Approximation Schemes for Information Acquisition and Exploitation in Multichannel Wireless Networks

Sudipto Guha, Kamesh Munagala, Saswati Sarkar

***Abstract*— Nodes in future wireless networks are likely to have access to multiple channels. A node can learn the instantaneous state of a channel only by probing it which in turn consumes both additional energy and time. A node therefore needs to not only optimally select the channel based on available information but also optimally determine the amount of information it should acquire about the instantaneous states of its available channels. The successful exploitation of the available channels is therefore contingent upon designing simple mechanisms for jointly optimizing both information acquisition and exploitation. We provide a joint channel probing and selection scheme that can approximate a utility function that captures both the cost and value of information. The approximation can be made arbitrarily close to the optimal while increasing the computation time of the solution. Specifically, given any positive ϵ, the proposed scheme can be tuned to attain a utility which is at most ϵ less than that of the optimal, and requires a computation time which is polynomial in the number of channels and the degree of this polynomial increases with decrease in ϵ.**

I. Introduction

Future wireless networks will provide each terminal access to a large number of channels. A channel can for example be a frequency in a frequency division multiple access (FDMA) network, or a code in a code division multiple access (CDMA) network, or an antenna or a polarization state (vertical or horizontal) of an antenna in a device with multiple antennas (MIMO). Several existing wireless technologies, e.g., IEEE 802.11a [1], IEEE802.11b [11], IEEE802.11h [2] propose to use multiple frequencies. For example, IEEE 802.11a protocol has 8 channels for indoor use and 4 channels for outdoor use in the 5GHz band, while the IEEE 802.11b protocol has 3 channels in the 2.4 GHz band. The potential deregulation of the wireless spectrum is likely to enable the use of a significantly larger number of frequencies. Due to significant advances in device technology, laptops with multiple antennas (antenna arrays) incorporated in the front lid, and devices with smart antennas have already been developed, and the number of such antennas are likely to significantly increase in near future.

The increase in the number of channels is expected to significantly enhance network capacity and enable several new bandwidth-intensive applications as multiple transmissions can now proceed simultaneously in a vicinity using different channels. Furthermore, the availability of multiple channels substantially enhances the probability (at any given time) of existence of at least one channel with acceptable transmission quality, since the transmission quality of the individual channels stochastically vary with time and location of the users. These benefits can however be realized only if the users can select the channels efficiently using intelligent control mechanisms. Acquiring the information utilized by such control mechanisms often constitutes an important bottleneck. Note that a user can only learn the instantaneous state of a channel by transmitting a control packet in it and subsequently the receiver informs the sender about the quality of the channel in a response packet (e.g., the RTS and CTS packet exchange in IEEE 802.11). The exchange of control packets in this probing process consumes additional energy, and prevents other neighboring users from simultaneously utilizing the channel. Probing a channel is therefore associated with a cost. When the number of available channels is large, the cost incurred in learning the instantaneous transmission qualities of all channels may become prohibitive. We therefore seek to develop a framework for joint optimization of information acquisition and exploitation which in accordance with the cost and the benefits of probing different channels, determines both (a) the amount of information a user must obtain about the instantaneous transmission qualities of the channels at its disposal and also (b) how to select the channels based on the acquired information.

We consider a single sender with access to n channels. The instantaneous transmission qualities of the channels can have K possible values and stochastically vary with time. The statistics of these temporal variations may be different for different channels. Every time the sender probes a channel it learns about the signal to noise ratio and thereby the probability of success in the channel, but also incurs a certain cost c. Before each transmission, the sender needs to determine how many and which channels it will probe and also the sequence in which these channels will be probed (*probing policy*). Note that depending on the available hardware (e.g., availability, or lack thereof, of multiple network interface cards, or compatible transmission circuits to appropriately distribute the power across the antennas), a sender may, or may not, be able to simultaneously transmit in multiple channels. In this paper, we consider the scenario where a sender can transmit in only one channel in a time slot and transmits one packet in each slot. Based on the outcomes of the probes, the sender selects one of the available channels

S. Guha is with the department of Computer and Information Science at the University of Pennsylvania. K. Munagala is with the department of Computer Science at Duke University. S. Sarkar is with the department of Electrical and Systems Engineering at the University of Pennsylvania. Their emails are sudipto@cis.upenn.edu, kamesh@cs.duke.edu, swati@seas.upenn.edu. The authors were supported in part by NSF grants NCR $02-38340$, CCF $04-30376$, CNS $04-543264$ and CNS $05-40347$ and an Alfred P. Sloan Fellowship.

(*channel selection policy*), which need not be those that it has probed.

