω or to transmit using one of the channels, based only on its *a priori* information about the distribution of X_j. If it transmits over one of the channels, the process is complete. Otherwise, the sender probes some channel $j \in \Omega$ and finds out the value of X_j. Based on this new information, the sender must now decide between using channel j for transmission, probing another channel in $\Omega - \{j\}$ (will also be denoted simply as $\Omega - j$ for the rest of the paper), or using a channel in $\Omega - j$ for transmission even though it has not been probed. This

---

[1]Many techniques can be used to estimate the distributions of $\{X_j\}$, e.g., via a moving average [7].

[2]It should be noted that this is assumed without loss of generality: When channel probing gives partial (or noisy) information about the channel state, we can let X_j denote the *expected* probability of transmission success (or data rate).

decision process continues until the user decides which channel to use for transmission.

The system thus operates in discrete steps. At each step, the transmitter has a set of unprobed channels $S \subseteq \Omega$ and has found out the states of channels in $\Omega - S$ through probing. It must decide between the following actions: 1) probe a channel in S; 2) use the best previously probed channel in $\Omega - S$, for which we say the user *retires*; or 3) use a channel in S for transmission, which we call *guessing* (also referred to as using a *backup channel* in [3]). Note that actions 2) and 3) can be seen as *stopping actions* that complete the process. The sequence of decisions on whether to continue to probe and which channel to probe or transmit in will be called a *strategy* or *channel selection policy*.

In practical situations, it could be the case that only a subset of channels in Ω may be guessed or retired to. For example, the transmitter may be allowed to transmit in the industrial, scientific, and medical (ISM) radio band without probing (perhaps within a power limit), but may be required to probe a TV band immediately before using it. In this paper, we will start by assuming that all channels may be guessed and retired to. We then show in Section V how the results derived under this assumption apply to the case where only a subset of Ω can be guessed or retired to and where the user is penalized for guessing on a busy channel.

The description above outlines a *one-shot* problem in that we are trying to make a decision for a one-time transmission. In this context, we will assume that the time it takes to go through the probing-transmission process (referred to as a *decision epoch*) is within the channel coherence time, which ensures that the realizations of X_i's remain constant in this time period. Later in Section V-A, we discuss how to handle fast fading channels in this framework. Since the problem is within a single decision epoch, we do not make any assumption on the temporal dependence of these channels from one epoch to another. If they are independent, then the same procedure can be repeated in each epoch; if they are not, then the distributions of X_i's will first need to be updated (e.g., using Bayesian methods) at the beginning of each epoch based on past observations and any information on the underlying correlation, and then the same procedure can be repeated.[3]

With the above assumptions, we now formulate two optimization problems corresponding to two different objectives. Justification and interpretation follow each formulation.

A. Problem P1

We start by describing the objective for our first problem, and then we provide justification for considering this problem.

Problem 1: Given a set of channels, their probing costs, and statistics on the channel transmission success probabilities, the sender's objective is to choose the strategy that maximizes transmission reward less the sum of probing costs, i.e., achieving the following maximum:

$$J^* = \max_{\pi \in \Pi} J^\pi = \max_{\pi \in \Pi} E\left[X_{\pi(\tau)} - \sum_{t=1}^{\tau-1} c_{\pi(t)}\right] \qquad (1)$$

[3]This essentially results in a greedy approach that optimizes associated objectives for each epoch; one can also try to optimize over a finite or infinite horizon (of these epochs) through a Markov decision process (MDP).

where π denotes a strategy that probes channels in the sequence $\pi(1), \cdots, \pi(\tau - 1)$, then transmits over channel $\pi(\tau)$ at time τ. Π denotes the set of all possible strategies for Problem 1 (referred to as P1 below), and the right-hand sum in (1) is set to 0 if $\tau = 1$.

Note that τ is a random stopping time that, in general, depends on the result of channel probes, and $P(1 \le \tau \le N+1) = 1$ since the longest strategy is to probe all N channels and then use one for transmission. For the rest of this paper, we will let π^* denote the strategy that achieves J^* in (1) and will refer to π^* as the *optimal (P1) strategy*. Such a strategy is guaranteed to exist since there are a finite number of strategies due to the finite number of channels.

We now provide two interpretations of P1.

Data maximization given constant data time (P1-DM): P1 may be seen as maximizing the total amount of data transmitted over a fixed amount of transmission time T, where each probe takes Δ_j amount of time not included in T.[4] To see this interpretation, let the random variable X_j denote the data rate associated with channel j. Thus, under strategy π, the user successfully transmits $X_{\pi(\tau)} \cdot T$ units of data after probing for $\sum_{t=1}^{\tau-1} \Delta_t$ amount of time.

Now, consider a baseline strategy that forgoes channel probing and in return gets to transmit at some constant data rate r over the same amount of time the it takes to probe and transmit under π. The total amount of data this baseline strategy transmits is $r(T + \sum_{t=1}^{\tau-1} \Delta_t)$. Suppose the user wants to maximize its advantage over the baseline strategy

$$\bar{J}^* = \max_{\pi \in \Pi} \bar{J}^\pi = \max_{\pi \in \Pi} E\left[X_{\pi(\tau)}T - r\left(T + \sum_{t=1}^{\tau-1} \Delta_{\pi(t)}\right)\right].$$

The above objective function reflects the desire to balance between obtaining a high rate through probing (the first term) and minimizing probing time (second term). Note that since the user has a constant transmission time T, simply maximizing $X_{\pi(\tau)}T$ will produce a trivial solution: The best strategy would be to probe every channel and use the best one.

It can be seen that since the term $-rT$ in $\bar{J}^\pi$ is the same for all strategies π, by letting $c_j = r(\Delta_j/T)$, the strategy maximizing $\bar{J}^\pi$ also maximizes J^* in P1, and $\bar{J}^* = -r + \max_{\pi \in \Pi} E[X_{\pi(\tau)} - \sum_{t=1}^{\tau-1} r(\Delta_{\pi(t)}/T)]$.

Throughput maximization given constant data (P1-TM): P1 can also be seen as maximizing throughput for a fixed amount of data. To see this interpretation, consider transmitting one unit of data and let X_j again denote the data rate associated with channel j. The throughput under a strategy π is given by $1/(1/X_{\pi(\tau)}) + (\sum_{j=1}^{\tau-1} \Delta_j)$. Maximizing this quantity is equivalent to maximizing $-(1/X_{\pi(\tau)}) - \sum_{t=1}^{\tau-1} \Delta_j$, which in turn is equivalent to solving P1 by setting $c_j = \Delta_j$ for each j and replacing random variables X_j with $Y_j = -(1/X_j)$.

Thus we have shown that P1 is equivalent to a data maximization problem and a throughput maximization problem, respectively. Due to this equivalence, in the rest of the paper, we will not make the distinction and will simply refer to this problem as P1.

[4]This is also called the *constant data time* (CDT) problem in [7].

Because the X_j's are bounded rewards in P1, then J^* is also upper-bounded by M. Thus, we will assume $0 < E[X_j] < M$ for all $j \in \Omega$. This is because if $E[X_j] = M$, then it is always optimal to use channel j without probing, and if $E[X_j] = 0$, then channel j is never probed or used; the optimal strategy becomes trivial if these assumptions are violated.

It can be shown that at any step, a sufficient information state (see, e.g., [12, ch. 6, pp. 82–84]) is given by the pair (u, S), where $S \subseteq \Omega$ is the set of unprobed channels and $0 \leq u \leq M$ is the highest probed value among channels in $\Omega - S$. The dynamic programming representation of the decision process is as follows. Let $V(u, S)$ denote the value function, i.e., maximum expected remaining reward given the system state is (u, S). This can be written mathematically as

$$V(u, S) = \max \left\{ \max_{j \in S} \left\{ -c_j + E\left[V\left(\max\{u, X_j\}, S - j\right)\right] \right\}, \right.$$
$$\left. u, \max_{j \in S} E[X_j] \right\} \quad (2)$$

where all of the above expectations are taken with respect to random variable X_j. The three terms on the right-hand side of (2) represent, respectively, the expected reward of probing the best channel in S, of using the best-probed channel, and of guessing the best unprobed channel. $V(0, \Omega)$ denotes the expected total reward of the optimal strategy.

B. Problem P2

An alternative formulation of the problem seeks to maximize the total amount of data transmitted within a fixed amount of time T available for both probing and transmission, when each probe takes Δ amount of time.[5] We assume $N\Delta < T$ so that the transmitter has the option of probing every channel. Since the total amount of time is fixed, this can be equivalently viewed as throughput maximization.

Problem 2: We seek the strategy maximizing the following:

$$\bar{J}^* = \max_{\lambda \in \Lambda} E\left[X_{\lambda(\tau)} \cdot (T - \Delta(\tau - 1))\right] \quad (3)$$

where $\lambda(\tau)$ is the channel that strategy λ uses for transmission after $\tau - 1$ probes, and Λ is the set of all possible P2 policies. We will denote by λ^* the strategy that maximizes the expectation given by (3).

Unlike in P1 where the information state is given by the pair (u, S), in P2 the value function also depends on Ω, or equivalently, the amount of time left, denoted by $\bar{T} = T - |\Omega - S| \cdot \Delta$. Consequently, the information state is the triple $(\bar{T}, u, S)$, while noting that $\bar{T}$ is obtainable from S if Ω is also given. The maximum expected remaining reward $H(\bar{T}, u, S)$, analogous to (2), is given by

$$H(\bar{T}, u, S) = \max \left\{ \bar{T}u, \max_{j \in S} \left\{ \bar{T}E[X_j] \right\}, \right.$$
$$\left. \max_{j \in S} \left\{ E\left[H\left(\bar{T} - \Delta, \max\{X_j, u\}, S - j\right)\right] \right\} \right\} \quad (4)$$

where the three terms represent, respectively, the reward of retiring, using channel j without probing, and probing j followed by the optimal strategy.

Note that while the dynamic programs are readily available in both P1 and P2, computing the value function $V(\cdot, \cdot)$ and $H(\cdot, \cdot, \cdot)$ for every state is very difficult and practically impossible because the state space is potentially infinite and uncountable since u can be any real number in $[0, M]$ if the X_j's are continuous random variables.[6] Rather than directly computing these values, the approach we take in this paper is to first derive fundamental properties of optimal strategies and then use them to construct simpler algorithms in Section IV.

For P1, any strategy can be defined by the set of actions it takes with respect to its entire set of information states, $\cup_S \cup_u (u, S)$. We let retire(u), probe(j), and guess(j), $j \in S$, denote the three options that the sender has and must choose from. We let $\pi(u, S)$ denote the action taken by strategy π when the state is (u, S). We use similar notations for P2: λ denotes the strategy, and $\lambda(\bar{T}, u, S)$ denotes the action under the strategy in state $(\bar{T}, u, S)$.

The detailed analysis in this paper primarily deals with P1 due to space limitation and its relative simplicity in presentation. Then, in Section VI, we show how our results on P1 strategies apply to P2 strategies.

III. PROPERTIES OF THE OPTIMAL STRATEGY

In this section, we establish key properties of the optimal P1 strategy. Unless otherwise stated, all proofs are given in the Appendix.

