Optimal Channel Probing and Transmission Scheduling in a Multichannel System

Nicholas B. Chang and Mingyan Liu
Department of Electrical Engineering and Computer Science
University of Michigan, Ann Arbor, MI 48109-2122
Email: {changn,mingyan}@eecs.umich.edu

***Abstract*— In this study we consider the problem of optimal channel selection policies for a transmitter in a multichannel wireless system where a channel can be in one of multiple states. Each channel state is associated with a probability of transmission success. In such systems, the sender typically has partial information concerning the channel states, but can deduct more information by probing individual channels, e.g. by sending control packets in the channels. On the other hand, while probing can help the sender determine the best channel to use, it consumes valuable resources. The main goal of this work is to derive optimal strategies for determining which channels to probe (in what sequence) and which channel to use for transmission. Our primary interest in this study is to derive key structural properties of the optimal strategy. In particular, we show that it exhibits a threshold structure and can be described by an index policy. We further show that the optimal strategy can only take on one of three structural forms. In addition, we explicitly derive optimal strategies for some practical channel models, and provide conditions under which some commonly used strategies are optimal. Using these properties, we present a two step look-ahead policy and show that it is optimal for a number of special cases of practical interest.**

***Index Terms*— channel probing, scheduling, stochastic control, wireless ad hoc networks, dynamic programming,**

I. Introduction

Effective transmission over wireless channels is a key component of wireless communication, and to achieve this one must address a number of issues specific to the wireless environment. One such key challenge is the time-varying nature of the wireless channel due to multi-path fading. This channel fading can be caused by factors such as mobility, interference, and environmental objects. The unreliability caused by such fading must be accounted for when designing robust transmission strategies. Recent works such as [1], [2] have studied opportunistically transmitting when channel conditions are better in order to exploit these channel fluctuations in time.

At the same time, many wireless systems also provide nodes 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, 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. Many IEEE 802.11 technologies also propose using multiple frequencies. In addition, software radio (SDR) systems [4] may provide nodes with multiple channels (e.g. tunable frequency bands and modulation techniques) by means of a programmable hardware which is controlled by software. Such systems are particularly important for applications such as military where flexible and interoperable communications are desired.

All of these systems share the common feature that the transmitter is generally supplied with more channels than needed for a single-transmission. Thus, it is possible for nodes to exploit the time-varying channels by opportunistically choosing the best one to use for transmission [5], [6], [7]. This may be viewed as an exploitation of channel fluctuations in space (i.e., across different channels), and is akin to the idea of multiuser diversity [8].

In order to utilize such channel variation, it is desirable for the transmitter and/or receiver to periodically obtain information on channel quality. One method of accomplishing this in a distributed manner is to allow nodes to probe channels by having the transmitter send control packets. Upon receiving a control packet, the receiver sends back a response packet to the sender that may indicate the channel quality. For example, recent works such as [5], [9] have proposed enhancing or exploiting the multi-rate capabilities of the IEEE 802.11 RTS/CTS handshake mechanism. [9] proposes a protocol called Receiver Based Auto Rate (RBAR) in which the receivers use physical-layer analysis of received RTS messages to find out the maximum possible transmission rate to achieve a specific bit error rate. The receiver controls the sender's transmission rate by inserting the maximum possible transmission rate into CTS message.

This probing process can help nodes obtain more information about channel quality and therefore make better decisions concerning which channel to use for transmission. However, channel measurement and estimation also incur overhead and consume valuable network resources. In particular, the exchange of control packets decreases the amount of time available to send actual data and consumes energy. In addition, sending control packets can also prevent other users from simultaneously using the channel. Thus, channel probing must be done efficiently in order to balance the trade-off between obtaining useful channel information and consuming valuable resources.

In this work, we study optimal strategies for a joint channel probing and transmission problem. Specifically, we consider

This work is supported by NSF award ANI-0238035 and a 2005-2006 MIT Lincoln Laboratory Fellowship.

a transmitter with multiple channels with known state distributions. The transmitter can sequentially probe any channel with channel-dependent costs. The problem 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], [10], [11], [6], [7], [5]. The commonality and differences between our study and the previous work are highlighted below within the context of our main contributions.

The main contributions of this work are as follows.

Firstly, we derive key properties of optimal strategies for the problem outlined above, and show that the optimal strategy has a threshold structure. We use this structure to derive necessary and sufficient conditions for certain strategies to be optimal. In contrast to [3], [10], [11], we do not restrict the channels to take a finite number of states; our work applies to both the case where the number of channel states is finite, and the case where they can take an infinite number of states. This generalization is important as many next generation physical layer technologies such as MIMO and Adaptive-Bit-Loading OFDM [12] are aiming to provide continuous range of data rates that can be adjusted according to channel quality. In this sense our work is more general than previous work.

Secondly, we derive explicitly the optimal channel probing strategy for a number of practical special cases. One of these special cases is when the channels are statistically identical with possibly different probing costs, and our results provide optimal strategies for *arbitrary* channel distributions and probing costs. In [6], [7], [5], a variant of this problem was studied when the channels are statistically identical. In particular, [6], [7] analyzed the situation with identically distributed channels, only allowed channels to be used immediately after probing (i.e. no recall of past channel probes), and did not allow unprobed channels to be used for transmission. In addition, their work restricts the transmitter to probe the channels in a particular order, while ours determines the optimal order for the transmitter. Meanwhile, [5] assumes independent Rayleigh fading channels and because all channels are independent and identically distributed, do not focus on which channels should be probed and in what order. In our work, similar to [3], [10], we allow the situation where channels are not statistically identical, and we provide results for a general class of channel distributions. Our work can thus be seen as a generalization of previous studies in [3], [10], [11], [6], [7], [5].

Finally, based on the key properties of optimal strategies, we propose strategies that perform well for arbitrary number of channels and arbitrary number of channel states (finite or infinite). To the best of our knowledge, this is the first channel probing algorithm for the combined scenario of an arbitrary number of channels, arbitrary channel distributions, and statistically non-identical channels. It should be noted that when probing costs are equal for all finite-state channels, [11] has derived a class of strategies whose computation time is polynomial in the number of channels and can arbitrarily approximate the performance of the optimal strategy. When the probing costs possibly differ between channels, [10] has derived the optimal strategies when channels can only have two states, while [3] has proposed and studied suboptimal algorithms with performance guarantees for finite-state channels. Our algorithm and results apply to arbitrary, possibly infinite, number of channel states and to arbitrary (possibly differing) probing costs.

The remainder of this paper is organized as follows. We formulate the problem and present preliminary results in Sections II and III, respectively. Section IV introduces a channel probing algorithm and show that its optimal for a number of special cases. Section V 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 transmitter who would like to send a message using exactly one of the channels. With each channel j, we associate a probability of successful transmission by X_j, a random variable (discrete or continuous) with some distribution over the interval $[0, 1]$. The randomness of the transmission probability comes from the time-varying and uncertain nature of the wireless medium. The transmitter knows *a priori* the distribution of X_j for all $j \in \Omega$. By sending a channel *probe* on the jth channel, the user can find out more information about X_j.