The sender seeks to maximize a utility function which depends on the probability of success of each transmission and the probing cost incurred before each transmission. Clearly, a meaningful utility function should (a) increase with increase in the former and (b) decrease with increase in the latter. As a starting point, we consider linear functions that satisfy the above criteria. Specifically, we seek to design a jointly optimal probing and channel selection policy that maximizes a system utility which is the difference between the probability of successful transmission and a suitably scaled expected probing cost before each transmission. Loosely, this utility function represents the gain or the profit of the sender if the sender receives credit from the receiver for each packet it delivers successfully and needs to additionally compensate the wireless provider for each probe packet it transmits[1].

We first enumerate the challenges in designing the jointly optimal strategy. The optimal policy needs to probe adaptively, i.e., the result of a probe determines the channels to be probed subsequently. For example, consider channels with 3 possible states $(0, 1, 2)$, each of which is associated with a different transmission quality. Clearly, the probing terminates if a probed channel is in the highest state. Now, let a probed channel be in the intermediate state (state 1). Then the subsequent probes should be limited to channels that have high probabilities of being in the highest state. However, if all channels that have been probed in a slot are in the lowest state, then the channels that have high probabilities of being in the intermediate state may also be subsequently probed. Also, the channel selection decision depends on the outcomes of the probes and the expectation and uncertainty of the transmission quality of the channels that have not been probed. The optimal policy is therefore a decision tree over n variables (Figure 1) – naive computations will require both exponential time and exponential storage space. Next, the policies may depend on the higher order statistics of the channels. This is because the optimum policy may not probe a channel if its quality has a low variance as probing it does not provide significant information but incurs additional cost.

In this paper, we obtain a parameterized probing and channel selection policy whose parameters can be appropriately selected so as to attain any desired tradeoff between performance guarantee and computation time. Specifically, given any $\epsilon > 0$, we obtain a policy that attains a utility which is at most ϵ less than that of the optimal and requires a computation time which is polynomial in the number of channels n, but does not depend on the number of states K. The degree of this polynomial however increases with decrease in ϵ. Our results are somewhat surprising given that optimal solutions for most partial information based control problems turn out to be computationally intractable, and standard approximation techniques either do not provide guaranteeable approximation ratios or require exponential computation times [5]. Our proofs therefore rely on exploitation of specific system characteristics and employ techniques that are not standard in context of stochastic control.

The paper is organized as follows. We review the related literature in Section II. We describe the system model in Section III. We present the probing and channel selection policy, and prove its performance guarantee in Section IV. We conclude in Section V.

II. Related Literature

We first discuss the relation of our problem with some classical problems like the stopping time and multi-armed bandit problems. The most well-researched version of the stopping time problem is a stochastic control problem that optimally selects between two possible actions at any given time: to continue or to stop [7]. Recently, the results for this problem have been used to solve partial information based control problems for statistically identical channels with equal probing costs [17], [22]. Since we consider channels that may have different statistics, the optimal action needs to be selected from multiple options at any given time - the options being (a) whether to continue probing (b) which channel to probe next if the decision is to probe and (c) which channel to transmit if the decision is to stop probing. Thus, the results from the above version of stopping time problem do not apply in our context. The optimal stopping time problem has also been considered in a more general setting where the number of available actions may be more than two; our problem is in fact a special case of this general version (Chapter IV, [5]). In this general case, the process terminates in certain states, which constitute the termination set, and selects the optimal action in other states. But, so far, only certain broad characterizations of the termination set are known in this general case, and the optimal actions when the decision is not to stop are also not known in close form [5]. Thus, these general results do not lead to the optimal policies we are seeking to characterize.

The stochastic multi-armed bandit problem considers a bandit with n arms [10]. The system can try one arm in each slot, and when it tries an arm, it receives a random reward which depends on the state of the arm. The state of an arm changes only when the system tries it. The reward of a system in T slots is the sum of the rewards in each slot. The goal is to maximize the expected reward in T slots. Our problem differs from the above in that (a) the state of a channel can change even when it is not probed or used for transmission and (b) a node can learn the states of multiple channels in an epoch while incurring additional probing costs for learning the state of each additional channel. The adversarial multi-armed bandit problem removes one of the above differences in that it allows the state of an arm to change even when the system does not try it [3]. But, it seeks to optimize the selection under the assumption that the sender uses the same arm in all slots. Note that we allow a sender to probe and transmit in different channels in different slots. In another version of the adversarial multi-armed bandit problem, the goal is to select the arms so as to minimize the

[1] The sender may have to share with the provider part of the credit it receives from the receiver for each successfully delivered packet. Then the credit we are considering here is the credit remaining after the sharing process.

"regret" or the difference in expected reward with the best policy in a collection of a certain number (say N) of given policies. As expected, the regret in T slots increases with increase in both N and T (the regret for the best known policy is $O(\sqrt{nT\ln(N)})$). In our context, the total number of possible probing and channel selection policies that can be used in T slots is large, e.g., the number of deterministic policies is $(n!)^T$. Thus the results available in this context do not apply in our problem, and we use different solution approach and obtain different performance guarantees.