A. Threshold Property of the Optimal Strategy

We first note that for all $S \subseteq \Omega$ and any $\tilde{u} \geq u$, $V(u, S) \leq V(\tilde{u}, S)$, i.e., $V(\cdot, S)$ is nondecreasing. This inequality follows from (1) and (2). In particular, any channel selection strategy cannot have smaller reward when starting from $(\tilde{u}, S)$ rather than (u, S) since the set of unprobed channels is the same for both cases, while the best probed channel for the latter case is better than that of the former scenario. Thus, $V(\cdot, S)$ is a nondecreasing function. Similarly, it can be established that $V(u, \cdot)$ is a nondecreasing function, i.e., for all $u \in [0, 1]$ and any $\check{S} \supseteq S$: $V(u, S) \leq V(u, \check{S})$. We have the following.

Lemma 1: Consider any state (u, S). If $V(u, S) = u$, then $V(\tilde{u}, S) = \tilde{u}$ for all $\tilde{u} \geq u$.

Lemma 2: Consider any state (u, S). If $V(u, S) = E[X_j]$ for some $j \in S$, then $V(\bar{u}, S) = E[X_j]$ for all $\bar{u} \leq u$.

Proof of Lemma 1 can be found in [13]. Lemma 2 follows directly from (2) and $V(\cdot, S)$ being nondecreasing since these equations imply $E[X_j] \leq V(\bar{u}, S) \leq V(u, S) = E[X_j]$. Its proof is therefore not included in the Appendix.

The above two lemmas imply that for fixed S, the optimal strategy has a *threshold* structure with respect to u. In particular, for any set $S \subseteq \Omega$, we can define the following:

$$a_S = \inf \{u : V(u, S) = u\} \quad (5)$$
$$b_S = \sup \{u : V(u, S) = E[X_j], \text{some } j \in S\} \quad (6)$$

---

[5]This is also called the *constant access time* (CAT) problem in [7].

[6]The direct computation of such problems usually involves approximation and discretizing the state space.

where the right-hand side of (5) is nonempty since $V(M, S) = M$ is always true. We set $b_S = 0$ if the set on the right-hand side of (6) is empty. Note that both a_S and b_S are completely determined given the set S. It follows from Lemmas 1 and 2 that $0 \leq b_S \leq a_S \leq M$. Thus, we have the following corollary.

Corollary 1: For any state (u, S), there exists an optimal strategy π^* and constants $0 \leq b_S \leq a_S \leq M$ satisfying

$$\pi^*(u, S) = \begin{cases} \text{retire}(u), & \text{if } u \geq a_S \\ \text{probe}(j_u), j_u \in S, & \text{if } b_S < u < a_S \\ \text{guess}(j), j \in S, & \text{if } u < b_S. \end{cases}$$

It should be noted for completeness[7] that at $u = b_S$, $\pi^*(u, S) = \text{guess}(j)$ if $b_S > 0$; otherwise, $\pi^*(u, S) = \text{probe}(j)$. Also, note that the optimal channel to probe, j_u, in general depends on the value of u. This corollary indicates that there exists an optimal strategy with the described threshold structure. It remains to determine these thresholds, which can be difficult especially for large S. It also remains to determine which channel should be probed in the "probe" region above.

To help overcome the difficulty in determining a_S and b_S for a general S, we first focus on quantities $a_{\{j\}}$ and $b_{\{j\}}$ (subsequently simplified as a_j and b_j) for a single element $j \in \Omega$, which can be determined relatively easily from (5) and (6), respectively, as shown below. These are thresholds (also referred to as indices below) concerning channel j that are *independent of other channels*. We will see that they are very useful for reducing the complexity of the problem.

We now take a closer look at a_j and b_j. Note that at state (u, j), $\text{probe}(j)$ results in expected reward $-c_j + E[\max\{u, X_j\}]$ since there are no more channels to probe after j. Action $\text{guess}(j)$ gives the expected reward $E[X_j]$, while retiring gives reward u. The assumptions $0 < E[X_j] < M$ and $c_j > 0$ imply that, for sufficiently small u, the probing reward becomes less than the guessing reward. By comparing the rewards of these three options, it can be seen that guessing is optimal if $E[(u - X_j)I_{\{X_j < u\}}] \leq c_j$ and $u \leq E[X_j]$, where $I_{\{\cdot\}}$ is the indicator function. We will adopt the notation that, for any random variable Z, $E[Z^+] = E[Z \cdot I_{\{Z > 0\}}]$.

Similarly, when u is sufficiently large, the probing and guessing reward become less than the reward for retiring, u. Thus, for any $j \in S$ we have the following:

$$a_j = \min\left\{u : u \geq E[X_j], c_j \geq E\left[(X_j - u)^+\right]\right\} \quad (7)$$
$$b_j = \max\left\{u : u \leq E[X_j], c_j \geq E\left[(u - X_j)^+\right]\right\}. \quad (8)$$

Note that $a_j \geq E[X_j] \geq b_j$. In addition, $a_j = b_j$ if and only if $E[X_j] = a_j = b_j$. It also follows that for $b_j < u < a_j$, probing is strictly an optimal strategy. It can be seen from the above that c_j essentially controls the width of this probing region; for larger c_j, a_j and b_j will be closer to $E[X_j]$.

The above discussion is depicted in Fig. 1, where we have plotted the expected reward of the three actions $\text{guess}(j)$

[7]It can be shown that $V(\cdot, S)$ is a continuous function. If $b_S > 0$, then by definition $V(u, S) = \max_{j \in S} E[X_j]$ for all $u < b_S$, which implies by continuity that $V(b_S, S) = E[X_j]$ for some j. Thus, $\pi^*(b_S, S) = \text{guess}(j)$. If $a_s > b_S = 0$, then it can be shown there exists $\epsilon > 0$ such that $\pi^*(u, S) = \text{probe}(j)$ for some j and all $0 < u < \epsilon$. Then, by continuity of $V(\cdot, S)$, we have $\pi^*(0, S) = \text{probe}(j)$.
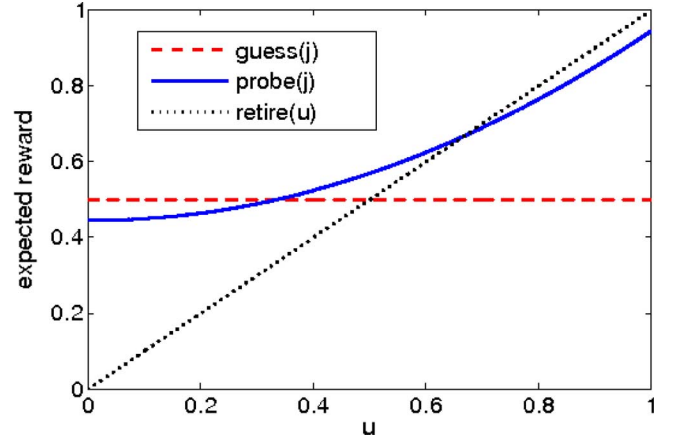

Fig. 1. As described in Section III-A, when j is the only unprobed channel and X_j is uniformly distributed in $[0,1]$, the expected reward from actions $\text{guess}(j)$, $\text{probe}(j)$, and $\text{retire}(u)$ as functions of u. Note that $a_j = 2/3$ (the crossing point of solid and dotted lines) and $b_j = 1/3$ (the crossing point of solid and dashed lines).

(dashed line), $\text{probe}(j)$ (solid line), and $\text{retire}(u)$ (dotted line) as functions of u when X_j is uniformly distributed in $[0,1]$ and $c_j = 1/18$. In this case, $a_j = 2/3$ and $b_j = 1/3$. Note that increasing (decreasing) c_j would shift the solid curve down (up), thus decreasing (increasing) the width of the middle region where $\text{probe}(j)$ is the optimal action.

This example demonstrates a method for computing a_j and b_j for any channel j. Notice that to determine these two constants, we simply need to take the intercepts between the following three functions of u: $f_1(u) = E[X_j]$, $f_2(u) = u$, and $f_3(u) = -c_j + uP(X_j \leq u) + E[X_j I_{\{X_j > u\}}]$. Thus, regardless of whether X_j is continuous or discrete, computing a_j and b_j is not very complex.

In the rest of this section, we derive properties of the optimal strategy expressed in terms of individual indices a_j and b_j.

B. Structure of the Optimal Strategy

We first present an algorithm to sort any set of channels S based on indices $\{a_j\}_{j \in S}$, which will help describe properties of the optimal strategy throughout this paper.

Algorithm 1: (Sorting Algorithm):

Initially: $k = 1$. Let $N = |S|$. The algorithm proceeds as follows:

1) Compute R and j^* according to the following equations:

$$R = \left\{ j \in S : a_j = \max_{l \in S} a_l \right\}. \quad (9)$$

$$j^* = \arg\max_{j \in R}\left\{ I_{\{b_j = a_j\}}E[X_j] + I_{\{a_j > b_j\}} \right.$$
$$\left. \times \left\{ E[X_j | X_j \geq a_j] - \frac{c_j}{P(X_j \geq a_j)} \right\} \right\}. \quad (10)$$

Let $s_k = j^*$; $k = k + 1$; $S = S - \{j^*\}$.
2) If $|S| \geq 1$, repeat Step 1; otherwise, stop and return the sorted set $\{s_1, \cdots, s_N\}$.
3) Relabel the sorted set as $\{1, 2, \cdots, N\}$.

We see that Algorithm 1 takes any set of channels S and replaces it with an equivalent sorted set $\{s_1, \cdots, s_N\}$, which is then relabeled $\{1, 2, \cdots, N\}$. The channels are sorted in decreasing order of a_j. Whenever $a_k = a_{k+1}$, then sorting proceeds according to (10), where the tiebreaker essentially sorts channels according to their one-step reward of probing/guessing channel j when $u = a_j$ and j is the only remaining channel.

We use this sorting to describe the following important result on the optimal strategy, which will be proven throughout various parts of this section, as described below.

Theorem 1: For any set of channels S sorted according to Algorithm 1, there exists a constant d_S such that $d_S \leq \max_{j \in S} b_j$ and the following holds:

1) For all $u \geq a_1$, $\pi^*(u, S) = \text{retire}(u)$. If $u < a_1$ then $\pi^*(u, S) \neq \text{retire}(u)$.
2) For all $d_S < u < a_1$, $\pi^*(u, S) = \text{probe}(1)$.
3) For all $0 \leq u < d_S$, exactly one of the following holds:
   a) $\pi^*(u, S) = \text{probe}(1)$;
   b) $\pi^*(u, S) = \text{guess}(1)$;
   c) $\pi^*(u, S) = \text{probe}(k)$, $k \in S - 1$;
   where channel k does not vary with u.

We note that d_S indicates the highest value of u such that one of the cases 3a, 3b, 3c of Theorem 1 holds. When case 3b is true, b_S and d_S coincide. For cases 3a and 3c, we have $d_S > b_S = 0$ since it is not optimal to guess for all u.

The proof of Theorem 1 is broken down separately in subsequent sections and in the Appendix as follows. Part 1 of Theorem 1 is proven in Section III-C. This result provides both a necessary and sufficient condition for the optimality of retiring and using a previously probed channel. A very appealing feature of this result lies in the fact that it allows us to decide when to retire based only on *individual* channel indices that are calculated independent of other channels, thus reducing the computational complexity.