We will assume temporal independence for the probability of transmission success on any given channel. That is, the channel state in any given time slot is independent of the state during other slots. This assumption allows us to focus on the transmission of a single message. In addition, we will assume independence between channels, i.e. $\{X_j\}_{j\in\Omega}$ are independent random variables. Thus, probing channel j does not provide any information about the state of any other channel in $\Omega - \{j\}$. These same assumptions were also made in [6], [3], [10], [11].

Note that in reality, the transmitter may not be directly probing to find the probability of transmission success. For example, channel probes may be used to measure the channel signal-to-noise ratio (SNR) in order to estimate channel conditions [5], [6]. This measured SNR, however, essentially affects the probability of transmission success 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. In addition, it should be noted that even if estimates obtained through probing are not perfect, X_j can nevertheless represent the *expected* probability of transmission success as a result of the probe. Thus without loss of generality we will simply assume the precise value (or the realization) of X_j is found after 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 his *a priori* information about the distribution of X_j. If it transmits over one of the channels, then 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 $\Omega - j$ in the rest of the paper), or using a channel in $\Omega - j$ for transmission even though it has not been probed. This decision process continues until the user decides which channel to use for transmission.

We can therefore think of the system as proceeding in discrete steps, where at each step the transmitter has a set of unprobed channels $S \subseteq \Omega$, and has determined the values 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; we call this *guessing* since the channel has not been probed. This last action was referred to as using a *backup channel* in [3]. Note that actions (2) and (3) can be seen as *stopping actions* that complete the entire probing and transmission process.

We will assume that the sender receives a reward of 1 upon a successful transmission over any channel. Thus X_j can also be thought of as the *reward* associated with using the jth channel for transmission. As mentioned earlier, in contrast to previous works [3], [10], we will not assume that X_j are discrete random variables. In particular, we allow X_j to be either continuous or discrete random variables. As mentioned earlier, this generalization allows us to model, for example, wireless systems with continuous data rates and channels with discrete state space.

We will also associate a cost c_j, where $c_j > 0$, with probing channel j. This cost may vary between channels, depending on the probing time, interference caused to other users, and so on. Though not free, probing channels can provide valuable information about the channel condition, particularly if the random variables $\{X_j\}$ have large entropy. Therefore, it is important that the sender efficiently balance the trade-off between probing too many channels and not using a good channel due to lack of probing. 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*.

With the assumptions and objectives outlined above, we have the following 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;

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

where π denotes the time-invariant strategy that probes channels in sequence $\pi(1), \cdots, \pi(\tau-1)$, and then transmits over channel π_τ at time τ. Π denotes the set of all possible strategies (all possible sequences of channel probes and transmissions), and the right hand sum in (1) is set to 0 if $\tau = 1$.

Note that τ is a random stopping time that may depend 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 strategy*. Such a strategy is guaranteed to exist since there are a finite number of strategies due to the finite number of channels.

Because the X_j's are probabilities of transmission success and thus upper bounded by 1, it can be seen that J^* is also upper bounded by 1. Thus we will further assume that $0 < E[X_j] < 1$ for all $j \in \Omega$, because if $E[X_j] = 1$, 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. That is, the optimal strategy becomes trivial if these assumptions are violated.

It can be shown that at any step, a sufficient information state [13] is given by the pair (u, S), where $S \subseteq \Omega$ is the set of unprobed channels and $0 \le u \le 1$ is the highest value among probed channels in $\Omega - S$. In other words, u is the probability of transmission success if we use the best probed channel. We can use dynamic programming [13] to represent the decision process 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:

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

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

Note that computing the value function $V(\cdot, \cdot)$ for every state is very difficult and practically impossible because there are possibly an infinite number of states, since u can be any real number in $[0, 1]$ if the X_j's are continuous random variables. For example, to compute $V(u, S)$, we may need to know $V(\tilde{u}, S - j)$ for all $j \in S$ and all $u \le \tilde{u} \le 1$. Nevertheless, it will be seen that the above formulation allows us to obtain important insights into optimal strategies, and helps us derive much simpler methods of determining optimal strategies that we will study in Section IV.

Any strategy can be defined by the set of actions it takes with respect to the set of information states, $\cup_S \cup_u (u, S)$. We thus use the following notation. We let retire(u) denote the action that the sender retires and uses the best previously probed channel in $\Omega - S$, which has value u; probe(j) denotes the action that channel j is probed, for some $j \in S$; and guess(j) denotes the action that channel j, for $j \in S$, is guessed (i.e., used even though it has not been probed). For state (u, S), a strategy must decide between retire(u), probe(j), and guess(j), for all $j \in S$. We let $\pi(u, S)$ denote the action taken by strategy π when state is (u, S). For example, $\pi(u, S) =$probe(j) means the sender probes channel j when the state is (u, S).

III. Preliminaries

In this section, we establish some properties of the optimal strategies which will be crucial for proving our main results later. 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 $\tilde{u} \geq u$,

$$V(u,S) \leq V(\tilde{u},S) \tag{3}$$

This inequality follows from (1) and (2). In particular, consider any channel selection strategy starting from state (u,S), and apply the same strategy starting from state $(\tilde{u},S)$. Clearly the expected reward of the strategy cannot be less in the latter starting scenario, since the set of unprobed channels is the same for both cases, while the best unprobed channels for the latter case is better than the best unprobed channel 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 $\tilde{S} \supseteq S$:

$$V(u,S) \leq V(u,\tilde{S}). \tag{4}$$

We have the following fundamental lemmas:

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 $\bar{u} \leq u$.

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 quantities:

$$a_S = \inf\{u : V(u,S) = u\} \tag{5}$$

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

where the right hand side of (5) is nonempty since $V(1,S) = 1$ is always true. We will 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 $0 \leq b_S \leq a_S \leq 1$. Thus we have the following corollary:

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

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

It should be noted for completeness that at $u = b_S$, $\pi^*(u,S) = \text{guess}(j)$ if $b_S > 0$; otherwise, $\pi^*(u,S) = \text{probe}(j)$. This corollary indicates that there exists an optimal strategy with the described threshold structure. It remains to determine these thresholds, which can be very difficult especially for large S. Secondly, it also remains to determine which channel should be probed if we are 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 indices concerning channel j that are *independent of other channels*. We will see that they are very useful for deciding the optimal strategies, thus significantly reducing the complexity of the problem.

It is thus worth taking a closer look at a_j and b_j. Note that at state (u,j), probing channel j incurs an expected reward $-c_j + E[\max\{u,X_j\}]$, since there are no channels to probe after j. Action guess(j) gives the expected reward $E[X_j]$ while retiring gives reward u. Because of the assumptions that $0 < E[X_j] < 1$ and $c_j > 0$, for sufficiently small u the probing reward becomes less than the guessing reward. By comparing the rewards of the three options, it can be seen that guessing is optimal if: $E[u - X_j | X_j < u]P(X_j < u) \leq c_j$ and $u \leq E[X_j]$. 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:

$$a_j = \min\{u : u \geq E[X_j],\ c_j \geq E[X_j - u | X_j > u]P(X_j > u)\} \tag{7}$$

$$b_j = \max\{u : u \leq E[X_j],\ c_j \geq E[u - X_j | X_j < u]P(X_j < u)\} \tag{8}$$

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, then a_j and b_j will be closer to $E[X_j]$.