Optimizing the order of evaluation of random variables so as to minimize the cost of evaluation ("pipelined filters") has been investigated in several different contexts like diagnostic tests in fault detection and medical diagnosis, optimizing conjunctive query and joint ordering in data-stream systems, web services [4], [6], [8], [9], [16], [18], [19], [20], [21]. However our work is different from all the above in that, we (a) consider multi-state channel models whereas pipeline filters consider two state models and (b) allow a node to transmit in a channel even if the channel has not been probed. Note that usually two state models can not capture the statistical variations of wireless channels [12]. Both the above generalizations significantly alter the decision issues and the optimal solutions.

Finally, opportunistic selection of channels with complete knowledge of channel states has been comprehensively investigated over the last decade (e.g., [23]). But, in general, the area of partial information based control problems, and in particular the joint optimization of the reward obtained from informed selections and the cost incurred in acquiring the required information, remains largely unexplored in wireless networks. The first results in this area have been obtained in [17], [22], but as mentioned above, they consider only statistically identical channels with equal probing costs. Recent statistical investigations indicate that different channels available to a sender may have different statistics [12]. We now describe our earlier results in the area. We have recently proved that when every channel has two states the joint optimization problem can be solved in polynomial time even when different channels have different statistics and probing costs [14]. In [13], we proved that for channels with multiple states the optimization can be approximated within a factor of $1/2$ using polynomial time algorithms. In another recent submission (technical report [15]) we prove that the joint optimization problem can be approximated within a factor of $4/5$ for channels with arbitrary states using polynomial time algorithms. In all these papers we considered channels with potentially different probing costs and different statistical distributions for the state processes. In the current paper, we focus on the important special case where all channels have equal probing costs, but potentially different distributions for the state processes, and present a joint probing and selection policy that attains any desired tradeoff between approximation guarantees and computation time.

III. SYSTEM MODEL AND PROBLEM DEFINITION

A sender U has access to n channels which are denoted as channels $1, 2, \ldots, n$, each of which has K possible states, $0, \ldots, K-1$. We assume that time is slotted. In any slot channel j is in state i with probability p_{ij} independent of its state in other time slots and the states of other channels in any slot. Without loss of generality, we assume that $p_{K-1j} < 1$ for each j, as otherwise the optimum policy is simply to transmit in j without probing any channel. In every slot, U transmits one data packet in a selected channel. If the channel selected for transmission is in state i, the transmission is successful with probability r_i. Without loss of generality we assume $0 \leq r_0 < r_1 < \cdots < r_{K-1} \leq 1$. For simplicity, we also assume that $r_0 = 0$; all analytical results can however be generalized to the scenario where $r_0 > 0$. Whenever U probes a channel j, it pays a cost of $c \geq 0$. We assume that the probing cost is the same for different channels, as in many cases of practical interest the probing cost is determined by the energy consumed in transmitting the probe packets which is again similar for different channels.

A *probing policy* is a rule that, given the set of channels the sender has already probed in a slot (which would be empty at the beginning of the slot) and the states of the channels probed in the slot, determines (a) whether the sender should probe any more channels and (b) if the sender probes additional channels which channel it should probe next. The sender knows the state of a channel in a slot if and only if it probes the channel in the slot.

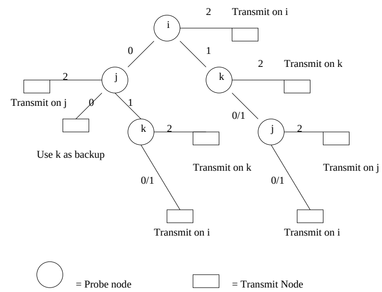
A *selection policy* is a rule that selects a channel for the transmission of a data packet in a slot on the basis of the states of the probed channels, after the completion of the probing process in the slot. The selection policy can select a channel even if it has not been probed in the slot, and in that case, the channel is referred to as a *backup* channel.

The *probing cost* is the sum of the costs of all channels probed in the slot. The probing cost is clearly a random variable that depends on the probing policy and the outcomes of the probes (as the sender may probe subsequent channels depending on the outcomes of the previous probes). The *expected probing cost* is the expectation of this random variable and depends on both the probing policy and the channel statistics.

In any slot, the *transmission reward* is 1 if there is a successful transmission and 0 otherwise. Therefore, the expected transmission reward is r_i in a slot t if U transmits in a channel in state i during t. We sometimes overload the terminology and denote the "reward" of transmitting in state i by r_i. The expected transmission reward of a policy is therefore $\sum_i q_i r_i$ where q_i is the probability that the selection policy decides to use a channel which is in state i; q_i depends on the channel statistics as well as the policy.