Part 2 of Theorem 1 is also proven in Section III-C. This result implies that by first ordering the individual channels by functions of the indices a_j, we can determine the optimal channel to probe for u in the interval (d_S, a_1).

Finally, Part 3 of Theorem 1 gives three possibilities on the structure of the optimal strategy. Parts 3a and 3c, proven in Section III-C, indicate that the optimal channel to probe does not vary with u in the region $(0, d_S)$. Meanwhile, part 3b narrows down the set of possible channels we can guess. The channel in S with the highest value of a_j and sorted according to (10), which we have called 1, is the only possible channel we can guess. This result is proven in Sections III-D. A key result in that section is that if there are multiple channels in R, then we can easily check whether $a_j = b_j$ is true in order to determine whether probing or guessing is the optimal action. Section III-D also provides some necessary and sufficient conditions for guessing to be optimal.

Theorem 1 significantly reduces the number of possibilities on the structure of the optimal strategy, but it remains to determine when cases 3a, 3b, 3c of Theorem 1 hold along with the value of d_S. In general, this structure will depend on the specific values of S and the indices a_j, b_j. One can use the results of Sections III-C and III-D to determine some necessary or sufficient conditions for any particular case of the above theorem
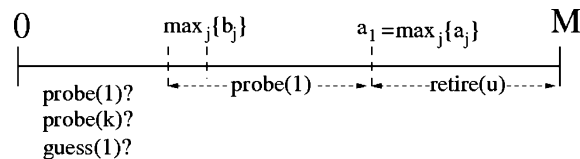

Fig. 2. Summary of main results from Section III. Figure depicts optimal strategy $\pi^*(u, S)$ as a function of u. For the middle and right regions of the line, the optimal strategy is well-defined for any S. For the left region, the optimal action may depend on S.

to hold. In Section IV, we will propose a suboptimal algorithm, based on these three possible forms, which we show to be optimal under a number of special cases of interest.

Fig. 2 summarizes the main results from Theorem 1. For all $u \geq a_1$, i.e., right region of the line, $\text{retire}(u)$ is optimal. For $\max_{j \in S} b_j < u < a_1$, i.e., the middle region of the line, $\text{probe}(1)$ is optimal. Note that it is possible this region may be empty if the probing costs become too high. Finally, the optimal action in the left region will depend on S and thus remains to be determined. Note that $\text{guess}(1)$ is the only possible guessing action for this region, as proven in Lemma 6 and Corollary 2.

C. Optimal Retiring and Probing

In this subsection, we prove Parts 1, 2, 3a, and 3c of Theorem 1 by deriving conditions under which it is optimal to retire or probe a channel.

We begin with the following lemma.

Lemma 3: For any (u, S), $\pi^*(u, S) = \text{retire}(u)$ if and only if $u \geq \max_{j \in S}\{a_j\}$. Equivalently, $a_S = \max_{j \in S} a_j$.

Proof of this lemma can be found in [13, Appendix 9.2]. This lemma provides both a necessary and sufficient condition for the optimality of retiring and using a previously probed channel. This lemma proves Part 1 of Theorem 1. As previously mentioned, a very appealing feature of this lemma lies in the fact that it allows us to decide when to retire based only on individual channel indices that are computed independent of other channels.

We now examine when it is optimal to probe and which channels to probe. In order to shed light on the best channels to probe, we present the optimal strategy for a separate but related problem. It will be seen that analysis on this problem is crucial for deriving useful properties of the optimal strategy.

No Guessing (NG) Problem: Consider Problem 1 with the following modification: At each step, the user must choose between the *two* actions: 1) probe an unprobed channel; or 2) retire and use the best probed channel. Therefore, the user is not allowed to transmit using an unprobed channel.

The NG Problem can be seen as a generalization of [3, Section IV, Theorem 4.1], which restricted X_j to be discrete random variables. Note that even though guessing is removed as a possible action, the resulting problem is still very different from the classical optimal stopping problem for two reasons. First, we allow recall in this problem, while it is typically not allowed in the latter. Second and more importantly, the NG problem is not only trying to decide when to stop, but also trying to figure out the best probing sequence. By contrast, in a classical stopping problem, the sequence is considered (randomly) given and not controlled. For instance, in [14], a multiuser single-channel

access problem was considered, where users competing for the channel decide when to use the channel when they gain the access depending on their perceived channel quality. This is in a sense "probing" the users (as opposed to probing the channels) to decide when to stop and let a user transmit, but in this case, the "probing" sequence is random (each user has a fixed probability of gaining access) and not up to the decision process. Interestingly, the problem studied in [14] was shown to reduce to an optimal stopping problem.

To describe the theorem, we use the following notation for any channel $j \in \Omega$:

$$\bar{a}_j = \min \left\{ u : c_j \geq E\left[(X_j - u)^+ \right] \right\} \qquad (11)$$

where $\bar{a}_j = 0$ if the above set is empty. Note that from (7) and (8), we see that $\bar{a}_j = a_j$ if $a_j > b_j$. If $a_j = b_j$, then we have $c_j \geq E[(X_j - u)^+]$ for all $u \geq a_j$, and thus $\bar{a}_j \leq a_j$ by (11). We use these indices in the following theorem, which can be seen as a generalization of [3, Theorem 4.1].

Theorem 2: For state (u, S), the optimal strategy $\hat{\pi}$ for the NG Problem is described as follows:
1) If $u \geq \max_{j \in S} \{ \bar{a}_j \}$, then $\hat{\pi}(u, S) = \text{retire}(u)$.
2) Otherwise, sort the set S according to Algorithm 1 by replacing a_j with $\bar{a}_j$ for all j. Then, $\hat{\pi}(u, S) = \text{probe}(1)$.

Even though the NG Problem is different from Problem 1, its optimal strategy will also be optimal for Problem 1 if guessing becomes nonoptimal for all future time steps. From Lemma 3 and definition of b_j, guessing is nonoptimal for all future time steps, and probing occurs if $\max_{j \in S} b_j < u < \max_{j \in S} a_j$. Thus, we have proven the following lemma.

Lemma 4: For any set S sorted according to Algorithm 1, $\pi^*(u, S) = \text{probe}(1)$ for all $\max_{j \in S} b_j < u < \max_{j \in S} a_j$.

This result completes the proof of Part 2 in Theorem 1 for $u > \max_{j \in S} b_j$.

To prove Parts 3*a* and 3*c* of Theorem 1, we prove the following result.

Lemma 5: Consider any S sorted according to Algorithm 1. If $\pi^*(u, S) = \text{probe}(k)$ for some $u > 0$ and $k > 1$, then $\pi^*(\tilde{u}, S) = \text{probe}(k)$ for all $0 \leq \tilde{u} \leq u$.

Proof of Lemma 5 can be found in [13, Appendix 9.4]. This result implies that if $\pi^*(u, S) = \text{probe}(k)$ for some $u > 0$ and $k > 1$, then $\pi^*(\bar{u}, S) \neq \text{probe}(l)$ for all $l \neq 1, k$ and $\bar{u} \geq u$.

We note that d_S in Theorem 1 satisfies $d_S \leq \max_{j \in S} b_j$ due to the following. From Lemma 4, we know that $\pi^*(u, S) = \text{probe}(1)$ for all $\max_{j \in S} b_j < u < \max_{j \in S} a_j$. In addition, from Lemma 5, if for some $\bar{u} \leq \max_{j \in S} b_j$ we have $\pi^*(\bar{u}, S) = \text{probe}(k)$ where $k \neq 1$, then $\pi^*(u, S) = \text{probe}(k)$ for all $u < \bar{u}$. For $\bar{u} \leq u \leq \max_{j \in S} b_j$, we know that $\pi^*(u, S) \neq \text{guess}(j), \text{retire}(u)$ due to Lemmas 1 and 2. Thus, it is only possible that $\pi^*(u, S) = \text{probe}(1)$ for $\bar{u} \leq u \leq \max_{j \in S} b_j$. Therefore, 3*a* and 3*c* are the only possible forms for the optimal strategy that involve probing a channel.

D. Optimal Guessing

We now prove Part 3b of Theorem 1 by deriving conditions for guessing to be optimal. Note that $b_S = 0$ implies guessing is not optimal for all $u \in [0, 1]$.

Lemma 6: Given a set of unprobed channels S, define R as in (9). Then, we have:
1) If there exists $j^* \in R$ such that $a_{j^*} > b_{j^*}$ and $b_{j^*} \leq \max_{j \in S - j^*} E[X_j]$, then $b_S = 0$.
2) If there exists $j^* \in R$ such that $b_{j^*} \geq \max_{j \in S - j^*} a_j$, then $b_S = b_{j^*}$.

Proof of this lemma can be found in [13, Appendix 9.5]. Conditions 1) and 2) of the lemma provide separate necessary and sufficient conditions for guessing to be optimal. Note that this lemma also has further implications. When $|R| \geq 2$, and $a_j = b_j$ for at least one $j \in R$, then condition 2) of Lemma 6 is always satisfied. Thus, $b_S = b_j$ in this case. Otherwise, $a_j > b_j$ for all $j \in R$, and condition 1) of Lemma 6 is always satisfied.

On the other hand, when $|R| = 1$ and letting $1 = R$, suppose $\pi^*(u, S) = \text{guess}(k)$ for some $k \neq 1, u > 0$. This implies $E[X_k] > E[X_1] \geq b_1$, which leads to condition 1) of Lemma 6 if $a_1 > b_1$. This lemma implies that we have $b_S = 0$, which contradicts the assumption that $\pi^*(u, S) = \text{guess}(k)$. Thus, if $|R| = 1$, then $\pi^*(u, S) \neq \text{guess}(k)$ for $k \notin R$. Similarly, if $a_1 = b_1$ and again $|R| = 1$, then we have $a_1 = E[X_1] > a_k \geq E[X_k]$, which is again a contradiction to $\pi^*(u, S) = \text{guess}(k)$. This leads to the following corollary.

Corollary 2: Given a set S, define R as in (9). Then, if $|R| \geq 2$ and $a_j = b_j$ for at least one $j \in R$, then $b_S = b_j$. Otherwise, $b_S = 0$. If $|R| = 1$, let $\{1\} = R$. Then, $\pi^*(u, S) \neq \text{guess}(j)$ for all u and $j \in S - 1$.

This corollary and its preceding lemma narrow the set of possible channels we can guess to a single channel, i.e., the channel j with the highest value of a_j. If there are multiple channels achieving this maximum, then we can easily check whether $a_j = b_j$ in order to determine whether probing or guessing is the optimal action.

In order to complete the proof of Part 3b in Theorem 1, it remains to show that if $b_S > 0$ (i.e., $\pi^*(u, S) = \text{guess}(1)$ for some $u > 0$), then $\pi^*(u, S) \neq \text{probe}(k)$ for all u and $k \neq 1$. This is easily proven by using Lemma 2 and the contrapositive of Lemma 5.

