To Stay Or To Switch: Multiuser Dynamic Channel Access

Yang Liu and Mingyan Liu
Electrical Engineering and Computer Science
University of Michigan, Ann Arbor
{youngliu, mingyan}@umich.edu

***Abstract*—In this paper we study opportunistic spectrum access (OSA) policies in a multiuser multichannel random access setting, where users perform channel probing and switching in order to obtain better channel condition or higher instantaneous transmission quality. However, unlikely many prior works in this area, including channel probing and switching policies for a single user to exploit spectral diversity, and probing and access policies for multiple users over a single channel to exploit temporal and multiuser diversity, in this study we consider the collective switching of multiple users over multiple channels. In addition, we consider finite arrivals, i.e., users are not assumed to always have data to send and demand for channel follow a certain arrival process. Under such a scenario, the users' ability to opportunistically exploit temporal diversity (the temporal variation in channel quality over a single channel) and spectral diversity (quality variation across multiple channels at a give time) is greatly affected by the level of congestion in the system. We investigate the optimal decision process in this case, and evaluate the extent to which congestion affects potential gains from opportunistic dynamic channel switching.**

I. Introduction

Dynamic and Opportunistic Spectrum Access (OSA) policies have been very extensively studied in the past few years against the backdrop of spectrum open access as well as advances in ever more agile radio transceivers. At the heart of opportunistic spectrum access is the idea of improving spectrum efficiency through the exploitation of *diversity*.

Within this context there are three types of diversity gains commonly explored. The first is *temporal diversity*, where the natural temporal variation in the wireless channel causes a user to experience or perceive different transmission conditions over time even when it stays on the same channel, and the idea is to have the user access the channel for data transmission when the condition is good, which may require and warrant a certain amount of waiting. Studies like [3] investigate the tradeoff involved in waiting for a better condition and when is the best time to stop.

The second is *spectral diversity*, where different channels experience different temporal variations, so for a given user at any given time a set of channels present different transmission conditions. The idea is then to have the user select a channel with the best condition at any given time for data transmission, which typically involves probing multiple channels to find out their conditions. Protocols like [5] does exactly this, and studies like [1], [14] further seek to identify the best sequential probing policies using a decision framework.

This work was partially supported by the NSF under grants CIF-0910765 and CNS 1217689, and the ARO under grant W911NF-11-1-0532.

The third is *user diversity* or *spatial diversity*, where the same frequency band at the same time can offer different transmission qualities to different users due to their difference in transceiver design, geographic location, etc. The idea is to have the user with the best condition on a channel use it. This diversity gain can be obtained to some degree by using techniques like stopping time rules whereby a user essentially judges for itself whether the condition is sufficiently good before transmitting, which comes as a byproduct of utilizing temporal diversity.

We note that the above forms of diversities are often studied in isolation. For instance, temporal diversity is studied in a multiuser setting but with a single channel in [12], [15]; spectral diversity is analyzed for a single user in [11], among others.

As the number of users and their traffic volume increase in such a multi-channel system, one would expect their ability to exploit the above diversity gains to decrease significantly due to the increased overhead, e.g., the time it takes to perform channel sensing or the time it takes to regain access right, or increased collision due to channel switching. This overhead has been captured in the form of penalty cost in prior work such as [11], but is often assumed to be independent of the traffic volume existing in the system.

With the above in mind, in this paper we set out to study opportunistic spectrum access policies in a multiuser multi-channel random access setting, where users are not assumed to always have data to send, demand for channel follows a certain arrival process, and collision and competition times are taken into account. Our focus is on the effect of collective switching decisions by the users, and how their decision process, in particular their channel switching decisions, are affected by increasing congestion levels in the system.

Toward this end we characterize the nature of an optimal access policy and identify conditions under which channel switching actually results in transmission gain (e.g. in terms of average data rate or throughput). Our qualitative conclusion, not surprisingly, is that with the increase in user/data arrival rate, the average throughput decreases and a user becomes increasingly more reluctant to give up a present transmission opportunity in hoping for better condition later on or in a different channel. Quantitatively we present algorithms that

978-1-4673-5946-7/13/$31.00

calculate optimal switching decisions and analyze the stability of the overall system.

The remainder of this paper is organized as follows. In the Section II, related works are present. The system model is given in Section III. In Section IV, we model each channel's evolution as an IID process and analyze the properties of an optimal stopping/switching rule. Numerical results are given in Section V, and Section VI concludes the paper.