The *expected utility* of the sender, denoted simply as *gain*, is the difference between the expected transmission reward, and the probing cost scaled by a factor κ. The gain depends on the probing and selection policies, the channel statistics and the scaling parameter κ. Since κ can be included in the probing cost themselves, we drop this parameter in the remaining discussion without loss of generality.

Problem Definition: Given $\{c_j\}, \{r_i\}$ and $\{p_{ij}\}$, find a probing and selection policy so as to maximize the expected gain. Let OPT denote the optimal policy and G_{OPT} its gain.



(a) Decision Tree

Fig. 1. The figure shows the decision tree for an example policy. A channel is probed at each probe node, and the letter inside it indicates which channel is probed at the node. The numbers next to the branches indicate the outcome of the probe. The number r/s next to a branch indicates that both states r and s of the previously probed channel lead to the same action. For example, the sender first probes channel i. If i is in state 2, it transmits in i. If i is in state 1 and 0, it probes k and j respectively.

Every joint probing and selection policy can be represented by a unique decision tree (Figure 1); we therefore use policies and decision trees interchangeably.

The optimal probing policy does not probe any further in a slot if a probed channel is in state $K-1$. Since channels are temporally independent, the optimal probing and selection strategies in a slot need not depend on the decisions and the observations in other slots. Also, the optimal probing and selection strategies remain the same in all slots, though the specific choices made by each policy may be different in different slots depending on the outcome of the probes. Using these observations, the optimal policy can be computed using a bottom-up dynamic program. However, the computation time for the optimal is $\Omega(K2^n)$ time as the dynamic program has $K2^n$ states, and the storage space is $\Omega(n^K)$.

IV. Arbitrary Tradeoff between Approximation Guarantee and Computation Time

We present a recursive policy, APPROX(ϵ) that given any $\epsilon > 0$ selects its parameters so as to attain a gain of at least $G_{\text{OPT}} - \epsilon$. For any given ϵ, APPROX(ϵ) selects the optimum among a class of recursive policies which is guaranteed to contain at least one policy whose gain is at least $G_{\text{OPT}} - \epsilon$. The size of this class increases with decrease in ϵ, and the search in this class needs to evaluate the gain of each policy in the class. Hence, the computation time of the search increases with decrease in ϵ.

First, consider the case that $\epsilon \geq c$. Note that if a policy A that transmits in a backup channel for certain outcomes of the probes attains a gain of G, then there exists another policy that transmits only in probed channels and attains a gain of at least $G - c$; the second policy is obtained by altering A so as to probe the channel in which it transmits if that channel has not been probed already. Thus, the gain of the policy, OptNoBkup (which we presented in [13]), that attains the maximum gain among all policies that transmit in only probed channels is at most c less than that of the optimum gain. Thus, since $c \leq \epsilon$, OptNoBkup attains the desired gain. For completeness, we present OptNoBkup here.

Definition 1: For $i = 1, \ldots, n$, let $\tilde{r}_i[u] = \frac{\sum_{v=u}^{v=K-1} p_{vi} r_{vi}}{\sum_{v=u}^{v=K-1} p_{vi}}$ and $\tilde{p}_i[u] = \sum_{v=u}^{v=K-1} p_{vi}$. Let $\tilde{r}_0[0] = -1$.

Definition 2: Let $H_u = \phi$ for all $u \geq K$, and for $u = K-1$, down to $u = 1$ $H_u = \left\{ i | i \notin \bigcup_{v:v>u} H_v, \text{ and } \tilde{r}_i[u] - \frac{c}{\tilde{p}_i[u]} > r_{u-1} \right\}$. Assume $c/\tilde{p}_i[u] = +\infty$ when $\tilde{p}_i[u] = 0$.

OptNoBkup

Consider each H_u in decreasing order of u starting from $u = K-1$ down to $u = 1$.

Within each H_u probe in non-increasing order of $\tilde{r}_j[u] - \frac{c}{\tilde{p}_j[u]}$, and stop if any channel is found to be in state u or above (*probing process*).

Transmit in the channel, which is in the highest state, among all probed channels (*selection process*).

We will next show that a recursive policy can be used to obtain the desired approximation guarantees even when $\epsilon < c$. We first introduce a new definition.

Consider a fictitious system $\mathcal{F}(Q, j)$ that can probe or use as backup only channels in the set $\{1, \ldots, n\} \setminus Q$. A user is allowed not to select any channel in this system - its gain then is 0. The *incremental gain* of a policy A, $G_I^A(\mathcal{Q}, j)$, is its expected gain in this fictitious system assuming that the reward obtained by transmitting in a channel that is in state i is $r_i - r_j$. This is the additional gain of a policy that acts in accordance with policy A after probing channels in $\mathcal{Q}$ and observing the highest state of probed channels as j, as compared to selecting a probed channel that is in state j (note that if A does not select any channel then this policy transmits in a channel in $\mathcal{Q}$ that is in state j). The *optimum incremental gain* $G_I^*(\mathcal{Q}, j)$ is the maximum value of the incremental gain for any policy A. Let OptIG$(\mathcal{Q}, j)$ denote the policy that that attains this optimum incremental gain.