The above discussion is depicted in Figure 1 where we have plotted the expected reward of the three actions guess(j) (dashed line), probe(j) (solid line), and 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 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 $f_3(u) = -c_j + uP(X_j < u) + E[X_j | X_j > u]P(X_j > u)$. Thus regardless of whether X_j is continuous or discrete, we do not expect computing a_j and b_j to be very complex.

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

B. Optimal Stopping

In this subsection we derive conditions for which it is optimal to stop, i.e., either to retire and use the best previously probed channel, or to guess and use an unprobed channel.

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,

$$a_S = \max_{j \in S} a_j\ . \tag{9}$$

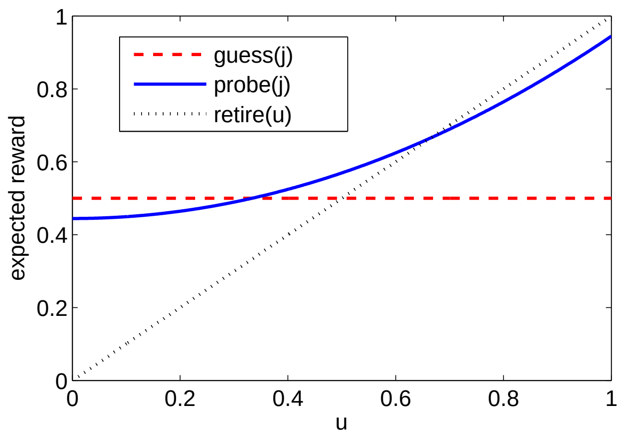

Fig. 1. As described in Section III-A: when j is the only unprobed channel and X_j is uniformly distributed in $[0,1]$, the expected reward from actions guess(j), probe (j) and retire(u) as functions of u. Note that $a_j = 2/3$ (the crossing point of solid and dotted lines) and $b_j = 1/3$ (the crossing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

This lemma provides both a necessary and sufficient condition for the optimality of retiring and using a previously probed channel. A very appealing feature of this lemmas 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.

The following lemma provides conditions for guessing to be optimal. Note that $b_S = 0$ implies guessing is not optimal for all $u \in [0, 1]$.

Lemma 4: Given a set of unprobed channels S, define R as follows:

$$R = \left\{ j \in S : a_j = \max_{k \in S} a_k \right\} . \qquad (10)$$

Then we have the following:

1) If there exists $j^* \in R$ satisfying $a_{j^*} > b_{j^*}$ and $b_{j^*} \leq \max_{j \in S - j^*} E[X_j]$, then there exists an optimal strategy with $b_S = 0$, i.e., it never guesses for any u.
2) If there exists $j^* \in R$ such that $b_{j^*} \geq \max_{j \in S - j^*} a_j$, then there exists an optimal strategy with $b_S = b_{j^*}$.

Thus 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 4 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 4 is always satisfied.

On the other hand, when $|R| = 1$ and letting $j^* = R$, suppose $\pi^*(u, S) =$guess(k) for some $k \neq j^*$, $u > 0$. This implies $E[X_k] > E[X_{j^*}] \geq b_{j^*}$, which leads to condition 1) of Lemma 4. This lemma implies that we have $b_S = 0$, which implies guessing is not optimal and thus contradicts the assumption that $\pi^*(u, S)$ =guess(k). Thus, we have shown that if $|R| = 1$ then $\pi^*(u, S) \neq$guess(k) for $k \notin R$. This narrows the possible channels to guess, and leads to the following corollary:

Corollary 2: Given a set of unprobed channels S, define R as in Lemma 4.

1) If $|R| \geq 2$ and $a_j = b_j$ for at least one $j \in R$, then $b_S = b_j$. Otherwise, $b_S = 0$.
2) If $|R| = 1$, let $\{j^*\} = R$. Then $\pi^*(u, S) \neq$guess(j) for all u and $j \in S - j^*$. Furthermore, $b_S \leq b_{j^*}$, i.e. $\pi^*(u, S)$ =guess(j^*) only if $u \leq b_{j^*}$.

This corollary and its preceding lemma are very useful as they allow us to narrow down the set of possible channels we can guess. In words, channel j in S with the highest value of a_j is the only possible channel we can guess. If there are multiple channels achieving this maximum, then we can easily check whether $a_j = b_j$ is true in order to determine whether probing or guessing is the optimal action.

C. Optimal Probing

In this subsection we 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 will help us derive useful properties of the optimal strategy for Problem 1.

Problem 2: Consider Problem 1 with the following modification: at each step, the user must choose between the following *two* actions: (1) probe a channel that has not yet been probed, or (2) retire and use the best previously probed channel. Therefore, the user is not allowed to use a channel that has not yet been probed.

This problem can be seen as a generalization of the problem considered in [3], which restricted X_j to be discrete random variables. To describe the theorem, we use the following notation for any channel $j \in \Omega$:

$$\bar{a}_j = \min\{u : c_j \geq E\left[X_j - u | \, X_j > u\right] P(X_j > u)\} ,$$

where $\bar{a}_j = 0$ if the above set is empty. Note that from equations (7) and (8), we see that $\bar{a}_j = a_j$ if and only if $a_j > b_j$. If $a_j = b_j$, then $\bar{a}_j < a_j$. We use these indices in the following theorem, which can be seen as a generalization of Theorem 4.1 in [3].

Theorem 1: For state (u, S), the optimal strategy $\hat{\pi}$ for Problem 2 is described as follows:

1) If $u \geq \max_{j \in S} \{\bar{a}_j\}$, then $\hat{\pi}(u, S)$ =retire(u).
2) Otherwise, define the following:

$$R = \left\{ j \in S : \bar{a}_j = \max_{k \in S} \bar{a}_k \right\} .$$

$$j^* = \text{argmax}_{j \in R} \left\{ E[X_j \, | X_j \geq \bar{a}_j] - \frac{c_j}{P(X_j \geq \bar{a}_j)} \right\} .$$

Then $\hat{\pi}(u, S)$ =probe(j^*) .

This theorem implies that by first ordering the individual channels by functions of the indices $\bar{a}_j$, we can determine the optimal channel to probe. Again note that the $\bar{a}_j$'s are individual channel indices computed independently of other channels, thus reducing the computational complexity significantly.