II. Related work

Opportunistic Spectrum Access(OSA) has been quite extensively studied in recent years; it aims at various diversity harvesting with the objective of improving spectrum efficiency. Example include [10], where centralized scheduling strategies are examined for a class of OSA problems, and [15], where temporal diversity is used in a multi-user wireless network and optimal stopping policies are developed. In particular, using optimal stopping theories [4], optimal strategies for different types of user are derived, including selfish and collaborate users. In [12] a distributed opportunistic scheduling problem for ad-hoc communications under delay constraints is considered. The above works consider only multi-user and temporal diversities but not spectral diversity.

In [11] authors exploit spectral diversity in OSA for the single user with sensing errors. The user's average throughput is maximized under the optimal policy. However, in this framework the multi-channel overhead is captured by a generic penalty on each channel switching. This becomes insufficient in a multi-user setting as such overhead will obviously depend on the level of congestion in the system that results in different amount of collision and the time it takes to regain access to a channel. In [5] an opportunistic auto rate multi-channel MAC protocol MOAR is presented to exploit spectral diversity for a multi-channel multi-rate IEEE 802.11-enabled wireless ad hoc network. However, this scheme does not allow parallel use of multiple channels by different users due to its reservation mechanism. Other works that study multi-channel access by a single user include [1]–[3], [6], [7], [13].

III. Model, assumptions and preliminaries

A. Model and assumptions

Consider a wireless system with N channels indexed by the set $\Omega = \{1, 2, ..., N\}$. We associate each channel with a reward of transmission (e.g., transmission rate) X^j, which is a random variable with distribution characterized by $f_{X^j}(x)$. There are m cognitive users (or radio transceivers) each equipped with a single transmitter attempting to send data to a base station. Our model also captures direct peer-to-peer communication, where m *pairs* of users communicate and each pair can rendezvous and perform channel sensing and switching together through the use of a control channel [8]. However, for simplicity of exposition, for the rest of the paper we will take the view of m users transmitting to a base station. We will assume these m users are within a single interference domain, so that at any given time each channel can only be occupied by one user. Considering spatial reuse will make the problem considerably more challenging and remains an interesting direction of future research.

We consider discrete time with a suitably chosen time unit, and with all other time values integer multiples of this underlying (and possibly very small) unit. We will assume that the channel conditions over time form an IID process defined on this time unit. Conditions in different channels are independent and are in general not identically distributed. Parallel and similar results may be obtained for channels described by Markovian models, though the technical details are quite different. [1]

The system operates in a way similar to a multi-channel random access network like IEEE 802.11, with the following modifications related to dynamic and opportunistic channel access. Each user has a pre-assigned (or self-generated) random sequence of channels; this sequence determines in which order the user performs channel switching, an approach similar to that used in [11]. Each time a user enters a new channel, it must perform carrier sensing and compete for access as in a regular 802.11 channel. As soon as it gains the right to transmit, it finds out the instantaneous channel quality it would get if it transmits immediately. Upon finding out the channel condition, this user faces the following choices:

1) Transmit on the current channel right away. Intuitively this happens if the current channel condition is deemed sufficiently good. This action will be referred to as *STOP*.
2) Forego this transmission opportunity, presumably due to poor channel condition, but remain on the same channel and compete for access again in the near future hoping to come across better condition then. This happens if the current channel condition is poor, so the user will risk waiting for possibly better condition later. This action will be referred to as *STAY*.
3) Give up the current channel and switch to the next one on its sequence. This happens if the current channel condition is poor, and the prospect of staying on the same channel to wait for better condition later is not as good as switching to the next channel. This action will be referred to as *SWITCH*.

Note that option (1) above allows the system to exploit both multiuser diversity (the transmission opportunity is given to another user in the random access) and temporal diversity (the user in question waits for better condition to appear in time), while option (3) allows the system to exploit spectral diversity as users seek better conditions on other channels. These options, in particular (1) and (2) are similar to those used in a stopping time framework, e.g., [15].

In the above decision process a user is not allowed *recall*, i.e., once the user decides to leave a channel it cannot use the channel for transmission without going through carrier sensing and random access competition again. More importantly from a technical point of view, the user cannot claim the same channel condition once it returns to a previously visited channel. As we

[1] A more complete version including analysis over slow changing model with Markovian assumptions can be found in our technical report.

shall see, due to the IID assumption, under an optimal policy, once a user leaves a channel it will never return.