Let $k(\epsilon, x) = \lceil \frac{\ln \epsilon/(r_{K-1} - r_x)}{\ln(1-\epsilon/(r_{K-1} - r_x))} \rceil$.

We now show how to design a policy ApproxIntermediate $(\mathcal{Q}, x, \epsilon)$ that attains an incremental gain of at least $G_I^*(\mathcal{Q}, j) - \epsilon$. We use a recursive design and assume that for each $j > x$ and $\mathcal{Q}'$ such that $\mathcal{Q} \subseteq \mathcal{Q}' \subseteq \mathcal{Q} \cup S$, where $S \subseteq \{1, \dots, n\} \setminus \mathcal{Q}$ and $|S| \leq n^{k(\epsilon,x)}$, we have a policy APPROXINTERMEDIATE $(\mathcal{Q}, j, \epsilon/2)$ that attains an incremental gain of at least $G_I^*(\mathcal{Q}', j) - \epsilon/2$. We first introduce the following definition.

Definition 3: Let $\vec{\tau}^{k,\mathcal{Q}}$ be a k-dimensional vector with components in $\{0, \dots, n\} \setminus \mathcal{Q}$ which satisfies the property that if $\tau_i^{k,\mathcal{Q}} = 0$ then $\tau_{i+1}^{k,\mathcal{Q}} = 0$.

SEQUENCE $(\vec{\tau}, \ell, \mathcal{Q}, x, \epsilon)$

Let k_1 be the number of elements in $\vec{\tau}$. If $\vec{\tau}_j > 0$ for all j, $k_2 = k_1$, else $k_2 = \min\{j : \vec{\tau}_j = 0\}$.

From $i = 0$ until $i = k_2 - 1$, probe channels $\vec{\tau}_i$ unless a probed channel is in a state higher than x.

Let $\mathcal{Q}'$ be the set of channels probed in this step. If $\mathcal{Q}' = \phi$, $r_j = -1$. Otherwise, let the highest state of a channel in $\mathcal{Q}'$ be j, and let t be the probed channel which is in state j.

If $\max(r_j, \tilde{r}_\ell) \leq r_x$, transmit in a probed channel that is in state x. Else, if $i = k_2$, (a) if $r_j > \tilde{r}_\ell$ transmit in t and (b) otherwise transmit in ℓ. Else, follow the same decisions as APPROXINTERMEDIATE $(j, \mathcal{Q} \cup \mathcal{Q}', \epsilon/2)$.

Definition 4: Let SEQUENCE$^*(\mathcal{Q}, x, \epsilon)$ be the policy that attains the maximum gain among SEQUENCE $(\vec{\tau}, \ell, x, \mathcal{Q}, \epsilon)$ for different $\vec{\tau}^{k(\epsilon,x),\mathcal{Q}\cup\{\ell\}}$ for all possible $\ell \in \{0, \dots, n\} \setminus \mathcal{Q}$.

The following lemma proves that if $c > \epsilon$ SEQUENCE$^*(\mathcal{Q}, x, \epsilon)$ attains an incremental gain of at least $G_I^*(\mathcal{Q}, j) - \epsilon$. Thus, APPROXINTERMEDIATE $(\mathcal{Q}, x, \epsilon)$ is in fact SEQUENCE$^*(\mathcal{Q}, x, \epsilon)$, and hence APPROXINTERMEDIATE $(\mathcal{Q}, x, \epsilon)$ can be constructed in a recursive manner provided we know APPROXINTERMEDIATE $(\mathcal{Q}', K-2, \epsilon/2^{K-2-x})$ for certain subsets $\mathcal{Q}'$. Since $G_I^*(\phi, 0)$ is the optimum gain, the recursive construction will in turn lead to the desired overall approximation guarantee.

Lemma 4.1: Let $c > \epsilon$. Suppose for each $j > x$ and $\mathcal{Q}'$ such that $\mathcal{Q} \subseteq \mathcal{Q}' \subseteq \mathcal{Q} \cup S$, where $S \subseteq \{1, \dots, n\} \setminus \mathcal{Q}$ and $|S| \leq n^{k(\epsilon,x)}$, a policy APPROXINTERMEDIATE $(\mathcal{Q}', j, \epsilon)$ attains an incremental gain of at least $G_I^*(\mathcal{Q}, j) - \epsilon$. Then SEQUENCE$^*(\mathcal{Q}, x, \epsilon)$ attains an incremental gain which is at least $G_I^*(\mathcal{Q}, j) - 2\epsilon$.