Note that even though Problem 2 is different from Problem 1, its optimal strategy will also be optimal for Problem 1 if

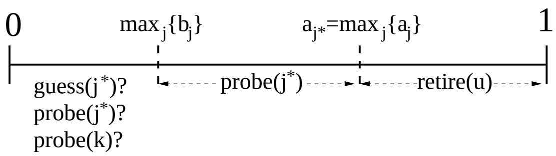

Fig. 2. Summary of main results from Section III: figure depicts optimal strategy $\pi^*(u, S)$ as a function of u. For the middle and right regions of the line, the optimal strategy is well-defined for any S. For the left region, the optimal action may depend on S.

guessing becomes non-optimal for all future time steps. From Lemma 3 and Corollary 2, probing occurs if $\max_{j \in S} b_j < u < \max_{j \in S} a_j$. Thus we have the following result:

Corollary 3: For any state (u, S) where $\max_{j \in S} b_j < u < \max_{j \in S} a_j$, the optimal strategy π^* for Problem 1 is $\pi^*(u, S) =$probe(j^*), where j^* is determined as follows: first, define set R as in Theorem 1 but replace $\bar{a}_j$ with a_j. Then define j^* as in Theorem 1 with $\bar{a}_j$ again replacing a_j.

Figure 2 summarizes the main results from Section III. We have shown that for all $u \geq a_{j^*}$, i.e. right region of the line, retire(u) is optimal. For $\max_{j \in S} b_j < u < a_{j^*}$, i.e. the middle region of the line, probe(j^*) 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 guess(j^*) is the only possible guessing action for this region, as proven in Lemma 4 and Corollary 2.

IV. Optimal Strategies

As stated earlier, the state space for our problem is the set of all (u, S), which is infinite. This makes it very difficult to recursively apply dynamic programming to evaluate all $V(u, S)$ and solve for the optimal strategy. In this section we propose an algorithm for channel probing for arbitrary number of channels with arbitrary distributions. This algorithm is motivated by the properties derived in the previous section. We show that this algorithm is optimal for a number of special cases.

A. A Channel Probing Algorithm

In order to motivate our algorithm, which we will call γ, consider when there are two unprobed channels $S = \{j \cup k\}$. As described in the previous section, the ordering of the constants $a_j,\ b_j,\ a_k$, and b_k will help determine the optimal strategy. Note that due to Corollary 1, it is not hard to calculate the expected reward of probing j or k for state (u, S). For example, if $u < b_k$, then probe(j) at state (u, S) incurs the following expected reward:

$$\begin{aligned}
&- c_j + E[V(\max(u, X_j), k)] \\
&= -c_j + P(X_j < b_k)E[X_k] + P(X_j \geq a_k)E[X_j|\, X_j \geq a_k] \\
&+ P(b_k \leq X_j < a_k)\left(-c_k + E[\max(X_k, X_j)|\, a_k > X_j \geq b_k]\right) .
\end{aligned}$$

The above calculation can similarly be applied to the other two separate cases of $u > a_k$ and $a_k > u > b_k$, and they can also similarly be applied to determine the expected reward of action probe(k). Note that this procedure computes the expected probing reward in a finite number of steps, whereas not using the threshold properties given by Corollary 1 would first require the computation of $V(u, j)$ and $V(u, k)$ for all $u \in [0, 1]$, thus requiring an infinite number of computations.

Motivated by the above, the proposed algorithm is as follows. It essentially finds two channels indexed by j^* and k, and use these to define the strategy γ.

Algorithm 1: (A two-step lookahead policy γ for a given set of unprobed channels $S \subseteq \Omega$)

Step 1: Compute R and j^* as follows:

$$\begin{aligned}
R &= \left\{ j \in S : a_j = \max_{k \in S} a_k \right\} \\
j^* &= \arg\max_{j \in R} \Big\{ I_{\{b_j = a_j\}} E[X_j] + \\
&\quad I_{\{a_j > b_j\}} \left(E[X_j\, |X_j \geq a_j] - \frac{c_j}{P(X_j \geq a_j)} \right) \Big\} ,
\end{aligned}$$

where $I_{(\cdot)}$ is the indicator function.

Step 2: Replace S with $S - j^*$ and repeat Step 1. The result of the second equation above is k.

Then strategy γ is defined as follows for state (u, S):

1) If $u \geq a_{j^*}$, then $\gamma(u, S) =$retire(u).
2) If $a_{j^*} > u > b_{j^*}$, then $\gamma(u, S) =$probe(j^*).
3) If $u \leq b_{j^*}$ then we have the following cases:
    a) If $b_{j^*} \geq a_k$, then $\gamma(u, S) =$guess(j^*).
    b) If either $b_k \geq b_{j^*}$ or $-c_{j^*} + E[V(X_{j^*}, k)] \geq \{E[X_{j^*}], -c_k + E[V(X_k, j^*)]\}$, then $\gamma(u, S) =$probe(j^*).
    c) Otherwise, there exists a unique b_0, where $b_{j^*} > b_0 > b_k$, such that $-c_{j^*} + E[V(\max(b_0, X_{j^*}), k)] = \max\{E[X_{j^*}], -c_k + E[V(X_k, j)]\}$. Then for $b_0 \leq u \leq b_{j^*}$, we have $\gamma(u, S) =$probe(j^*). For $u < b_0$, we have $\gamma(u, S) =$guess(j^*) if $E[X_{j^*}] \geq -c_k + E[V(X_k, j^*)]$. Otherwise, $\gamma(u, S) =$probe(k).

It is worth describing the motivation behind this strategy. For u satisfying case 1) of the algorithm description, γ is optimal from Lemma 3. For some of the u values described in Case 2), γ is optimal from Corollary 3. For case 3a), γ is optimal from Lemma 4 and Corollary 2. Thus γ is optimal for most values of u. For cases 3b) and 3c) of Algorithm 1, the procedure essentially computes the expected probing cost if we are forced to retire in two steps.

We next consider a few special cases and show that γ is optimal in these cases.

B. Special Cases

We first consider a two channel system. Since Algorithm 1 is essentially a two-step lookahead policy, we have the following result (the proof is omitted for brevity):

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

We next consider the case when all channels are statistically identical, but there are an arbitrary number of them. The channels can have different probing costs.

Theorem 3: 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 j^* being a channel in S satisfying $c_{j^*} = \min_{j \in S} \{c_j\}$.

Case 1: If $a_{j^*} > b_{j^*}$, we have:

$$\pi^*(u, S) = \begin{cases} \text{retire}(u) & \text{if } u \geq a_{j^*} \\ \text{probe}(j^*), & \text{otherwise} \end{cases}$$

Case 2: If $a_{j^*} = b_{j^*}$, the optimal strategy is: retire(u) if $u \geq a_{j^*}$; otherwise, $\pi^*(u, S) =$guess(j^*).

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_{j^*} and b_{j^*}, where j^* is the channel with smallest probing cost. If $a_{j^*} = b_{j^*}$, then $\pi^*(u, \Omega) =$guess(j^*) and it is not worth probing any channels. If $a_{j^*} > b_{j^*}$, then we should first probe j^*. Let k denote the channel with the smallest probing cost in $S - \{j^*\}$. If the probed value of X_{j^*} is higher than a_k, then it is optimal to retire and use j^* 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 $|S| = 1$. When $|S| = 1$, then we compare the best probed channel, which has value u, to a_j and b_j, where $S = \{j\}$. If $u \geq a_j$, then we retire; if $a_j > u \geq b_j$, then we probe the last remaining channel; finally, if $u < b_j$ then we should just use the last remaining channel without probing it.