Once a user gets the right to transmit on a certain channel, it can transmit for a period of T time units, which is a constant. For simplicity a single time unit is also assumed to be the amount of time to transmit a control packet.

B. Capturing congestion with a "mean field" approach

As mentioned earlier our focus in this paper is on understanding how the users' channel access decision process is affected by increasing traffic load or congestion in the system. To model this we will first take the view of a single user, and introduce user arrival rates in each channel as well as the amount of delay involved in STAY and SWITCH as parameters that need to be taken into consideration in its decision process. Note that these parameter values are the result of the collective switching decisions of all users, and therefore cannot be obtained prior to defining the switching policies. Indeed later on we show that the system under the optimal switching policy converges and that these parameters have well-defined averages, thereby justifying such an assumption. In other words, policies derived under the assumption that these parameters have well-defined averages lead to a stable system with well-defined averages for these parameters. This not unlike the Markov mean field approach where a single user operates against a background formed by all other users in a system over which this single user has no control or influence. In practice these values can be obtained empirically through learning.

More specifically, we assume users with transmission needs arrive at a given channel (either as their random starting position or as a result of switching from other channels) as a Poisson process, with the rate vector given by $\mathbf{G} = [G_1, G_2, ..., G_N]$ and a sum rate $\sum_{i=1}^{N} G_i = G$. At any given time a user may or may not have data to send, so the rate G is the aggregate rate of data arrival from all users.

The level of congestion on any channel is captured by two parameters. The first is the average *contention delay* t_j^r on channel j, which is the average time from carrier sense to gaining the right to transmit on channel j. The more competing users there are on channel j, the higher this quantity is. The second is the average *switching delay* t_j^c of channel j, which is the time from a user switching into channel j (from another channel) to its gaining the right to transmit on channel j. Compared to t_j^r, the switching delay includes the additional time it takes for the radio to perform channel switching. We adopt the following two natural assumptions on these quantities.

Assumption 1: Both t_j^r and t_j^c are non-decreasing functions of arrival rate G_j, $\forall j \in \Omega$.

Assumption 2: Both t_j^r and t_j^c are non-decreasing functions of the data transmission time T, $\forall j \in \Omega$.

Remark 3.1: We will show later that these assumptions are in general true under the optimal access policies derived in the next section. It is intuitively clear that with increasing arrival rate and time for data transmission, the contention time increases in general.

C. Problem formulation

For simplicity and without loss of generality, for the single user under consideration we will relabel the channels in its sequence in ascending order: $1, 2, \cdots, N$. We now define the following rate-of-return problem with the objective of maximizing the effective data rate over one successful data transmission.

Specifically, let π denote a policy $\pi = \{\alpha_1, \alpha_2, \cdots \alpha_{\gamma(\pi)}\}$, where α_k denotes the k-th actions taken, $\alpha_k \in \{\text{STAY, SWITCH}\}$, $k = 1, \cdots, \gamma(\pi) - 1$, and $\alpha_{\gamma(\pi)} = \text{STOP}$. $\gamma(\pi)$ is the stopping time at which this decision process terminates with a transmission action. Note that an action is only taken upon gaining the right to transmit in a channel.

Let X_k^π denote the data rate obtained during the k-th decision epoch (or stage) under policy π. If under policy π a channel is not used in the k-th stage (i.e., $\alpha_k = $ SWITCH or STAY for some k) then $X_k^\pi = 0$. Note that by this definition only the last decision epoch results in a positive data rate. Let T_k^π denote the amount of time spent during decision epoch k under policy π. The goal is to maximize the effective rate over the duration of this decision process given by:

$$J^* = \max_{\pi \in \Pi} E\{\frac{\sum_{k=1}^{\gamma(\pi)} X_k^\pi \cdot T}{\sum_{k=1}^{\gamma(\pi)} T_k^\pi}\} \tag{1}$$

with Π denoting the admissible set of policies.

We begin by making an important observation of the optimal decision process, which is that once we leave a channel we will never return due to the IID assumption. Note that when we decide to leave a channel, say i, it is because its projected reward is less than the projected reward from the next channel, $i+1$, subject to the difference in delay, t_i^r vs. t_{i+1}^c. Because of the IID assumption, both projected rewards are essentially their statistical means which do not change over time. Therefore if we decide to leave for a better channel it cannot be optimal to ever return.