E. Decomposition of Problem 1

The following result on the structure of the optimal strategy allows us to decompose Problem 1 into N subproblems. To begin, define $\Phi(j)$, $j = 1, 2, \cdots, N$, to be the set of strategies that do not guess any channel except possibly channel j. Within each set $\Phi(j)$, we define the best strategy [achieves the value function in (2)] by π_j^*: $\pi_j^*(u, S) = \arg\max_{\pi \in \Phi(j)} V^\pi(u, S)$, where $V^\pi(u, S)$ is the expected remaining reward under policy π given the system state (u, S). We can show the optimal strategy satisfies $\pi^* \in \cup_{j \in \Omega} \{\Phi(j)\}$. This result was proven in [3, Theorem 5.2] for a three-channel system and with discrete channel rewards.

Lemma 7: For any (u, S), there exists an optimal strategy $\pi^*(u, S) = \arg\max_{\pi \in \Pi} V^\pi(u, S)$, which also satisfies $\pi^*(u, S) = \arg\max_{j \in S} V^{\pi_j^*}(u, S)$.

That is, the optimal strategy will only guess one channel (if it guesses at all) over all possible realizations of channel rewards. Thus, the optimal strategy π^* among all strategies is the best π_j^* among all $j \in \Omega$. This result again reduces the number of possible optimal strategies. As the proof of this lemma is similar to

that of [3, Theorem 5.2], it is omitted for brevity. It can be shown that Lemma 7 can be extended to the case where the transmitter is only allowed to guess a subset $\bar{S} \subseteq S$ of the channels. In this case, one can replace S under the argmax in Lemma 7 with $\bar{S}$.

Finally, it remains to determine the structure of π_j^*. We have the following useful result.

Lemma 8: For any S and $j \in S$, define R as in (9), replacing a_k with $\bar{a}_k$ [defined in (11)] for every channel except j. If $j \in R$, then let $1 = j$. Otherwise, define 1 according to Algorithm 1, again replacing a_k with $\bar{a}_k$ for all $k \neq j$. Then, if $1 \neq j$, the optimal strategy π_j^* is

$$\pi_j^*(u, S) = \begin{cases} \text{retire}(u), & \text{if } u \geq a_1 \\ \text{probe}(1), & \text{if } u < a_1. \end{cases}$$

It can be shown that Theorem 1 also holds for each strategy π_j^*. Thus, Lemma 8 can be seen as arising from Part 3a in Theorem 1. Lemma 8 uniquely describes the optimal strategy π_j^* for any set of channels S, if $j \neq 1$. When $j = 1$, then the optimal strategy has a more complicated structure. In Section IV, we propose a suboptimal algorithm that approximates the optimal strategy when $j = 1$.

IV. JOINT PROBING AND TRANSMISSION STRATEGIES

As stated earlier, it is very difficult to recursively apply dynamic programming to evaluate all $V(u, S)$ and solve for the optimal strategy due to the uncountability of the state space. In this section, we first demonstrate how Theorem 1 can be used to derive a dynamic program that computes the optimal strategy in a finite number of steps even when the channel rewards are continuous random variables. This gives one possible method of determining the optimal strategy. We further propose two faster and more computationally efficient algorithms, motivated by the properties derived in the previous section. We show that they are optimal for a number of special cases of practical interest.

A. Value Function Parameterization

In this section, we show that Theorem 1 leads to a parameterization of the value function which can help determine $V(\cdot, \cdot)$ in a finite number of steps even if channel rewards are continuous, with the following corollary to Theorem 1.

Corollary 3: For any set S sorted according to Algorithm 1, let $f_1(u) = -c_1 + E[V(\max\{u, X_1\}, S - \{1\})]$ denote the expected reward of probing 1 at state (u, S). Then, $V(\cdot, S)$ has the following structure for some constant $0 \leq d_S < \max_{j \in S} b_j$: $V(u, S) = u$ if $u \geq a_1$; $V(u, S) = f_1(u)$ if $a_1 > u \geq d_S$; and $V(u, S) = f_1(d_S)$ if $u < d_S$.

We see that $V(\cdot, S)$ is uniquely determined by the constant d_S. Furthermore, for $u < d_S$, $V(u, S)$ is a constant. Thus, to determine the optimal strategy, it only remains to determine this constant for every S. We now explain how to calculate d_S for each S. From Theorem 1, if $V(u, S - \{j\})$ is determined for all j then d_S for $V(u, S)$ can be calculated by determining $V(0, S)$ as follows: $V(0, S) = \max\{E[X_1], \max_{1 \leq k \leq N} f_k(0)\}$, where $f_k(0)$ is defined in Corollary 3 by replacing 1 with k.

Then, d_S is the unique number satisfying the following: $V(0, S) = -c_1 + E[V(\max\{d_S, X_1\}, S - \{1\})]$. Therefore, determining d_S simply requires taking the intersection between constant $V(0, S)$ and the function

$-c_1 + E[V(\max\{u, X_1\}, S - \{1\})]$. Note that computing $V(0, S)$ also determines $\pi^*(0, S)$. Thus, from Theorem 1, $\pi^*(u, S) = \pi^*(0, S)$ for all $u < d_S$, and we have computed $\pi^*(\cdot, S)$.

$V(\cdot, \cdot)$ can thus be recursively determined by first calculating d_S for each singleton channel $j \in \Omega$, then using the above procedure to determine $V(\cdot, S)$ for all $|S| = 2$, etc. This procedure therefore gives a method to calculate the optimal strategy in a *finite* number of steps even if the channel rewards are continuous random variables. Note that this procedure does require considering all combinations of subsets of Ω, a total of $2^{|\Omega|} - 1$ of them. Thus, in practice this procedure is only applicable when the number of channels is not too large. In the next subsection, we propose faster algorithms which may be suboptimal but avoid computing over the power set of S and are thus computationally more efficient.

B. Channel Probing Algorithms

To motivate our first algorithm, recall that Theorem 1 shows that for fixed S, as u varies there can be at most two possible channels to probe, one of which must be 1. This gives rise to the following two-step look-ahead policy γ that only considers the two best channels 1 and 2 (i.e., pretending that $S = \{1 \cup 2\}$) and decides on the action by comparing the constants a_1, b_1, a_2, and b_2 using Corollary 1. To describe it, we use the same notation in the previous section: $f_1(u) = -c_1 + E[V(\max(X_1, u), 2)]$, which is the expected reward of probing 1 at state $(u, 1 \cup 2)$, and $f_2(u)$ is defined similarly by switching 1 and 2.

Algorithm 2: (A Two-Step Look-Ahead Policy γ for a Given Set of Unprobed Channels $S \subseteq \Omega$):

Step 1: Use Algorithm 1 to sort S and determine 1,2.
Step 2: Define strategy γ as follows for state (u, S):
    1) If $u \geq a_1$, then $\gamma(u, S) = \text{retire}(u)$.
    2) If $a_1 > u > \max\{b_1, b_2\}$, then $\gamma(u, S) = \text{probe}(1)$.
    3) If $u \leq \max\{b_1, b_2\}$, then we have the following cases:
        a) If $b_1 \geq a_2$, then $\gamma(u, S) = \text{guess}(1)$.
        b) If either $b_2 \geq b_1$ or $f_1(0) \geq \max\{E[X_1], f_2(0)\}$, $\gamma(u, S) = \text{probe}(1)$.
        c) Otherwise, there exists a unique b_0, where $b_1 > b_0 > b_2$ and $f_1(b_0) = \max\{E[X_1], f_2(0)\}$. Then, for $b_0 \leq u \leq b_1$, we have $\gamma(u, S) = \text{probe}(1)$. For $u < b_0$, we have $\gamma(u, S) = \text{guess}(1)$ if $E[X_1] \geq f_2(0)$. Otherwise, $\gamma(u, S) = \text{probe}(2)$.

It is worth describing this strategy in the context of results derived in the previous section. For u satisfying Case 1 of the algorithm description, γ is optimal from Theorem 1, Part 1, and Lemma 3. For u values described in Case 2, if $\max\{b_1, b_2\} = \max_{j \in S} b_j$, then γ is optimal from Theorem 1 and Lemma 4. For Case 3a, γ is optimal from Theorem 1, Lemma 6, and Corollary 2. Thus, γ is optimal for most values of u. For Cases 3b and 3c, the procedure essentially computes the expected probing cost if we are forced to retire in two steps.

We also propose a second two-step look-ahead algorithm, called β, that is motivated by Algorithm γ and Lemmas 7 and 8. Due to its similarity to γ, we present only a brief description.

Algorithm 3: (Two-Step Look-Ahead Policy β): For each channel $j \in S$ and the corresponding set of strategies $\Phi(j)$ defined in Section III-B, first find the best two channels indexed by 1 and 2 (analogous to Algorithm 2). If $1 \neq j$, then from Lemma 7, we can set β_j to be strategy π_j^* of that lemma. Otherwise, determine the best two-step strategy in $\Phi(j)$ using the two channels 1 and 2, similar to Algorithm 2, but replacing a_2 with $\bar{a}_2$ and setting $b_2 = 0$. Call this β_j. After β_j has been determined for all $j \in S$, using Lemma 7 take the best strategy among all $\{\beta_j\}$ to determine β.

When the transmitter can only guess a subset $\bar{S} \subseteq S$ of channels, we can modify Algorithm 3 by replacing S with $\bar{S}$.

Note that determining algorithm β requires running a similar algorithm to γ for each channel in S, thus requiring more computation. However, this strategy generally performs better than γ as we will show in Section VII. We next consider a few special cases and show that γ is optimal in these cases. It can also be shown that these results hold for β as well.

C. Special Cases

We first consider a two-channel system. Since Algorithm 2 is essentially a two-step look-ahead policy, we have the following.

Theorem 3: For any given set of unprobed channels S, where $|S| = 2$, γ is an optimal strategy.

The proof is omitted for brevity. We next consider the case of statistically identical channels with different probing costs.

Theorem 4: Suppose $|S| \geq 2$, and all channels in S are identically distributed, with possibly different probing costs. Then, the optimal strategy π^* is described as follows, with 1 being a channel in S satisfying $c_1 = \min_{j \in S}\{c_j\}$. If $u \geq a_1$, then $\pi^*(u, S) = \text{retire}(u)$. For all $u < a_1$ we have two cases:

Case 1) If $a_1 > b_1$, then $\pi^*(u, S) = \text{probe}(1)$.
Case 2) If $a_1 = b_1$, then $\pi^*(u, S) = \text{guess}(1)$.

Proof of Theorem 4 can be found in [13]. This theorem implies that if we have a set of statistically identical channels Ω, then the initial step of the optimal strategy is uniquely determined by a_1 and b_1, where 1 is the channel with smallest probing cost. If $a_1 = b_1$, then $\pi^*(u, \Omega) = \text{guess}(1)$, and it is not worth probing any channels. If $a_1 > b_1$, then we should first probe 1. Let k denote the channel with the smallest probing cost in $S - \{1\}$. If the probed value of X_1 is higher than a_k, then it is optimal to retire and use 1 for transmission. Otherwise, if $a_k > b_k$, then probe(k) is optimal; if $a_k = b_k$ then guess(k) is the optimal action. This process continues until we retire, guess, or $|S| = 1$, in which case the decision is straightforward by comparing u with a_j and b_j, $j \in S$.