Note that the optimal strategy described above is the same as strategy γ of Algorithm 1 applied to statistically identical channels. This is true because within case 3) in the description of Algorithm 1, 3b) will occur whenever $a_{j^*} > b_{j^*}$ for statistically identical channels, and case 3a) occurs whenever $a_{j^*} = b_{j^*}$. Collectively, 1), 2), 3a) and 3b) all describe the optimal strategy of Theorem 3. 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. It should be noted that [5], [6], [7] have all considered variants of Problem 1 where the set of channels are statistically identical. However, their results do not allow unprobed channels to be used for transmission. In addition, [6], [7] do not determine the optimal order for probing channels.

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

Problem 3: Consider Problem 1 with the following modification: the wireless system consists of N different types of channels, but an infinite number of each channel type.

Note that Theorem 3 solves this problem if $N = 1$. When referring to the state space for this problem, we will let S denote the set of available channel *types*. Theorem 3 says that if we have many statistically identical channels of one type, then whether $a_j > b_j$ or $a_j = b_j$ determines if we will probe or guess channel j. Analogously, we have the following result:

Theorem 4: For any set of channels S, define j^* according to *Step 1* of Algorithm 1. Then for Problem 3, there exists an optimal strategy π^* satisfying the following:

1) If $u \geq a_{j^*}$, then $\pi^*(u, S) =$retire(u).
2) If $u < a_{j^*}$ and $b_{j^*} = 0$, then $\pi^*(u, S) =$probe(j^*).
3) If $u < a_{j^*}$ and $b_{j^*} > 0$, then $\pi^*(u, S) =$guess(j^*).

Due to space limitations, proof of the above theorem is not included; however, it should be noted that it essentially follows from Theorem 3. 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, i.e. the other channels become irrelevant. In addition, note that Algorithm 1 is also the optimal strategy for Problem 3, because case 3c) of the description of Algorithm 1 does not occur. For all the other cases, Algorithm 1 reduces to the optimal strategy described in Theorem 4.

To summarize, in this subsection we have shown that Algorithm 1 reduces to the optimal strategies for the above special cases based on Theorems 2, 3, and 4.

V. 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. We also proposed a channel probing algorithm, shown to be optimal for some cases of practical interest, including statistically identical channels, a few nonidentical channels, and a large number of nonidentical channels.

References

[1] X. Liu, E. Chong, and N. Shroff, “Transmission scheduling for efficient wireless network utilization,” *Proceedings of IEEE INFOCOM*, 2001, anchorage, AK.

[2] X. Quin and R. Berry, “Exploiting multiuser diversity for medium access control in wireless networks,” *Proceedings of IEEE INFOCOM*, 2003, san Francisco, CA.

[3] S. Guha, K. Munagala, and S. Sarkar, “Jointly optimal transmission and probing strategies for multichannel wireless systems,” *Proceedings of Conference on Information Sciences and Systems*, March 2006, princeton, NJ.

[4] J. Kennedy and M. Sullivan, “Direction finding and ”smart antennas” using software radio architechtures,” *IEEE Communications Magazine*, pp. 62–68, May 1995.

[5] Z. Ji, Y. Yang, J. Zhou, M. Takai, and R. Bagrodia, “Exploiting medium access diversity in rate adaptive wireless lans,” *ACM MOBICOM*, September 2004, philadelphia, PA.

[6] A. Sabharwal, A. Khoshnevis, and E. Knightly, “Opportunistic spectral usage: Bounds and a multi-band csma/ca protocol,” *IEEE/ACM Transactions on Networking*, 2007, accepted for publication.

[7] V. Kanodia, A. Sabharwal, and E. Knightly, “Moar: A multi-channel opportunistic auto-rate media access protocol for ad hoc networks,” *Proceedings of Broadnets*, October 2004.

[8] R. Knopp and P. Humblet, “Information capacity and power control in a single cell multiuser environment,” *Proceedings of IEEE ICC*, 1995, seattle, WA.

[9] G. Holland, N. Vaidya, and P. Bahl, “A rate-adaptive mac protocol for multi-hop wireless networks,” *Proceedings of ACM MobiCom*, 2001, rome, Italy.

[10] S. Guha, K. Munagala, and S. Sarkar, “Optimizing transmission rate in wireless channels using adaptive probes,” *Poster paper in ACM Sigmetrics/Performance Conference*, June 2006, saint-Malo, France.

[11] ——, "Approximation schemes for information acquisition and exploitation in multichannel wireless networks," *Proceedings of 44th Annual Allerton Conference on Communication, Control and Computing*, September 2006, monticello, IL.
[12] J. Heiskala and J. Terry, "Ofdm wireless lans: A theoretical and practical guide," *SAMS*, 2001.
[13] P. Kumar and P. Karaiya, *Stochastic Systems: Estimation, Identification, and Adaptive Control*. Prentice-Hall, Inc, 1986, englewood Cliffs, NJ.

VI. Appendix

A. Proof of Lemma 1

It suffices to prove that the following holds for all S and $0 \le u \le \tilde{u} \le 1$:

$$V(\tilde{u}, S) - V(u, S) \le \tilde{u} - u \ . \tag{11}$$

If (11) is true, then $V(u, S) = u$ implies that $V(\tilde{u}, S) \le \tilde{u}$. Combining this with (2), which says that $V(\tilde{u}, S) \ge \tilde{u}$, proves the lemma. We prove (11) by induction on the cardinality of S.

Induction Basis: Consider any $S \subseteq \Omega$ such that $|S| = 1$. Let $j = S$. From equation (2) defining the value function, $V(\tilde{u}, \{j\})$ (also simplified as $V(\tilde{u}, j)$ below) has three possible values. We show (11) holds for all three cases.

Case 1: $V(\tilde{u}, j) = \tilde{u}$. From (2), $V(u, j) \ge u$. Therefore, equation (11) easily follows.

Case 2: $V(\tilde{u}, j) = -c_j + E[\max(X_j, \tilde{u})]$. From (2), $V(u, j) \ge -c_j + E[\max(X_j, u)]$. Therefore,

$$\begin{aligned} V(\tilde{u}, j) - V(u, j) &\le E\left[\max(X_j, \tilde{u})\right] - E\left[\max(X_j, u)\right] \\ &= E[\tilde{u} - \max(X_j, u) | X_j < \tilde{u}] P(X_j < \tilde{u}) \le \tilde{u} - u \ , \end{aligned}$$

which proves that (11) holds.

Case 3: $V(\tilde{u}, j) = E[X_j]$. From (2), $V(u, j) \ge E[X_j]$. Therefore, equation (11) easily follows.