Proof: Clearly, at every node in its decision tree OPTIG$(\mathcal{Q}, x)$ takes the same decisions irrespective of the state of the probed channel provided the state is x or lower. So, the sub-tree starting from the root node in which the probed channels are in state x or a lower state corresponds to a path, which we refer to as the x-path. Note that the incremental gain of SEQUENCE$^*(\mathcal{Q}, x, \epsilon)$ is at most ϵ less than that of any policy in the fictitious system $\mathcal{F}(\mathcal{Q}, x)$ that probes at most $k(\epsilon, x)$ channels in its x-path. Thus, if OPTIG$(\mathcal{Q}, x)$ probes $k(\epsilon, x)$ or fewer channels in its x-path, then clearly the incremental gain of SEQUENCE$^*(\mathcal{Q}, x, \epsilon)$ is at least $G_I^*(\mathcal{Q}, j) - \epsilon$. The result follows. Otherwise, OPTIG $(\mathcal{Q}, x)$ probes at least $k(\epsilon, x) + 1$ channels in its x-path. Now, change OPTIG $(\mathcal{Q}, x)$ so that it terminates the probing process after probing the first $k(\epsilon, x)$ channels in the x-path and takes the same decision as OPTIG $(\mathcal{Q}, x)$ does at the end of its x-path. Let the new policy be denoted as OPTMOD. Since OPTMOD probes $k(\epsilon, x)$ channels in its x-path and can be used in $\mathcal{F}(\mathcal{Q}, x)$, its incremental gain exceeds that of SEQUENCE$^*(\mathcal{Q}, x, \epsilon)$ by at most ϵ. We now show that the gain of OPTMOD is at most ϵ less than that of OPT. The result follows.

Let the probability that all channels probed by OPTMOD in its x-path are in state x or in a lower state be α. Clearly, the difference between the incremental gains of OPTMOD and OPTIG $(\mathcal{Q}, x)$ is at most $\alpha(r_{K-1} - r_x)$. We now show that $\alpha \leq \epsilon/(r_{K-1} - r_x)$. Note that if OPTIG probes a channel i in its x-path, then $\sum_{k=x+1}^{K-1} p_{ki}(r_k - r_x) - c \geq 0$. Thus, if OPTMOD probes a channel i in its x-path, then $\sum_{k=x+1}^{K-1} p_{ki}(r_k - r_x) - c \geq 0$. Thus, $\sum_{k=x+1}^{K-1} p_{ki}(r_k - r_x) \geq \epsilon$. Thus, $(r_{K-1} - r_x) \sum_{k=x+1}^{K-1} p_{ki} \geq \epsilon$. Thus, $r_{K-1} - r_x \geq \epsilon$. Also, $(1 - \sum_{k=0}^{x} p_{ki})(r_{K-1} - r_x) \geq \epsilon$. Hence, $\sum_{k=0}^{x} p_{ki} \leq 1 - \epsilon/(r_{K-1} - r_x)$. Since OPTMOD probes $k(\epsilon, x)$ channels in the 0-path, $\alpha \leq (1 - \epsilon/(r_{K-1} - r_x))^{k(\epsilon,x)}$. Since $k(\epsilon, x) \geq \frac{\ln \epsilon/(r_{K-1} - r_x)}{\ln(1 - \epsilon/(r_{K-1} - r_x))}$, $(1 - \epsilon/(r_{K-1} - r_x))^{k(\epsilon,x)} \leq \epsilon/(r_{K-1} - r_x)$. ∎

Now, note that since $r_{K-1} \leq 1$, $k(\epsilon, x) \leq g(\epsilon)$ where $g(\epsilon) = \lceil (1/\epsilon) \ln(1/\epsilon) \rceil$.

Now, given lemma 4.1, the recursive design will be complete once we determine APPROXINTERMEDIATE $(\mathcal{Q}, K-2, \epsilon)$ for certain subsets $\mathcal{Q}$ of $\{1, \dots, n\}$. Note that the channels in the fictitious system $\mathcal{F}(\mathcal{Q}, K-2)$ effectively have two states $K-2, K-1$. We have presented a policy that determines the optimum gain in any system where each channel has 2 states in polynomial time [14]. OPTIG$(\mathcal{Q}, K-2)$, and hence APPROXINTERMEDIATE $(\mathcal{Q}, K-2, \epsilon)$ for any ϵ, can be obtained in polynomial time by minor modification of the policy presented in [14]. The modified version is presented as TWOSTATEOPT$(\mathcal{Q})$.

TWOSTATEOPT $(\mathcal{Q})$

For each $\ell \in \{0, \dots, n\} \setminus \mathcal{Q}$, determine TWOSTATEOPT$(\mathcal{Q}, \ell)$ as follows.