Note that the optimal strategy described above is the same as strategy γ of Algorithm 2 applied to statistically identical channels. This is true because within Case 3 in the description of Algorithm 2, 3b will occur whenever $a_1 > b_1$ for statistically identical channels, and Case 3a occurs whenever $a_1 = b_1$. Collectively, Cases 1, 2, 3a, and 3b all describe the optimal strategy of Theorem 4. Note that this theorem applies to *all* cases of statistically identical channels, regardless of their distribution or probing costs. Changing the channel distribution and probing

costs will affect the values of a_j or b_j, but they do not alter the general structure of the optimal strategy as given by the theorem.

Finally, we consider the case where the number of channels is very large and not statistically identical.

Infinite Number of Channels (INC) Problem: Consider P1 with the following modification: We have N different types of channels, but an infinite number of each channel type.

Note that Theorem 4 solves this problem if $N = 1$. When referring to the state space for this problem, we let S denote the set of available channel *types*. Then, we have the following.

Theorem 5: For any set of channels S, the optimal strategy for Theorem 4 is also optimal for the INC Problem.

This theorem implies that when the number of channels is infinite, and there are an arbitrary number of channel types, then we will only probe or guess one channel. Note that Algorithm 2 is also the optimal strategy for the INC Problem since it is also the optimal strategy in Theorem 4.

Thus, we have shown Algorithm 2 is the optimal strategy for the special cases based on Theorems 3–5.

V. POLICY CONSTRAINTS

In this section, we discuss three generalizations of P1 that incorporate practical regulatory constraints. The first involves channels that must be probed immediately before transmission (i.e., cannot be recalled), the second involves channels that cannot be guessed, and the last incorporates a random penalty associated with guessing.

A. Probing Regulations

As mentioned in Section II, in practical systems it is possible that some channels cannot be guessed or retired to unless they were the last probed channel (i.e., no recall). This could be because these channels change conditions more rapidly and thus must be probed immediately before transmission.

To incorporate this scenario, we modify the problem formulation as follows. For any set S, let S_f and S_s denote the fast and slow fading channels, respectively; thus, $S = S_f \cup S_s$. We assume the coherence time for any channel in S_f is very short such that this channel's probing result is only valid immediately after probing. Beyond this, the channel reward is i.i.d, so the values are independent of probing results from earlier in the cycle. The user can probe this channel multiple times, with each probe resulting in a different independently drawn value. Channels in S_s behave as in previous sections.

Letting $\bar{V}$ denote the value function for this modified problem, $\bar{V}(u, S)$ is the maximum of the three terms in (2) and the additional term $\max_{j \in S_f}\{-c_j + E[\max\{X_j, \bar{V}(u, S)\}]\}$, where u denotes the best probed *slow* fading channel thus far. This equation can be explained as follows. The first three terms in (2) correspond to the rewards of actions involving slow fading channels: probing, retiring, or guessing, as described in (2). The additional term describes the expected reward of probing a fast fading channel because the user can either use them for transmission immediately after probing or not transmit on such channels, thus returning the system to state (u, S). The set of channels remains S because the user can probe fast fading channels multiple times. We have the following result.

Lemma 9: Consider any set of channels $S = S_s \cup S_f$. If $\pi^*(u, S) = \text{probe}(j)$ for some $j \in S_f$, then $\bar{V}(u, S) = \bar{a}_j$. This lemma can be proven as follows. Suppose at state (u, S), $\pi^*(u, S) = \text{probe}(j)$ for some $j \in S_f$. Then: $\bar{V}(u, S) = -c_j + E[\max\{\bar{V}(u, S), X_j\}]$. Comparing this to the definition of $\bar{a}_j$ in (11) yields $\bar{V}(u, S) = \bar{a}_j$, proving the result.

Comparing this lemma to (2), we have the following equivalence. Suppose we replace any fast fading channel $j \in S_f$ with a slow fading $k \in S_s$ such that $P(X_k = \bar{a}_j) = 1$. Because the reward X_k is constant, then an optimal strategy will never probe channel k, but might guess it and obtain a reward $\bar{a}_k$. If we replace all fast fading channels with slow fading channels, each with a constant reward $\bar{a}_j$, the value function for this modified problem is equivalent to the value function (2) for a system with only slow fading channels. Therefore, $\bar{V}(u, S) = V(u, S')$, where $S' = S_s \cup S'_s$ and S'_s denotes the set of slow fading channels created from fast fading ones. Thus, the original P1 formulation can solve a modified problem formulation that includes fast fading channels.

B. Guessing Constraints

As described in Section III-E, P1 can be extended to analyze constraints where only a subset of channels may be guessed. We summarize this extension in this subsection.

Recall that Lemma 7 describes how P1 can be decomposed into N subproblems. For each channel $j \in S$, we compute the optimal strategy π_j^* if no channel besides j can be guessed. Then, π^* is the best strategy among $\{\pi_j^*\}_{j \in S}$. If, in Problem 1, only a subset $\bar{S} \subseteq S$ of the channels can be guessed, then Lemma 7 can be modified by only determining π_j^* for each $j \in \bar{S}$ and then taking the best among these strategies.

The results of Section III can be generalized for a subset of guessable channels $\bar{S} \subseteq S$ as follows. For each $j \in \bar{S}$, define a_j, b_j as in (7) and (8). For each $j \in S - \bar{S}$, set $a_j = \bar{a}_j$, where $\bar{a}_j$ was defined in (11), and $b_j = 0$. Then, it can be shown that the results of Section III apply to this new scenario by using these new channel indices. Similarly, one can modify the algorithms of Section IV to use these new channel indices.

C. Guessing Penalty

In a practical system, not probing channels before transmission—i.e., guessing—could lead the user to transmit on a channel which is in fact busy, thereby causing interference to other users. To model a penalty associated with this potential scenario, we modify the problem formulation as follows.

For each channel j, we associate a guessing penalty Y_j that is a random variable that may depend on X_j. The user receives a reward $X_j - Y_j$ from guessing. For example, consider when $P(X_j \in \{0, 1\}) = 1$, i.e., the channel is either available or busy. To assign a penalty to the user for guessing on a busy channel, Y_j can be defined as follows: $P(Y_j = g | X_j = 0) = 1$, $P(Y_j = 0 | X_j = 1) = 1$, which models a positive guessing penalty g that is incurred if and only if the channel is busy. Note that $P(Y_j = 0) = 1$ implies no guessing penalty as in the original P1 formulation.

For general $\{Y_j\}$, incorporating this guessing penalty only adjusts the guessing reward in (2). It can be shown the results of Sections II, –IV hold with each channel having new indices $\tilde{a}_j, \tilde{b}_j$ that replace a_j, b_j defined in (7) and (8): $\tilde{a}_j = \min\{u : u \geq E[X_j] - E[Y_j], c_j \geq E[(X_j - u)^+]\}$, $\tilde{b}_j$ is the maximum u such that $u \leq E[X_j] - E[Y_j]$ and $c_j \geq E[(u - X_j)^+] + E[Y_j]$. Thus, the guessing penalty shifts the channel indices. Since the change is only in the channel indices, but not in the structural properties of the optimal strategy, the main results of Section IV continue to hold by using these new indices.

VI. STRATEGIES FOR P2

In this section, we present results on the optimal P2 strategy.

Similarly to Corollary 1, we can show that for any state $(\bar{T}, u, S)$, there exists an optimal strategy λ^* and constants $0 \leq b_{S, \bar{T}} \leq a_{S, \bar{T}} \leq 1$ such that

$$\lambda^*(\bar{T}, u, S) = \begin{cases} \text{retire}(u), & \text{if } u \geq a_{S, \bar{T}} \\ \text{probe}(j_u), j_u \in S, & \text{if } b_{S, \bar{T}} < u < a_{S, \bar{T}} \\ \text{guess}(j'), j' \in S, & \text{if } u < b_{S, \bar{T}}. \end{cases}$$

Thus, for each channel j, we can define indices $a_{j, \bar{T}}$ and $b_{j, \bar{T}}$ similar to (7) and (8). Even though these indices are now time-variant, which makes the analysis significantly more complex, we show a similarity between λ^* and π^*. For any $\bar{T}$, threshold $a_{j, \bar{T}}$ is the smallest u such that $\lambda(\bar{T}, u, j) = \text{retire}(u)$. Thus, $a_{j, \bar{T}}$ is the smallest u such that $u \geq E[X_j]$ and $u \geq E[\max(X_j, u)] \cdot (\bar{T} - \Delta / \bar{T})$. Index $b_{j, \bar{T}}$ can be calculated similarly. We have the following result.

Lemma 10: For any S and $\bar{T} > 0$: $a_{S, \bar{T}} = \max_{j \in S} a_{j, \bar{T}}$. Thus, the index $a_{S, \bar{T}}$ and the set of states where retirement is optimal, can be determined using only individual channel indices from time $\bar{T}$. Similar to P1, these indices do not depend on other indices $\{a_{j, \bar{T} - k\Delta}\}$, which simplifies computation.

In general, due to the time-varying nature of these indices, it becomes very difficult to determine the structure of the optimal strategy. However, the similarity in index properties between P1 and P2 policies leads to the following two-step look-ahead algorithm, similar to Algorithms γ and β. For any $\bar{T}$ and set of channels S, we first determine the two channels with the highest indices $a_{j, \bar{T}}$. Then, the optimal strategy is computed if we are forced to retire within two steps, similarly to Algorithm 2. We evaluate this strategy's performance in the next section.

VII. NUMERICAL RESULTS

In this section, we examine the performance of the proposed algorithms under a practical class of channel models.

For both P1 and P2 policies, we consider a two-state channel model where, for each channel $P(X_j = r_j) = p_j = 1 - P(X_j = 0)$ for some $r_j > 0$. This models, for example, when channels are either on with available data rate r_j or off.[8] Under this setting, the set of information states is $\cup_j \cup_S (r_j, S)$.

We chose parameters r_j, p_j, c_j for each channel as follows. First, r_j and p_j were modeled as independent random variables,

---

[8][9] has considered optimal P1 strategies for two-state channels, each with identical data rate. When the parameters r_j differ between different channels, it can be shown the strategies of [9] are not necessarily optimal.
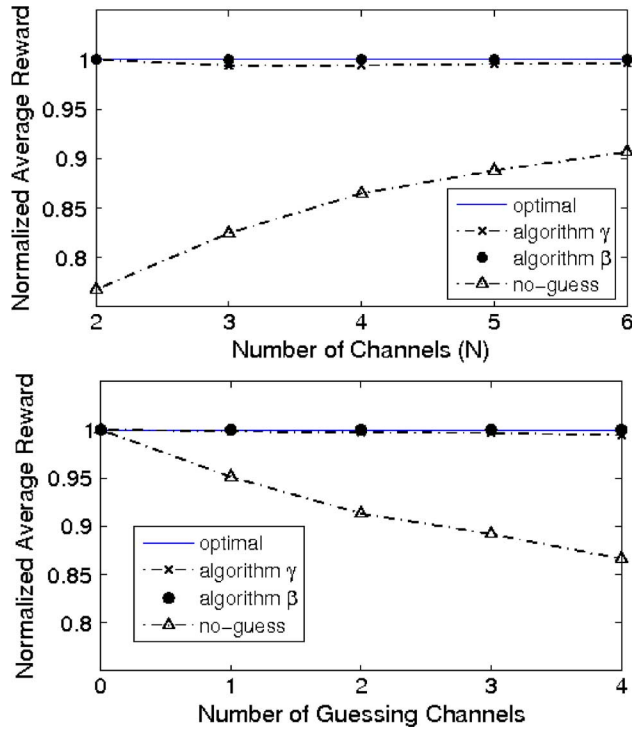
Fig. 3. (Top) Average performance of optimal P1 strategy, algorithms γ and β of Section IV-B, and the optimal strategy without guessing. Rewards are normalized by the average reward of the optimal strategy. (Bottom) Average performance of these strategies for a four-channel system where the number of channels that can be guessed varies between 0 and 4.