This maximization problem can be solved by using dynamic programming. Here we denote the value function at channel i, time t under the observed channel state x by $V_{i,t}(x)$; the value function $V_{i,t}(x)$ represents the maximum average throughput obtainable in stage i, time t facing current channel state x (current transmission rate). For stage $i < N$, the value function is the maximum over all possible actions, given as follows:

$$\begin{aligned} V_{i,t}(x) = \max\{X_t^i(x), &\frac{T}{T+t_i^r} E\{V_{i,t+1}(y)|x\}, \\ &\frac{T}{T+t_{i+1}^c} E\{V_{i+1,t+1}(y)|x\}\} \qquad (2) \\ = \max\{X_t^i(x), &\frac{T}{T+t_i^r} E\{V_{i,t+1}\}, \\ &\frac{T}{T+t_{i+1}^c} E\{V_{i+1,t+1}\}\} \qquad (3) \end{aligned}$$

where the second equality is due to the IID nature of the channel condition evolution and thus $E\{V_{i,t+1}(y)|x\}$ and $E\{V_{i+1,t+1}(y)|x\}$ are both independent of the time index t and state x. The first term is the transmission rate realized if we transmit in the current stage, and the second term indicates the expected throughput we get if we decide to give up the current transmission opportunity but stay in the same channel. The last term models the expected average throughput if we decide to leave the channel and explore the next. For $i = N$, channel switching is no longer an option and the value function is given by

$$V_{N,t}(x) = \max\{X_t^N(x), \frac{T}{T+t_N^r}E\{V_{N,t+1}\}\} \ , \tag{4}$$

In the next subsection, we give a quantitative analysis of this value function.

IV. Characterizing the optimal access policy

A. Uniqueness of optimal stopping rule

As discussed above, in our model after re-competition time t_i^r, the new value function for a single user is independent of previous actions; in addition, channels' characteristics are independent of each other. We can thus simplify the dynamic programming problem as follows:

$$V_{i,t}(x) = \max\{X_t^i(x), \frac{T}{T+t_i^r}E\{V_i\}, \frac{T}{T+t_{i+1}^c}E\{V_{i+1}\}\} \tag{5}$$

Note that $\frac{T}{T+t_{i+1}^c}E\{V_{i+1}\}$ (discounted reward from next stage $i+1$) is essentially a constant (only depends on i). We make the following substitution

$$\hat{X}_t^i(x) = \max\{X_t^i(x), \frac{T}{T+t_{i+1}^c}E\{V_{i+1}\}\} \ . \tag{6}$$

We then have the following equivalent dynamic program:

$$V_{i,t}(x) = \max\{\hat{X}_t^i(x), \frac{T}{T+t_i^r}E\{V_i\}\} \ . \tag{7}$$

Next we prove the threshold property of each stage's decision process.

Theorem 4.1: Under certain arrival rate vector, the optimal action at stage i of deciding between {STOP, SWITCH} and {STAY} is given by a stopping rule, i.e., the state space of the channel condition can be divided into a stopping set Δ^s and continuation set Δ^c, such that whenever channel condition is observed to be in either set above, the corresponding action is taken. Furthermore, these two sets are given by the following threshold property, i.e., the stopping set at stage i is given by

$$\Delta_i^s = \{x : \hat{X}_t^i(x) \geq x^*\}, \tag{8}$$

and the threshold x^* is unique.

Proof. First we prove the existence of the threshold. From [4] we know that an optimal stopping rule exists if the following two conditions are satisfied:

1. $E\{\sup_t w_t\} < \infty$.
2. $\lim_{t\to\infty} w_t = -\infty, a.s.$

As in our case, we have

$$w_t = \max\{X_t^i(x), \frac{T}{T+t_{i+1}^c}E\{V_{i+1}\}\}T - \lambda(t \cdot t_i^r + T) \tag{9}$$

As we assume $|X_t^i(x)| < \infty$; and we know $\frac{T}{T+t_{i+1}^c}E\{V_{i+1}\}, T$ are both finite. We know the second condition is satisfied easily. Similarly, we have

$$w_t^{'}(x) = \max\{X_t^i(x), \frac{T}{T+t_{i+1}^c}E\{V_{i+1}\}\}T - \lambda T \ . \tag{10}$$

As $X_t^i(x), \frac{T}{T+t_{i+1}^c}E\{V_{i+1}\}, \lambda, T$ are finite, $w_t^{'}(x)$ is finite. Therefore we have

$$E\{\max\{w_t^{'}(x), 0\}\} < \infty, E\{(\max\{w_t^{'}(x), 0\})^2\} < \infty \ . \tag{11}$$

Meanwhile, $w_t^{'}(x)$'s are IID, thus using [4], we have

$$E\{\sup_t w_t\} = E\{\sup_t w_t^{'}(x) - t \cdot \lambda t_r\} < \infty \ . \tag{12}$$

Existence is thus proved. ∎

Next we prove the *uniqueness* of the threshold.