Induction Hypothesis: Consider any $S \subseteq \Omega$ such that $|S| \ge 2$ and suppose (11) holds for all $\tilde{S} \subseteq \Omega$ such that $|\tilde{S}| < |S|$. Again we prove that (11) holds for all possible values of $V(\tilde{u}, S)$. If $V(\tilde{u}, S) = \tilde{u}$, then (2) implies $V(u, S) \ge u$ which implies (11) holds. Similarly, if $V(\tilde{u}, S) = E[X_j]$ for some $j \in S$, then $V(u, j) \ge E[X_j]$ implies equation (11). Finally, suppose $V(\tilde{u}, S) = -c_j + E[\max(X_j, \tilde{u})]$ for some $j \in S$. Then because $V(u, S) \ge -c_j + E[V(\max(X_j, u), S - j)]$ for the same j we have:

$$\begin{aligned} &V(\tilde{u}, S) - V(u, S) \\ &\quad \le E\left[V(\max(X_j, \tilde{u}), S - j) - V(\max(X_j, u), S - j)\right] \\ &\quad = E\left[(V(\tilde{u}, S - j) - V(\max(X_j, u), S - j))\, I_{(X_j < \tilde{u})}\right] \\ &\quad \le V(\tilde{u}, S - j) - V(u, S - j) \le \tilde{u} - u \ , \end{aligned}$$

where the last two inequalities follow from (3) and the induction hypothesis, respectively, and $I_{(\cdot)}$ is the indicator function. Therefore we have proven (11). □

B. Proof of Lemma 2

This lemma follows from (3). In particular, suppose $V(u, S) = E[X_j]$. Then (3) implies $V(\bar{u}, S) \le E[X_j]$. However, from (2) we know $V(\bar{u}, S) \ge E[X_j]$. Combining the two implies $V(\bar{u}, S) = E[X_j]$ and proves the lemma. □

C. Proof of Lemma 3

We prove the lemma by contradiction, on two separate cases:

Case 1: Suppose $a_S < \max_{j \in S} a_j$. Equivalently, $a_S < a_k$ for at least one $k \in S$. Fix u such that $a_S < u < a_k$. By definition of a_S, we have $V(u, S) = u$. On the other hand, the definition of a_k and $u < a_k$ implies:

$$V(u, k) = \max\left\{E[X_k], -c_k + E\left[\max(X_k, u)\right]\right\} > u \ ,$$

Finally, (4) gives $V(u, S) \ge V(u, k) > u$, which contradicts the assumption that $u > a_S$. Thus, $a_S < a_k$ is not possible for any $k \in S$.

Case 2: Suppose $a_S > \max_{j \in S} a_j$. Fix any u such that $\max_{j \in S} a_j < u < a_S$. By definition of a_S, we have the optimal strategy at state (u, S) is to either probe or guess a channel in S, but retiring is not optimal. Suppose the optimal strategy is to probe a channel $k \in S$. This implies:

$$\begin{aligned} -c_k + E\left[V(\max(u, X_k), S - k)\right] &\ge V(u, S) \\ &\ge V(u, S - k) \ , \end{aligned} \tag{12}$$

where the last inequality follows from (4).Since $u > a_k$, then by definition of a_k we have:

$$-c_k + E\left[\max(u, X_k)\right] < u \tag{13}$$

Combining (12) and (13), we have:

$$\begin{aligned} E\left[V(\max(u, X_k), S - k) - V(u, S - k)\right] &> c_k \\ &> E\left[\max(u, X_k) - u\right] \end{aligned}$$

Conditioning the above expectations on the events $\{X_k > u\}$ and $\{X_k \le u\}$ gives us:

$$\begin{aligned} &E\left[\left.V(X_k, S - k) - V(u, S - k)\right| X_k > u\right] P(X_k > u) \\ &\quad > E\left[\left.X_k - u\right| X_k > u\right] P(X_k > u) \ , \end{aligned} \tag{14}$$

which contradicts (11).

If the optimal strategy is to guess a channel $k \in S$, then $V(u, S) = E[X_k]$. However, since $V(u, k) \le V(u, S)$ then $V(u, k) = E[X_k]$ as well. This implies $a_k \ge u$, which contradicts the assumption that $u > a_k$.

Thus combining Case 1 and Case 2, we have proven Lemma 3. □

D. Proof of Lemma 4

Case 1: For notation, let $E[X_k] = \max_{j \in S - j^*} E[X_j]$ for $j^* \in R$ satisfying $a_{j^*} > b_{j^*}$ as described by the lemma. From the lemma, we know that $b_{j^*} \le E[X_k]$. If we can show that for every u, probing some channel in S or retiring is better than guessing any channel in S, then this will prove there exists an optimal strategy with $b_S = 0$. Note that the expected reward of guessing the best channel is $\max_{j \in S} E[X_j] = \max\{E[X_k], E[X_{j^*}]\}$. Thus it suffices to show that for all $u \le a_{j^*}$, there exists a probing strategy with higher expected reward than $E[X_k]$ and $E[X_{j^*}]$.

As described in Section III-A, $E[X_k] \le a_k$. Thus we have $b_{j^*} \le E[X_k] \le a_{j^*}$ since j^* is in R. From the definition of b_{j^*}, a_{j^*} and by the assumption that $a_{j^*} > b_{j^*}$, then $\pi^*(u,j^*)$ =probe(j^*) whenever $b_{j^*} \le u \le a_{j^*}$. Therefore, $\pi^*(E[X_k],j^*)$ =probe(j^*) and we have have:

$$V(E[X_k],j^*) = -c_{j^*} + E[\max(X_{j^*},E[X_k])] \ge E[X_k] \ ,$$

However, note that the lefthandside of the above equation is the expected reward of the following strategy: probe j^* first, and use this channel for transmission if its value is higher than $E[X_k]$; if its value is lower than $E[X_k]$, then $guess(k)$, i.e. use channel k for transmission. Thus the expected reward of this two-step strategy is always at least the reward of simply using channel k for transmission. This result holds for all u. In addition, by definition of a_{j^*} and b_{j^*}, $\pi^*(u,j^*)$ =guess(j^*) for all $u < b_{j^*}$. Thus $V(u,j^*) = E[X_{j^*}]$ for all such u. However, from equation (3), we also know that $V(E[X_k],j^*) \ge V(u,j^*) = E[X_{j^*}]$. Thus we have shown that for all u, there exists a strategy of probing j first which does at least as good as the strategy of guess(k) or guess(j^*). As explained earlier, these are the two best guessing actions; thus, there exists an optimal strategy which never guesses for all u, i.e. $b_S = 0$.