Fig. 4. (Top) Performance of optimal P2 strategy and a two-step look-ahead P2 policy as the number of channels (N) varies. (Bottom) Performance of the two-step look-ahead, one-channel, two-channel, and four-channel algorithms of Section VII, when all channels are statistically identical with $P(X_j \le x) = x^\alpha$ and different probing costs $c_j = 0.01j$.

uniformly distributed in the interval[9] (0,1). After the realization of these parameters was chosen, the channel cost c_j was uniformly chosen in the interval[10] $(0, p_j(1 - p_j)r_j + 0.01)$.

For each realization of r_j, p_j, c_j, the expected rewards of the following strategies were computed for P1: the optimal strategy (determined via dynamic programming), algorithms γ and β from Section IV-B, and the optimal algorithm if guessing is not allowed (no-guess), as described in Section III-C and Theorem 2. A total of 10^3 random realizations are generated and then averaged for each value N. Fig. 3 (top) depicts the performance of these strategies as the number of channels N varies. The average rewards of these strategies are normalized by dividing the average reward of the optimal strategy. We note that both Algorithm γ and β perform very close to the optimal, with β performing slightly better. This is because Algorithm γ and β are optimal when Case 3a from Theorem 1 holds. In general, this case holds for most values of p_j, r_j, and c_j. When Case 3b or 3c of Theorem 1 holds, Algorithms γ and β only differ with the optimal algorithm in the parameter d_S. Thus, in general they are very close numerically to the optimal strategy.

As mentioned in Section II, it may be the case that some regulatory spectrum policies do not allow all channels to be guessed.

Fig. 3 (bottom) analyzes the performance when $N = 4$ and only a subset $\bar{S} \subseteq S$ of these channels can be guessed. For this case, we modify Algorithm γ as follows. If $j \notin \bar{S}$, we set $a_j = \bar{a}_j$ as given by (11), and set $b_j = 0$. For $j \in \bar{S}$, the indices remain unchanged. These changes remove guess(j) as a possible action. For Algorithm β, we replace S in its definition with $\bar{S}$. The relative performance between the optimal strategy, γ, and β does not change as N_g, the number of channels that can be guessed, varies. On the other hand, by definition the no-guess strategy is optimal for $N_g = 0$ but as expected its average reward decreases as N_g increases.

Similarly, Fig. 4 (top) analyzes the optimal strategy and a two-step look-ahead algorithm (similar to γ, as described in the previous section) for P2. As can be seen, the two-step look-ahead algorithm performs similarly to the optimal strategy.

Fig. 4 (bottom) analyzes performance when channels are statistically identical with cdf $P(X_j \le x) = x^\alpha$ for all j and with different probing costs $c_j = 0.01j$. Performance of the two-step look-ahead algorithm, which is optimal from Theorem 4, is compared in the figure to the following algorithms. The *one-channel* algorithm does not probe and simply transmits using the "best" channel (lowest cost). Comparing the two-step look-ahead algorithm to this strategy gives an indication of the gain from using probing. The *two-channel* (*four-channel*) algorithm depicted in the figure probes the best two (four) channels and

---

[9]The upper bound on r_j of 1 is chosen for simplicity; it could be any positive M, which simply scales the reward and the cost simultaneously.

[10]This upperbound on c_j ensures that some channels will be probed, as it can be shown that if $c_j > p_j(1 - p_j)r_j$, then channel j should never be probed and only guessed. The additional 0.01 to $p_j(1 - p_j)r_j$ is to ensure that some channel will be guessed, but the value 0.01 is an arbitrary choice.
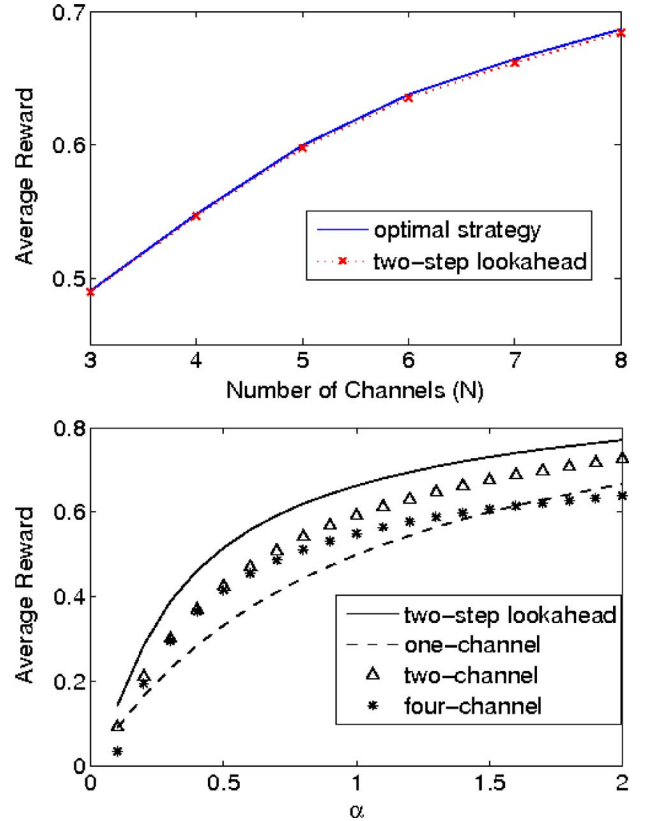
then uses the best channel (among those probed) for transmission. Thus, the results indicate the gain from using a more efficient probing algorithm over simple heuristics.

In all cases, these results confirm that the two-step look-ahead policy performs very similarly to the optimal strategy, even though it has much less computational overhead. From the dynamic programming formulation given in (2), even when the channel rewards are discrete random variables, computing the optimal strategy at state (u, S) still requires us to take combinations of all subsets of S. By comparison, the two-step look-ahead policy only considers the best two channels in S.

VIII. CONCLUSION

In this paper, we analyzed the problem of channel probing and transmission scheduling in wireless multichannel systems. We derived some key properties of optimal channel probing strategies and showed that the optimal policy has a threshold structure and can only take one of a few forms. Using these properties, we proposed two channel probing algorithms that we showed are optimal for some cases of practical interest, including statistically identical channels, a few nonidentical channels, and a large number of nonidentical channels. These algorithms were also shown to perform very well compared to the optimal strategy under a practical class of channel models.

APPENDIX

A. Proof of Theorem 2

The proof that $\hat{\pi}(u, S) = \text{retire}(u)$ for all $u \geq \max_{j \in S}\{\bar{a}_j\}$ uses the same steps as proving Lemma 3 and is thus omitted for brevity. For $u < \max_{j \in S}\{\bar{a}_j\}$, we prove the result by induction on the cardinality of S.

Induction Basis: Suppose $|S| = 1$. Let $S = \{j\}$. Then, $\hat{\pi}(u, j) = \text{probe}(j)$ follows from the definition of $\bar{a}_j$.

Induction Hypothesis: Let $|S| = n \geq 2$. Suppose the result holds for all $\tilde{S} \subseteq \Omega$ such that $|\tilde{S}| < n$. We proceed in steps to show $\hat{\pi}(u, S) = \text{probe}(1)$ for all $u < \max_{j \in S} a_j$.

Step 1 (Show $\hat{\pi}(u, S) = \text{probe}(1)$, for all $\max_{j \in S-R} \bar{a}_j \leq u < \bar{a}_1$): First, we show $V(u, S) \neq \text{probe}(j)$ for all all $j \in S - R$ by contradiction. Suppose there exists some $m > 1$, $m \in S - R$, such that $\hat{\pi}(u, S) = \text{probe}(m)$, where u satisfies $\bar{a}_m \leq u < \bar{a}_1$. By following the same exact steps as (16)–(17) in [13], we arrive at a contradiction.

From Lemma 3, retiring cannot be optimal (removing guessing does not change this result).

Therefore, $\hat{\pi}(u, S) = \text{probe}(m)$ for some $m \in R$. We show the remainder of the proof by contradiction. Suppose $\hat{\pi}(u, S) = \text{probe}(m)$ for some $m \in R$, $m > 1$. Note $\bar{a}_m = \bar{a}_1$ by definition of R. Let $\hat{V}_m(u, S)$ denote the expected reward of probing m first; this probe incurs cost c_m and then by the induction hypothesis, at state $S - m$, we probe 1 if $X_m < \bar{a}_1$; otherwise, we retire. Similarly, $\hat{V}_1(u, S)$ is the expected reward of probing 1 first, and then probing m in the second step if $X_1 < \bar{a}_m$. Since $\hat{\pi}(u, S) = \text{probe}(m)$, then $\hat{V}_m(u, S) - \hat{V}_1(u, S) > 0$. This inequality gives: $-c_m + E[X_m I_{\{X_m \geq \bar{a}_1\}}] - c_1 P(X_m < \bar{a}_1) > -c_1 - c_m P(X_1 < \bar{a}_1) + E[X_1 I_{\{X_1 \geq \bar{a}_1\}}]$. Rearranging yields an inequality that contradicts the definition of 1 in Theorem 2,

and thus contradicts $\hat{\pi}(u, S) = \text{probe}(m)$. We have therefore shown $\hat{\pi}(u, S) = \text{probe}(1)$ for all $\max_{j \in S-R} \leq u < \bar{a}_1$.

Step 2 (Show $\hat{\pi}(u, S) = \text{probe}(1)$ for all $u < \max_{j \in S-R} \bar{a}_j$): Again, let $\hat{V}_m(u, S)$ denote the expected reward of probing m first at state (u, S). By the induction hypothesis, if $X_m \geq \bar{a}_1$, then we retire; otherwise, we continue. Letting $\hat{V}_1(u, S)$ denote the expected reward of probing 1 first, it suffices to show for all $u \leq \bar{a}_1$ and m, $\hat{V}_1(u, S) - \hat{V}_m(u, S)$ does not depend on u. If this holds, then $\hat{\pi}(u, S) = \text{probe}(1)$ for all $u \leq \bar{a}_1$ since we have already shown $\hat{\pi}(u, S) = \text{probe}(1)$ for all $\max_{j \in S-R} \bar{a}_j \leq u < \bar{a}_1$.