1) Let $\mathcal{S}_\ell = \{i : p_{K-1i}(r_{K-1} - \max(\tilde{r}_\ell[0], r_{K-2})) > c_i\} \setminus \mathcal{Q}$.
2) Probe all channels in S_ℓ in non-increasing order of p_{K-1i}/c_i and stop if one channel is in state $K-1$.
3) If a probed channel is in state $K-1$ transmit in it, else if $\tilde{r}_\ell[0] > r_{K-2}$ transmit in ℓ, else transmit in a probed channel that is in state $K-2$.

Select TWOSTATEOPT$(\mathcal{Q}, \ell)$ that has the maximum gain among all $\ell \in \{0, \dots, n\} \setminus \mathcal{Q}$.

It now follows that APPROXINTERMEDIATE $(\mathcal{Q}, j, \epsilon)$ can be computed using the following recursive approach (the recursion occurs as SEQUENCE$^*(\mathcal{Q}, j, \epsilon)$ can be computed using APPROXINTERMEDIATE $(\mathcal{Q}', i, \epsilon/2)$ for certain subsets $\mathcal{Q}'$ and $i > j$.

APPROXINTERMEDIATE $(\mathcal{Q}, j, \epsilon)$

If $j = K-2$, use TWOSTATEOPT$(\mathcal{Q})$.
If $j < K-2$, use SEQUENCE$^*(\mathcal{Q}, j, \epsilon)$.

We now present the policy APPROX (ϵ) which attains a gain of at least $G_{\text{OPT}} - \epsilon$.

> APPROX (ϵ)
>
> If $c \leq \epsilon$, use OPTNOBKUP.
>
> Otherwise, for $j = K - 2$ to $j = 1$
> for all subsets $\mathcal{Q}$ such that $|\mathcal{Q}| \leq jg(\epsilon/2^j)$
> compute APPROXINTERMEDIATE($\mathcal{Q}, j, \epsilon/2^j$).
> Compute and use APPROXINTERMEDIATE($\phi, 0, \epsilon$).

Theorem 4.2: The gain of APPROX (ϵ) is at least $G_{\text{OPT}} - \epsilon$.

Proof: As discussed before, if $c \leq \epsilon$, the gain of OPTNOBKUP is at most ϵ less than that of the optimal policy. Let $c > \epsilon$. From lemma 4.1 and since TWOSTATEOPT($\mathcal{Q}$) is OPTIG($\mathcal{Q}, K - 2$), the incremental gain of APPROXINTERMEDIATE$(0, \phi, \epsilon)$ is at most ϵ less than that of the optimum incremental gain $G_I^*(\phi, 0)$. Note that the incremental gain of any policy in the fictitious system $\mathcal{F}(\phi, 0, \epsilon)$ is the same as its gain in the overall system, and any such policy always selects a channel. Thus, the incremental gain of APPROXINTERMEDIATE$(0, \phi, \epsilon)$ equals its gain, and $G_I^*(\phi, 0) = G_{\text{OPT}}$. The result follows. ■

We now evaluate the computation time for APPROX (ϵ). Now, APPROX (ϵ) can be computed in $O(g(\epsilon/2^K)n^{g(\epsilon/2^K)(K+2)+1})$ time. Given a positive ϵ, the computation time is therefore polynomial in n but exponential in K. This is usually acceptable since the number of states K is small. Nevertheless, we now describe how the computation time can be made independent of K. First, divide $[0, r_{K-1}]$ in disjoint intervals of size $\epsilon/2$. Then, consider a new system where the probability of success in each state i equals $(\epsilon/2)\lfloor 2r_i/\epsilon \rfloor$. This new system effectively consists of $K_0 = 2r_{K-1}/\epsilon \leq 2/\epsilon$ states. In this system, APPROX ($\epsilon/2$) approximates the optimum gain within an additive factor of $\epsilon/2$. The gain of the optimum policy in this system is at least $G_{\text{OPT}} - \epsilon/2$. Thus, APPROX ($\epsilon/2$) computed in this system attains a gain of at least $G_{\text{OPT}} - \epsilon$. Note that the time required for computing APPROX ($\epsilon/2$) in this system is $O(g(\epsilon/2^{2/\epsilon})n^{2g(\epsilon/2^{2/\epsilon})(1/\epsilon+1)+1})$ irrespective of the number of states in the original system.