Case 2: From the lemma, we have $a_{j^*} \ge b_{j^*} \ge \max_{j\in S-j^*} a_j$. In addition, from Lemma 3 and from the threshold properties described in Section III-A, we have $b_S \le a_S = a_{j^*}$. Thus, $V(u,b_S) = u$ for all $u \ge a_{j^*}$. Now we have two cases for the relationship between a_{j^*} and b_{j^*}. First, suppose $a_{j^*} = b_{j^*}$. From the equations in Section III-A for individual channel indices, we see that this equality implies that $E[X_{j^*}] = b_{j^*}$. Finally, using (3), we see that $V(u,S) \le V(b_{j^*},S) = b_{j^*} = E[X_{j^*}]$ for all $u < b_{j^*}$. On the other hand, from (2) we have $V(u,S) \ge E[X_{j^*}]$ for all $u < b_{j^*}$. Thus we have shown that $V(u,S) = E[X_{j^*}]$ for all $u < b_{j^*}$. This implies that $b_S = b_{j^*}$.

Now suppose $a_{j^*} > b_{j^*}$. This implies $a_{j^*} > \max_{j\in S-j^*} a_j$. Thus, by using Corollary 3, to be proven later, we have that $\pi^*(u,S)$ =probe(j^*) for all $a_{j^*} > u \ge b_{j^*}$. Thus, $V(b_{j^*},S) = -c_{j^*} + E[\max(X_{j^*},b_{j^*})]$, where we do not probe anything after j^* because of Lemma 3. Finally, from Section III-A, we have that $a_{j^*} > b_{j^*}$ implies $-c_{j^*} + E[\max(X_{j^*},b_{j^*})] = E[X_{j^*}]$. Thus, $V(b_{j^*},S) = E[X_{j^*}]$, which again implies that $V(u,S) = E[X_{j^*}]$ for all $u < b_{j^*}$. Thus we have shown there exists an optimal strategy with $b_S = b_{j^*}$. □

E. Proof of Theorem 1

The proof that $\hat{\pi}(u,S)$ =retire(u) for all $u \ge \max_{j\in S}\{\bar{a}_j\}$ follows from the same steps as proving Lemma 3 (the fact that guessing is not an option does not alter the result of this proof).

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 $\hat{\pi}(u,j)$ =probe(j) follows from the definition of $\bar{a}_j$, because $-c_j + E[\max(X_j,u)] > u$ for all $u < \bar{a}_j$.

Induction Hypothesis: Let $|S| = n \ge 2$, and suppose the result holds for all $\tilde{S} \subseteq \Omega$ such that $|\tilde{S}| < n$. Define R and j^* as in the theorem. For notational convenience, we index channels by the set $\{j_1, j_2, \cdots, j_n\}$, where $\bar{a}_{j_n} \ge \bar{a}_{j_{n-1}} \ge \cdots \ge \bar{a}_{j_1}$ and when $\bar{a}_{j_m} = \bar{a}_{j_{m-1}}$, then

$$\begin{aligned} &E[X_{j_m}|X_{j_m} \ge \bar{a}_{j_m}] - \frac{c_{j_m}}{P(X_{j_m} \ge \bar{a}_{j_m})} \\ &\quad \ge E[X_{j_{m-1}}|X_{j_{m-1}} \ge \bar{a}_{j_{m-1}}] - \frac{c_{j_{m-1}}}{P(X_{j_{m-1}} \ge \bar{a}_{j_{m-1}})} \ . \end{aligned}$$

Thus $j_n = j^*$ as defined in the theorem.

Now consider any j_m where $1 \le m < n$ such that $j_m \in S - R$. Let u satisfy $\bar{a}_{j_m} \le u < \bar{a}_{j_n}$, and suppose $\hat{\pi}(u,S)$ =probe(j_m). We will show by contradiction that this cannot be true. In fact, following the same exact steps as (12),(13), and (14), we arrive at a contradiction to (11). Therefore, $V(u,S) \neq$probe(j) for all $u > \bar{a}_j$ and all $j \in S - R$.

From Lemma 3, retiring cannot be optimal (again, removing guessing as an option does not change this result). Therefore, the optimal strategy at (u,S), $\bar{a}_{j_m} \le u < \bar{a}_{j_n}$, must be to probe a channel in R. To see which channel in R to probe, we prove by contradiction that $\hat{\pi}(u,S)$ =probe(j_n). Suppose $\hat{\pi}(u,S) \neq$probe(j_n), so that $\hat{\pi}(u,S)$ =probe(j_m) for some $j_m \in R, m \neq n$. Note that $\bar{a}_{j_m} = \bar{a}_{j_n}$ by definition of R. Probing j_m first will cost c_{j_m}. By the induction hypothesis, at state $S - j_m$ we will probe j_n if $X_{j_m} < \bar{a}_{j_n}$, or we will retire if $X_{j_m} \ge \bar{a}_{j_n}$. Similarly, we can compute the expected reward of probing j_n first, and then probing j_m in the second step if $X_{j_n} < \bar{a}_{j_m}$. Since we are assuming probing j_n is not optimal, then the expected gain of probing j_m first minus the expected gain of probing j_n first must be positive. Taking this difference and cancelling terms gives us:

$$\begin{aligned} &-c_{j_m} + P(X_{k_m} \ge \bar{a}_{j_n})E[X_{j_m}|X_{j_m} \ge \bar{a}_{j_n}] \\ &\quad -c_{j_n}P(X_{j_m} < \bar{a}_{j_n}) > -c_{j_n} - c_{j_m}P(X_{j_n} < \bar{a}_{j_n}) \\ &\qquad\qquad + P(X_{j_n} \ge \bar{a}_{j_n})E[X_{j_n}|X_{j_n} \ge \bar{a}_{j_n}] \ , \end{aligned}$$

Rearranging, we get a contradiction to the definition of j^* in Theorem 1 and the fact that $j^* = j_n$. Thus, we arrive at a contradiction to the assumption that probing j_n is not optimal.

This holds for all j_m not in R. Therefore, we have shown that $\hat{\pi}(u,S)$ =probe(j_n) for all $\max_{j\in S-R} \le u < \bar{a}_{j_n}$.

Now to show that $\hat{\pi}(u,S)$ =probe(j_n) for all $u < \max_{j\in S-R} a_j$, we note that for all $u \le \bar{a}_{j_n}$ the expected reward of probing j_m first can again be calculated by using the induction hypothesis. In particular, if $X_{j_m} \ge \bar{a}_{j_n}$ then we retire. Otherwise, $X_{j_m} < \bar{a}_{j_n}$ and by the induction hypothesis we continue. It suffices to show for all $u \le \bar{a}_{j_n}$, the difference in expected rewards between probing j_n first and probing k, where k is any other channel, does not depend on u. If this holds, then probe(j_n) must be optimal for all $u \le \bar{a}_{j_n}$ since we have already shown that probe(j_n) is optimal for $\max_{j\in S-R} \bar{a}_j \le u < \bar{a}_{j_n}$. Due to space limitations, we will only consider the alternate strategy of probing j_{n-1} first, but we note that the steps generalize to any other channel $j_m \neq j_n$.