According to the principle of optimality in Chapter 2 of [4], the optimal stopping rule of the transformed problem is given by

$$\Delta_i^s = \{x : \hat{X}_t^i(x) \geq V^*\} \tag{13}$$

Here V^* denotes the expected return from an optimal stopping rule; it satisfies the following optimality equation

$$V^* = E\{\max\{\hat{X}_t^i, V^*\} - \lambda t_i^r\} \tag{14}$$

or

$$E\{\max(\hat{X}_t^i - V^*, 0)\} = \lambda t_i^r \tag{15}$$

Next we introduce the following lemma.

Lemma 4.2: Under certain arrival rate vector, the optimal threshold x^* is an unique solution to

$$E[\hat{X}_t^i - x]^+ = \frac{x \cdot t_i^r}{T} \tag{16}$$

Proof. From the optimal stopping rule, we know the λ which gives V^* 0 will be the solution. Thus we have

$$E\{(\max\{X_t^i, \frac{T}{T+t_{i+1}^c}E\{V_{i+1}\}\}T - \lambda T)^+\} = \lambda t_i^r \tag{17}$$

Denote $c_i = \frac{T}{T+t_{i+1}^c}E\{V_{i+1}\}$. Then above formula can be expressed as

$$\begin{aligned} &E(\max\{X_t^i T - \lambda T, 0\}|X_t^i > c_i) \cdot P(X_t^i > c_i) \\ &+ E(\max\{c_i T - \lambda T, 0\}|X_t^i \leq c_i) \cdot P(X_t^i \leq c_i) = \lambda t_r \end{aligned} \tag{18}$$

If there exists a solution with $\lambda < c_i$, we have the following

$$\int_{c_i}^{\infty}(x-\lambda)f_{X_t^i}(x)dx + (c_i - \lambda) \cdot P(X_t^i \leq c_i) = \lambda \bar{t}_i^r$$

Here define normalized re-competition time $\bar{t}_i^r = \frac{t_i^r}{T}$. Above gives us

$$\lambda^* = \frac{\int_{c_i}^{\infty} x f_{X^i}(x)dx + c_i \cdot P(X_t^i \leq c_i)}{1 + \bar{t}_i^r} \tag{19}$$

On the other hand, if there exists a solution with $\lambda^* \geq c_i$, we have

$$\int_{\lambda^*}^{\infty} (x-\lambda) f_{X^i}(x) dx = \lambda^* \bar{t}_i^r$$

which gives us

$$\lambda^* = \frac{\int_{\lambda^*}^{\infty} x f_{X^i}(x) dx}{P(X_t^i \geq \lambda^*) + \bar{t}_i^r} \tag{20}$$

For the $\lambda^* < c_i$ case, we see that there could be at most one solution as $\frac{\int_{c_i}^{\infty} x f_{X^i}(x)dx + c_i \cdot P(X_t^i \leq c_i)}{1+\bar{t}_i^r} < c_i$ may not hold.

For $\lambda^* \geq c_i$ case, rewrite (20) we have

$$\lambda^* \bar{t}_i^r = \int_{\lambda^*}^{+\infty} (x-\lambda^*) f_{X^i}(x) dx \tag{21}$$

The LHS is an increasing function w.r.t λ^*.

By taking derivative we know that the RHS is a decreasing function w.r.t to λ^*. Therefore we will have at most one solution of this case; again "at most" is due to the fact that $\frac{\int_{\lambda^*}^{\infty} x f_{X^i}(x) dx}{P(X_t^i \geq \lambda^*) + \bar{t}_i^r} \geq c_i$ does not hold necessarily. Next we will show that there cannot exist solution for both cases above.