Finally, we comment on the design of an approximation algorithm that approximates the optimum gain within a multiplicative factor of ϵ. Let G' be the gain of the policy APPROXBKUP proposed in [13] that is guaranteed to attain a gain of at least half that of the optimum policy. APPROX ($\epsilon G'/2$) computed in the above system attains a gain of at least $G_{\text{OPT}} - \epsilon G'$, which is at least $(1 - \epsilon)G_{\text{OPT}}$. But, the computation time of APPROX ($\epsilon G'/2$) is $O(g(\epsilon G'/2^{2/\epsilon G'})n^{2g(\epsilon/2^{2/\epsilon G'})(1/\epsilon G'+1)+1})$, and therefore increases with decrease in G'. The computation time is still a polynomial in n once ϵ is specified provided G_{OPT} is $\Omega(\epsilon^k)$ for some constant k (as then G' is $\Omega(\epsilon^k)$). Thus, the computation time becomes unacceptable only when G_{OPT} is $o(\epsilon^k)$ for all constants k. But, in systems with such low gain, the channel qualities are so poor that from a practical perspective any optimization becomes useless.

V. CONCLUSION

The area of optimization of joint information acquisition and exploitation strategies in wireless networks poses several open problems which have received limited attention until recently. In this paper, we have provided a policy that attains arbitrary desired tradeoffs between approximation guarantee and computation time. The performance guarantees have however been obtained for the special case of equal probing costs and temporally and spatially independent channel state processes. Generalizing the results for networks where these assumptions do not hold constitute interesting problems for future research.

REFERENCES

[1] IEEE 802.11a Working Group. *Wireless LAN Medium Acces Control(MAC) and Physical Layer(PHY) Specifications-Amendment 1:High speed Physical Layer in the 5GHz Band*, 1999.

[2] IEEE 802.11h Working Group. *Wireless LAN Medium Acces Control(MAC) and Physical Layer(PHY) Specifications-Amendment 5:Spectrum and Transmit Power Management Extensions in the 5GHz Band in Europe*, 2003.

[3] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. Schapire. The non-stochastic multi-armed bandit problem. *SIAM Journal for Computing*, 32(1):48–77, November 2002.

[4] S. Babu, R. Motwani, K. Munagala, I. Nishizawa, and J. Widom. Adaptive ordering of pipelined stream filters. In *2004*, pages 407–418, June 2004.

[5] D. P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, 2nd edition, 2001.

[6] J. Burge, K. Munagala, and U. Srivastava. Ordering pipelined query operators with precedence constraints. *Submitted*, 2005.

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

[8] E. Cohen, A. Fiat, and H. Kaplan. Efficient sequences of trials. In *2003*, 2003.

[9] U. Feige, L. Lovász, and P. Tetali. Approximating min-sum set cover. *Algorithmica*, 2004.

[10] J. Gittins and D. Jones. A dynamic allocation index for the sequential allocation of experiments. *Progress in Statistics*. edited by J. Gani.

[11] IEEE 802.11 Working Group. *Wireless LAN Medium Acces Control(MAC) and Physical Layer(PHY) Specifications*, 1997.

[12] R. Guha and S. Sarkar. Characterizing temporal snr variation in 802.11 networks. *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC)*, April 3-6 2006.

[13] S. Guha, K. Munagala, and S. Sarkar. Jointly optimal transmission and probing strategies for multichannel wireless systems. In *Proceedings of Conference on Information Sciences and Systems*, Princeton, NJ, March 22-24 2006.

[14] S. Guha, K. Munagala, and S. Sarkar. Optimizing transmission rate in wireless channels using adaptive probes (poster paper). *Proc. of SIGMETRICS/Performance*, 2006.

[15] S. Guha, K. Munagala, and S. Sarkar. Performance guarantees through partial information based control in multichannel wireless networks. Technical report, Univ. of Pennsylvania, Philadelphia, PA, Jul 2006. http://www.seas.upenn.edu/∼swati/publication.htm.

[16] J. Hellerstein. Optimization techniques for queries with expensive methods. *ACM Transactions on Database Systems*, 23(2):113–157, 1998.

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

[18] H. Kaplan, E. Kushilevitz, and Y. Mansour. Learning with attribute costs. In *STOC '05: Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pages 356–365, 2005.

[19] M. S. Kodialam. The throughput of sequential testing. *Lecture Notes in Computer Science*, 2081:280–292, 2001.

[20] K. Munagala, S. Babu, R. Motwani, and J. Widom. The pipelined set cover problem. *Proc. Intl. Conf. Database Theory*, 2005.

[21] K. Munagala, U. Srivastava, and J. Widom. Optimization of continuous queries with shared expensive filters. *Submitted*, 2005.

[22] A. Sabharwal, A. Khoshnevis, and E. Knightly. Opportunistic spectral usage: Bounds and a multi-band csma/ca protocol. *IEEE/ACM Transactions on Networking*, 2006. Accepted for publication.

[23] A. Stolyar, S. Shakkottai, and R. Srikant. Pathwise optimality of the exponential scheduling rule for wireless channels. *Advances in Applied Probability*, 36(4):1021–1045, 2004.