By the induction hypothesis, $\hat{V}_2(u, S) = -c_2 + E[X_2 I_{\{B\}}] - P(B^c)c_1 + E[\max(X_1, X_2)I_{\{A \cap B^c\}}] + E[\hat{V}(\max(u, X_1, X_2), S - \{1, 2\})I_{\{A^c \cap B^c\}}]$, where $\hat{V}(\cdot, \cdot)$ is the value function for Problem 2, defined similarly to (2), A is the event $\{\max(X_1, X_2) \geq \bar{a}_{n-2}\}$, A^c is its complement, and B is the event $\{X_2 \geq \bar{a}_1\}$. $\hat{V}_1(u, S)$ can be calculated similarly by interchanging 1 and 2, and replacing B with the event $\{X_n \geq \bar{a}_2\}$. We see that $\hat{V}_1(u, S) - \hat{V}_2(u, S)$ is invariant to u (only the term with $\hat{V}(\cdot, \cdot)$ contains u, and this cancels out during the subtraction by conditioning the expectations on events B and D). Similar steps can be taken for other $m \neq 1$, by calculating $\hat{V}_m(u, S)$ until only channels $\{1, \cdots, m-1\}$ are left, and showing that $\hat{V}_1(u, S) - \hat{V}_m(u, S)$ does not change with u. Therefore, $\hat{\pi}(u, S) = \text{probe}(1)$ for all $u < \max_{j \in S-R}\{a_j\}$, and we have shown $\hat{\pi}(u, S) = \text{probe}(1)$ for all $u < \bar{a}_1$. ∎

B. Proof of Lemma 8

From Lemma 3, $\pi^*(u, S) = \text{retire}(u)$ if and only if $u \geq \bar{a}_1$. We thus need to prove $\pi^*(u, S) = \text{probe}(1)$ for $u < \bar{a}_1$. From Corollary 2, we know that $\pi^*(u, S) \neq \text{guess}(j)$. Thus, it only remains to determine which channel to probe for $u < \bar{a}_1$. Given a fixed channel $j \in S$, we prove the result by backward induction on the cardinality of S.

Induction Basis: Suppose $|S| = 2$. Then $S = 1 \cup 2$ where $\bar{a}_1 > a_2$ and $2 = j$ from the conditions of the lemma.

Lemma 5 implies that $\pi^*(u, S) = \text{probe}(1)$ for $b_j < u < \bar{a}_1$. It can be shown similar to the proof in Appendix A that for all $b_j \leq u < \bar{a}_1$, the difference in expected reward in probe(1) and probe(2) is invariant to u. Thus, $\pi^*(b_j, S) = \text{probe}(1)$. Meanwhile, the expected reward of probe(1) does not depend on u if $0 \leq u \leq b_j$. Since $V(\cdot, S)$ is nondecreasing, then $\pi^*(u, S) = \text{probe}(1)$ for all $0 \leq u \leq b_j$.

Induction Hypothesis: Suppose $|S| = n$ for some $n > 2$ and the lemma holds for all $\bar{S}$ such that $|\bar{S}| < n$. Sort S according to Algorithm 1, i.e., $|S| = \{1, 2, \cdots, n\}$. From the conditions stated in the lemma, $j = l$ for some $l \geq 2$, where $\bar{a}_1 > a_j$. We prove the induction hypothesis by further backward induction on j, i.e., we first prove the result for $j = N$ and then show this implies the result for $j = N - 1$, etc.

Step 1 (Prove the Result for $j = N$): Suppose $j = N$, which implies $b_j \leq a_j \leq \bar{a}_l$ for all $l < N$, and consider any $b_N < u < \bar{a}_1$. From Lemma 4, $\pi^*(u, S) = \text{probe}(1)$. We can show similarly to the proof of Theorem 2 in Appendix A that the difference in expected reward in probe(1) and probe(l) is invariant to u for $b_j \leq u < \bar{a}_1$, implying $\pi^*(b_j, S) = \text{probe}(1)$.

From the induction hypothesis, the optimal strategy after probing 1 is to retire if $X_1 > \bar{a}_2$; otherwise, probe(2). Then, retire if $\max\{X_1, X_2\} \geq \bar{a}_3$; otherwise, probe(3) and continue until the transmitter retires or N is the last channel, in which case the optimal strategy is given by Corollary 1 with $u = \max\{1, \cdots, N-1\}$ and $S = \{j\}$. Note the optimal expected reward $V(u, S)$ is constant for all $u \leq b_j$, because the transmitter never retires and collects b_j, since action guess(j) yields higher reward. Thus, $V(\cdot, S)$ being nondecreasing and $\pi^*(b_j, S) = \text{probe}(1)$ collectively imply $\pi^*(u, S) = \text{probe}(1)$ for all $0 \leq u \leq b_j$. This proves the result when $j = N$.

Step 2 (Prove the Result for $j = l$, $1 < l < N$): Now, suppose $j = l$ for some $1 < l < N$ and the hypothesis holds for all values of j in $\{l+1, \cdots, N\}$. We prove $\pi^*(u, S) = \text{probe}(1)$ by contradiction.

First, we prove $\pi^*(u, S) \neq \text{probe}(l)$. A strategy that first probes l never guesses, and from Theorem 2 cannot do better than first probing 1. Thus, $\pi^*(u, S) \neq \text{probe}(l)$.

We now prove by contradiction that $\pi^*(u, S) \neq \text{probe}(i)$ for some $l < i \leq N$. Suppose $\pi^*(u, S) = \text{probe}(i)$ for $l < i \leq N$.

Case 1: $P(X_i \leq a_j) = 1$. From the induction hypothesis, after probing i the optimal strategy probes 1 if $\max\{X_i, u\} < \bar{a}_1$, otherwise retires. Because $u < \bar{a}_1$ and $P(X_i \leq a_j < \bar{a}_1) = 1$, the optimal strategy obtains expected reward: $-c_i - c_1 + E[V(\max\{X_1, X_i, u\}, S - 1 - i)]$. Now, consider the modified strategy that acts similarly to the optimal strategy, except that it exchanges the roles of 1 and i. Its expected reward is: $-c_1 + P(A^c)\{-c_i\} + E[V(\max\{X_1, X_i, u\}, S-1-i)]$, where A is the event that $\{X_1 \geq \bar{a}_1\}$ and A^c denotes its complement. Since $P(A^c) > 0$ from the definition of $\bar{a}_1$ and $c_1 > 0$, the modified strategy obtains higher expected reward than the optimal one. This contradicts the definition of optimal strategy, proving case 1.

Case 2: $P(X_i > a_j) > 0$. In this case, $E[(X_i - a_j)^+] > 0$. For any $1 \leq k \leq N$, let ϕ_k denote the expected reward of probing k and then proceeding optimally. As assumed, $\phi_l \geq \phi_k$ for all $k \neq l$. Since $\bar{a}_i \leq a_j$, then $c_i \geq E[(X_i - a_j)^+]$.

We modify the original scenario (called scenario 1) to generate a modified problem (scenario 2). Under scenario 2, all channels have the same rewards and probing costs as scenario 1, *except* for channel i, whose probing cost (denoted by c_i^t) is decreased to satisfy $E[(X_i - \bar{a}_1)^+] < c_i^t < E[(X_i - a_j)^+]$, where the inequalities are strict because $\bar{a}_i > a_j$ and $P(X_i > a_j) > 0$ as assumed. Let $\bar{a}_i^t$ denote the new index of channel i under scenario 2. We see that $\bar{a}_i^t > a_j \geq \bar{a}_i$. Thus, because $\bar{a}_k > a_l = a_j$ for all $k \leq l$, we can apply the induction hypothesis to show $\pi^*(u, S) = \text{probe}(1)$ for scenario 2.

Now, we prove a contradiction, first for $i < l$. Let ϕ_l^t denote the expected reward of probing k under scenario 2, and then proceeding according to the optimal strategy. It can be seen that $\phi_l^t - \phi_l = c_i - c_i^t \geq \phi_k^t - \phi_k, k \neq l$. Because $\phi_l \geq \phi_k^t$, then $\phi_l^t \geq \phi_k^t$. Thus, probe(l) is also optimal for scenario 2. However, this contradicts $\pi^*(u, S) = \text{probe}(1)$ as shown earlier. Thus we have shown $\pi^*(u, S) \neq \text{probe}(i)$ for $i < l$.

Finally, we show contradiction for $1 < i < l$. Suppose $\pi^*(u, S) = \text{probe}(i)$ for some $1 < i < l$. The induction hypothesis gives the optimal strategy after probing i. We can use a similar proof to Theorem 2 in Appendix A to show that this

strategy's expected reward is less than reward obtained by first probing 1. Thus, $\pi^*(u, S) = \text{probe}(1)$ for these u. ∎

C. Proof of Theorem 5

For the INC problem, the available channel types are not changing. Let $V(u)$ and $\pi(u)$ denote the maximum expected reward and a strategy, respectively, given the best probed channel has value u. We prove the result for different u.

Case 1 ($u \geq a_1$): We first prove that retire(u) is optimal if $u \geq a_1$. Let V_n^q denote the expected reward after n time-steps of a strategy q that retires if $u \geq a > a_1$. From Lemma 3, there exists a strategy π that retires if $u \geq a_1$ and $V_n^\pi \geq V_n^q$. Both V_n^q and V_n^π converge because they are monotonically increasing in n and bounded above by M. Thus, $\lim_{n \to \infty} V_n^\pi \geq \lim_{n \to \infty} V_n^q$. However, the left- and right-hand sides of this inequality are the expected rewards of a strategy that retires if $u \geq a_1$ and $u > a$, respectively. Since this holds for all $a > a_1$, we have shown there exists an optimal strategy that retires if $u \geq a_1$.

Case 2 ($u < a_1$, $a_1 = b_1$): Suppose $a_1 = b_1$, and consider the strategy π such that: $\pi(u) = \text{guess}(1)$ if $u \leq a_1$, otherwise $\pi(u) = \text{retire}(u)$. From Corollary 2, this strategy is optimal for a finite number of channels. Let V_n^q denote the expected reward after n time-steps of a strategy that probes a channel instead of guessing. Since V_n^q is monotonically increasing in n and bounded above by M, it converges as $n \to \infty$. Meanwhile, from Corollary 2 we have $V_n^q \leq E[X_1]$ for all n. Thus, $\lim_{n \to \infty} V_n^q \leq E[X_1]$, which says π is optimal.

Case 3 ($u < a_1$, $b_1 < a_1$): When $a_1 \neq b_1$, proving that probe(1) is optimal uses the same steps as proving Lemma 6 and Theorem 2. The proof is omitted for brevity. ∎

D. Proof of Corollary 3

Parts 1) and 2) of Theorem 1 imply $V(u, S)$ for $u \geq a_1$ and $a_1 > u \geq d_S$. Thus, we only need to prove the corollary for $u < \alpha$. We use induction on the cardinality of S.

Induction Basis: Consider when $S = \{j\}$, i.e. a single channel. From (7) and (8), the corollary holds with $\alpha = b_S$.

Induction Hypothesis: Fix any S, $|S| > 1$ and suppose corollary holds for all $\tilde{S}$, $\tilde{S} \subset S$. We prove the corollary holds for all three possibilities of $\pi^*(u, S)$ given by Theorem 1.