By the induction hypothesis, probing j_{n-1} first gives expected reward:

$$- c_{j_{n-1}} + E[X_{j_{n-1}} I_{\{B\}}] - P(X_{j_{n-1}} < \bar{a}_{j_n}) c_{j_n} + E\left[\max(X_{j_n}, X_{j_{n-1}}) I_{\{A \cap B^c\}}\right] + E\left[\hat{V}(\max(u, X_{j_n}, X_{j_{n-1}}), S - j_n - j_{n-1}) I_{\{A^c \cap B^c\}}\right] ,$$

where $\hat{V}(\cdot,\cdot)$ is the value function for Problem 2, defined similarly to (2), A is the event $\{\max(X_{j_n}, X_{j_{n-1}}) \geq \bar{a}_{j_{n-2}}\}$, A^c is its complement, and B is the event $\left\{X_{j_{n-1}} \geq \bar{a}_{j_n}\right\}$. Similarly, probing j_n first gives expected reward:

$$- c_{j_n} + E[X_{j_n} I_{\{D\}}] - P(X_{j_n} < \bar{a}_{j_{n-1}}) c_{j_{n-1}} + E\left[\max(X_{j_n}, X_{j_{n-1}}) I_{\{A \cap D^c\}}\right] + E\left[\hat{V}(\max(u, X_{j_n}, X_{j_{n-1}}), S - j_n - j_{n-1}) I_{\{A^c \cap D^c\}}\right] ,$$

where D denotes the event $\left\{X_{j_n} \geq \bar{a}_{j_{n-1}}\right\}$. Taking the difference between this expected reward and the expected reward of probing j_{n-1} first, we see that the difference 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 whether events B and D occur). Similar steps can be taken for other $j_m \neq j_n$, by calculating the expected reward for any strategy until only channels $\{j_1, \cdots, j_{m-1}\}$ are left. It can then be shown that the difference in expected reward between actions probe(j_n) and probe(j_m) does not change with u. Therefore, $\hat{\pi}(u, S)$ =probe(j_n) for all $u < \max_{j \in S-R} \{a_j\}$.

Therefore, we have shown $\hat{\pi}(u, S)$ =probe(j_n) for all $u < \bar{a}_{j_n}$, which completes the proof. □

F. Proof of Theorem 3

Note that when probing costs are equal for all channels, then this theorem follows from Case 1) of Corollary 2. In particular, since all channels in S are statistically identical, R defined in Corollary 2 is equal to S. Thus, $|R| = |S| \geq 2$ and the optimal strategy is determined from whether $a_j = b_j$ for any $j \in S$.

When probing costs differ between channels, then we can use induction on the cardinality of S to prove the result. Note that from the discussion in Section III, when $c_j \leq c_k$ but X_j and X_k have the same distribution, then $a_j \geq a_k$ while $b_j \leq b_k$. We will use this fact throughout the proof.

Induction Basis: Suppose $|S| = 2$. From Theorem 2, the strategy given in Theorem 3 is optimal.

Induction Hypothesis: Consider any $S \subseteq \Omega$ such that $|S| \geq 3$ and suppose the theorem holds for all $\tilde{S} \subseteq \Omega$ such that $|\tilde{S}| < |S|$. For notational convenience, let $S = \{j_1, j_2, \cdots, j_n\}$ where $c_{j_1} \geq c_{j_2} \geq \cdots \geq c_{j_n}$. As mentioned earlier, this assumption implies $a_{j_n} \geq a_{j_{n-1}} \geq \cdots \geq a_{j_1}$ and $b_{j_n} \leq b_{j_{n-1}} \leq \cdots \leq b_{j_1}$.

From Lemma 3, we know that $a_S = a_{j_n}$. Thus, it only remains to determine the optimal strategy $\pi^*(u, S)$ for $u < a_{j_n}$.

Case 1: We first prove the theorem for the case where $a_{j_n} > b_{j_n}$. From Corollary 2, the only channel that can be guessed is j_n. However, because $b_{j_n} \leq b_{j_{n-1}}$, then from Lemma 2 we have $b_S = 0$. Thus, guessing j_n is not an optimal action for all u and therefore we only need to decide which channel to probe when $u < a_{j_n}$.

We derive the optimal strategy here for two separate subcases. First suppose $a_{j_l} > b_{j_l}$ for all $1 \leq l \leq n$. In particular, this implies $a_{j_n} > b_{j_1}$. Let $V^*(u, S)$ denote the expected reward of the following strategy: first probe j_n and then proceed according to the optimal strategy as determined by the induction hypothesis. Meanwhile, let $H(u, S)$ denote the expected reward of first probing some channel j_k, where $k < n$, and proceeding according to the optimal strategy. For any $b_{j_1} \leq u < a_{j_n}$, it can be shown similar to the proof of Theorem 1 that $V^*(u, S) - H(u, S)$ is invariant to u. However, from Corollary 3 we know that $\pi^*(u, S)$ =probe(j_n) for all $b_{j_1} \leq u < a_{j_n}$, which means $V^*(u, S) > H(u, S)$ for these values of u. Combining everything implies that $\pi^*(b_{j_1}, S)$ =probe(j_n) and $V(b_{j_1}, S) = V^*(b_{j_1}, S)$. Finally, it can be easily shown that $V^*(u, S) = V^*(b_{j_1}, S)$ for any $u < b_{j_1}$ because no channel is guessed unless j_1 is the only remaining channel. From (3), this implies that $V(u, S) = V^*(u, S)$ for all $u < b_{j_1}$; therefore, $\pi^*(u, S)$ =probe(j_n) for all $u < b_{j_1}$.

Now suppose $a_{j_l} = b_{j_l}$ for some $1 \leq l < n$ (we let l denote the largest index satisfying $a_{j_l} = b_{j_l}$). Consider probing any channel j_k where $k < n$ and $k \neq l$. Then from the induction hypothesis, after probing j_k we will either retire or continue to probe channels in decreasing order of the indices $\left\{a_{j_n}, a_{j_{n-1}} \cdots, a_{j_l}\right\}$. If the state is reached where channel j_l has the highest index value, then from the induction hypothesis we will retire if $\max\left\{X_{j_k}, X_{j_n}, \cdots, X_{j_{l+1}}\right\} \geq a_{j_l}$; otherwise, the optimal action is guess(j_l) which collects a reward of $E[X_{j_l}]$. Since j_l is never probed, the total expected reward of this strategy is exactly the same as the reward of a strategy in Problem 2 where initially $u = E[X_j]$, channels are probed in the order: $\{j_k, j_n, \cdots, j_{l+1}, j_l\}$ and retirement occurs according to Theorem 1. Similarly, first probing j_n has the same expected reward as a strategy in Problem 2 that probes channels in the order $\{j_n, j_{n-1}, \cdots, j_l\}$ and retires according to Theorem 1, where again $u = E[X_j]$. Given this equivalence, we can use Theorem 1 to show that the latter strategy must have higher expected reward. Similar steps can be used to show that probe(j_l) cannot be optimal for any u. Thus, $\pi^*(u, S)$ =probe(j_n) for all $u < a_{j_n}$.

Case 2: Now suppose $a_{j_n} = b_{j_n}$. This implies from Corollary 2 that $\pi^*(u, S)$ =guess(j_n) for all $u \leq a_{j_n} = b_{j_n}$. □