Step One: Suppose $\pi^*(u, S) = \text{probe}(k)$ for some $k > 1$. Lemma 5 implies that for $u < d_S$: $V(u, S) = -c_k + E[V(\max(X_k, u), S - k)]$. Thus, $V(u, S) = V(0, S)$ for all $u < d_S$, which implies $V(u, S)$ is a constant for all $u < d_S$. It can be shown $V(\cdot, S)$ is continuous (for fixed S), which implies $\lim_{u \to d_S^-} V(u, S) = V(d_S, S)$. However, $V(d_S, S)$ is the expected reward of probing 1 first, as given in the corollary; thus, $V(u, S) = V(d_S, S) = c_1 + E[V(\max\{X_1, d_S\}, S - 1)]$ for all $u < d_S$.

Step Two: If $\pi^*(u, S) = \text{guess}(1)$ for all $u < d_S$, then $V(u, S) = E[X_1]$, which is also a constant function with respect to u. Therefore, we similarly have $V(u, S) = V(d_S, S)$.

Step Three: Suppose $\pi^*(u, S) = \text{probe}(1)$ for all $u < d_S$. Then, for $u < d_S$: $V(u, S) = -c_1 + E[V(\max\{X_j, u\}, S - \{1\})] = -c_1 + E[V(\max\{X_j, d_S\}, S - \{1\})] = V(d_S, S)$. The second equality holds because $V(u, S - \{1\}) = V(d_S, S - \{1\})$ for all $u < d_S$ by the induction hypothesis. ∎

E. Proof of Lemma 10

Step 1: We first show that for any $S, \Delta, \bar{T}, \tilde{u}, u$

$$H(\bar{T} - \Delta, \tilde{u}, S) - H(\bar{T}, u, S) \le (\bar{T} - \Delta)\tilde{u} - \bar{T}u \qquad (12)$$

where $\Delta > 0$, $\tilde{u} \ge u$ and $\bar{T} - \Delta > 0$. We prove (12) for the three possible values of $H(\bar{T} - \Delta, \tilde{u}, S)$ in (4).

Case 1: If $H(\bar{T} - \Delta, \tilde{u}, S) = (\bar{T} - \Delta)\tilde{u}$, then (12) follows from $H(\bar{T}, u, S) \ge \bar{T}u$, as given in (4).

Case 2: If $H(\bar{T} - \Delta, \tilde{u}, S) = (\bar{T} - \Delta)E[X_j]$ for some $j \in S$, then $E[X_j] \ge \tilde{u}$. From (4), $H(\bar{T}, u, S) \ge \bar{T}E[X_j]$. Therefore, $H(\bar{T} - \Delta, \tilde{u}, S) - H(\bar{T}, u, S) \le (\bar{T} - \Delta)E[X_j] - \bar{T}E[X_j] = \Delta E[X_j] \le -\Delta\tilde{u} \le \bar{T}\tilde{u} - \Delta\tilde{u} - \bar{T}u = (\bar{T} - \Delta)\tilde{u} - \bar{T}u$, where the last inequality follows from $\tilde{u} \ge u$. Thus, (12) holds.

Case 3: If we have $H(\bar{T} - \Delta, \tilde{u}, S) = E[H(\bar{T} - 2\Delta, \max\{X_j, \tilde{u}\}, S - j)]$, then $H(\bar{T}, u, S) \ge E[H(\bar{T} - \Delta, \max\{X_j, u\}, S - j)]$. Thus, $H(\bar{T} - \Delta, \tilde{u}, S) - H(\bar{T}, u, S) \le E[H(\bar{T} - 2 \cdot \Delta, \max\{X_j, \tilde{u}\}, S - j)] - E[H(\bar{T} - \Delta, \max\{X_j, u\}, S - j)] = (*)$. Conditioning on $\{X_j \ge \tilde{u}\}$ and using induction, $(*) \le H(\bar{T} - 2\Delta, \tilde{u}, S - j) - H(\bar{T} - \Delta, u, S - j) \le (\bar{T} - 2\Delta)\tilde{u} - (\bar{T} - \Delta) = (\bar{T} - \Delta)\tilde{u} - \bar{T}u - \Delta(\tilde{u} - u) \le (\bar{T} - \Delta)\tilde{u} - \bar{T}u$.

Step 2: Using (12), we prove the lemma by contradiction on two cases. Let j be any channel achieving $\arg\max_{k \in S} a_{k,\bar{T}}$.

Case 1 ($a_{S,\bar{T}} < a_{j,\bar{T}}$): Fix any $a_{S,\bar{T}} \le u < a_{j,\bar{T}}$; thus, $\lambda^*(\bar{T}, u, S) = \text{retire}(u)$ and $H(\bar{T}, u, S) = \bar{T}u$. At the same time, $u < a_{j,T}$ implies: $H(\bar{T}, u, S) \ge H(\bar{T}, u, j) > \bar{T}u$, which contradicts the assumption $a_{S,\bar{T}} \le u < a_{j,\bar{T}}$.

Case 2 ($a_{S,\bar{T}} > a_{j,\bar{T}}$): Fix any $a_{S,\bar{T}} > u \ge a_{j,\bar{T}}$. Suppose $\lambda(\bar{T}, u, S) = \text{probe}(1)$ for some $1 \in S$. We know $\bar{T}u > E[H(\bar{T} - \Delta, \max\{X_1, u\}, \emptyset)] = (\bar{T} - \Delta)E[\max\{X_1, u\}]$. On the other hand, $E[H(\bar{T} - \Delta, \max\{X_1, u\}, S - \{1\})] \ge H(\bar{T}, u, S) \ge H(\bar{T}, u, S - \{1\})$. Combining these equations gives $\bar{T}u + E[H(\bar{T} - \Delta, \max\{X_1, u\}, S - \{1\})] \ge (\bar{T} - \Delta)E[\max\{X_1, u\}] + H(\bar{T}, u, S - 1)$, which implies $E[H(\bar{T} - \Delta, \max\{X_1, u\}, S - \{1\})] - H(\bar{T}, u, S - \{1\}) \ge (\bar{T} - \Delta)E[\max\{X_1, u\}] - \bar{T}u$, a contradiction to (12). If $\lambda(\bar{T}, u, S) = \text{guess}(1)$, then $E[X_1] > u$, and thus $a_{j,\bar{T}} > u$. Therefore, we again have a contradiction to $u \ge a_{j,\bar{T}}$. ∎

ACKNOWLEDGMENT

REFERENCES

[1] X. Liu, E. Chong, and N. Shroff, "Transmission scheduling for efficient wireless network utilization," in *Proc. 20th IEEE INFOCOM*, Anchorage, AK, 2001, pp. 776–785.

[2] X. Quin and R. Berry, "Exploiting multiuser diversity for medium access control in wireless networks," in *Proc. 22nd IEEE INFOCOM*, San Francisco, CA, 2003, pp. 1084–1094.

[3] S. Guha, K. Munagala, and S. Sarkar, "Jointly optimal transmission and probing strategies for multichannel wireless systems," in *Proc. CISS*, Princeton, NJ, Mar. 2006, pp. 955–960.

[4] J. Kennedy and M. Sullivan, "Direction finding and "smart antennas" using software radio architectures," *IEEE Commun. Mag.*, vol. 33, no. 5, pp. 62–68, May 1995.

[5] Y. Chen, Q. Zhao, and A. Swami, "Joint design and separation principle for opportunistic spectrum access," in *Proc. IEEE Asilomar Conf. Signals, Syst., Comput.*, Nov. 2006, pp. 696–700.

[6] Z. Ji, Y. Yang, J. Zhou, M. Takai, and R. Bagrodia, "Exploiting medium access diversity in rate adaptive wireless LANs," in *Proc. 10th ACM MobiCom*, Philadelphia, PA, Sep. 2004, pp. 345–359.

[7] A. Sabharwal, A. Khoshnevis, and E. Knightly, "Opportunistic spectral usage: Bounds and a multi-band CSMA/CA protocol," *IEEE/ACM Trans. Netw.*, vol. 15, no. 3, pp. 533–545, Jun. 2007.

[8] G. Holland, N. Vaidya, and P. Bahl, "A rate-adaptive MAC protocol for multi-hop wireless networks," in *Proc. 7th ACM MobiCom*, Rome, Italy, 2001, pp. 236–251.

[9] S. Guha, K. Munagala, and S. Sarkar, "Optimizing transmission rate in wireless channels using adaptive probes," presented at the ACM Sigmetrics/Perf. Conf., Saint-Malo, France, 2001.

[10] S. Guha, K. Munagala, and S. Sarkar, "Approximation schemes for information acquisition and exploitation in multichannel wireless networks," in *Proc. 44th Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, Sep. 2006, pp. 85–90.

[11] J. Heiskala and J. Terry, *OFDM Wireless LANs: A Theoretical and Practical Guide*. Indianapolis, IN: SAMS, 2001.

[12] P. R. Kumar and P. Karaiya, *Stochastic Systems: Estimation, Identification, and Adaptive Control*. Englewood Cliffs, NJ: Prentice-Hall, 1986.

[13] N. Chang and M. Liu, "Optimal channel probing and transmission scheduling for opportunistic spectrum access," in *Proc. 13th ACM MobiCom*, Montreal, Canada, Sep. 2007, pp. 27–38.

[14] D. Zheng, W. Ge, and J. Zhang, "Distributed opportunistic scheduling for ad-hoc communications: An optimal stopping approach," in *Proc. 8th ACM MobiHoc*, Montreal, Canada, Sep. 2007, pp. 1–10.
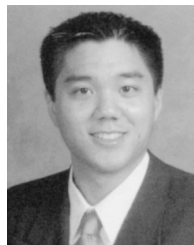
Nicholas B. Chang (S'05) received the B.S.E. degree (*magna cum laude*) in electrical engineering from Princeton University, Princeton, NJ, in 2002, and the M.S.E degree in electrical engineering: systems, M.S. degree in mathematics, and Ph.D. degree in electrical engineering: systems from the University of Michigan, Ann Arbor, in 2004, 2005, and 2007, respectively.

He is currently a Staff Member at MIT Lincoln Laboratory, Lexington, MA. His research interests include communication networks, wireless communication, stochastic control, stochastic resource allocation, and algorithms.

Dr. Chang is a Member of Tau Beta Pi and Sigma Xi, the Scientific Research Society. He is a recipient of the 2005–2006 MIT Lincoln Laboratory Graduate Fellowship.
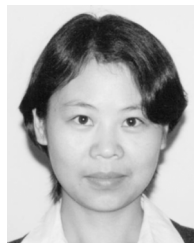
Mingyan Liu (M'00) received the B.Sc. degree in electrical engineering from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 1995, and the M.Sc. degree in systems engineering and Ph.D. degree in electrical engineering from the University of Maryland, College Park, in 1997 and 2000, respectively.

She joined the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, in September 2000, where she is currently an Associate Professor. Her research interests are in performance modeling, analysis, energy-efficiency and resource allocation issues in wireless mobile ad hoc networks, wireless sensor networks, and terrestrial satellite hybrid networks.

Dr. Liu is the recipient of the 2002 NSF CAREER Award and the University of Michigan Elizabeth C. Crosby Research Award in 2003. She serves on the Editorial Board of the IEEE/ACM TRANSACTIONS ON NETWORKING and the IEEE TRANSACTIONS ON MOBILE COMPUTING.