Suppose there exists solution for the case $\lambda^* < c_i$, therefore we get

$$\frac{\int_{c_i}^{\infty} x f_{X^i}(x) dx + c_i \cdot P(X_t^i \leq c_i)}{1+\bar{t}_i^r} < c_i \tag{22}$$

which gives us

$$\bar{t}_i^r > \frac{\int_{c_i}^{\infty} x f_{X^i}(x) dx + c_i \cdot P(X_t^i \leq c_i)}{c_i} - 1 \tag{23}$$

Now let us assume there exists solution for $\lambda^* \geq c_i$ as well, i.e.,

$$\frac{\int_{\lambda}^{\infty} x f_{X^i}(x) dx}{P(X_t^i \geq \lambda) + \bar{t}_i^r} \geq c_i \tag{24}$$

Substitute $\bar{t}_i^r > \frac{\int_{c_i}^{\infty} x f_{X^i}(x)dx + c_i \cdot P(X_t^i \leq c_i)}{c_i} - 1$ into the L.H.S we have (details of derivations omitted)

$$\text{L.H.S} < \frac{\int_{\lambda}^{\infty} x f_{X_t}(x) dx c_i}{\int_{c_i}^{\infty} x f_{X_t}(x) dx + c_i P(X_t \leq c_i) - c_i P(X_t \leq \lambda)} \leq c_i \tag{25}$$

which contradicts the fact $\frac{\int_{\lambda}^{\infty} x f_{X^i}(x) dx}{P(X_t^i \geq \lambda) + \bar{t}_i^r} \geq c_i$.

On the other hand, if both of the cases have no solution, we can follow the similar arguments above and reach the similar conclusion. Therefore we proved the uniqueness of the threshold. ∎

Remark 4.3: As we can see, λ^* and c_i are independent with the current state (observation) for each stage; also as they are constants, one of the options for each stage can be easily eliminated immediately. Thus conditioning on continuing decision process, the strategy at each stage is either {**STAY**} or {**SWITCH**}. Meanwhile, when $\lambda^* > c_i$, staying on the current will provide higher future reward while when $\lambda^* < c_i$ the continuation decision would be to **SWITCH** to the next channel in order. (the $\lambda^* = c_i$ case is trivial as either move is good.) Thus the optimal strategy is all clear so far.

B. Monotonicity of value functions

Next we examine the effect of the decrease and increase of arrival process $\mathbf{G}$.

Lemma 4.4: $E\{V_i\} \geq E\{V_i'\}$ if $G_i \leq G_i', \forall i \in \Omega$.

Proof. Proof can be found in Appendix-A. ∎

The results verify the intuition when the arrival rate increases, a user will experience longer competition time; thus the cost of seeking diversity gains by giving up the current transmission right will increase. Therefore a user's expected throughput will decrease in general.

C. Ergodicity of arrival process $\mathbf{G}$

We start the analysis by first stating an assumption.

Assumption 3: No channel is dominant.

Consider a dominant case: all arrival rate will drift to one channel, for example, channel i. Denote the maximum throughput for packets staying at channel i by $\lambda_i(G_i)$. Denote the maximum averaged throughput for one packet to stay in the channel right before channel i as $\lambda_{i-1}(0)$. We assume

$$\lambda_i(G) < \lambda_{-i}(0), \forall i \in \Omega. \tag{26}$$

Next we investigate the ergodicity of each channel.

Lemma 4.5: Channels' arrival processes are ergodic.

Proof. Without losing generality, consider channel i. According to our assumption there exits a threshold $\tilde{G}_i$ such that

$$\lambda_i(G_i) < \lambda_{-i}(G_{-i}), \forall i \in \Omega \tag{27}$$

for all $G_i \geq \tilde{G}_i$. Here G_{-i} means the aggregated arrival rate of all other users except user i. As the throughput of sticking with each channel i is a decreasing function with G_i. Therefore We can see for $G_i > \tilde{G}_i$, we have

$$\lambda_i(G_i) < \lambda_{-i}(G_{-i}), \forall i \tag{28}$$

Under this case, the arrivals on channel $i-1$ will NOT skip to channel i, i.e., for $\hat{G}_i > G_i$ the probability of skipping satisfies

$$\Pr\{\hat{G}_i | G_i\} = 0 \tag{29}$$

Define any increasing, unbounded Lyapunov function $L(G_i)$ on $[0, G]$(for example, $L(G_i) = \frac{1}{G-G_i}$), we have

$$E_{\hat{G}_i}[L(\hat{G}_i)|G_i] \leq L(G_i) \tag{30}$$

By Foster-Lyapunov criteria [9] we establish the ergodicity of our system's arrival process. From this point on, we will use $\mathbf{G}$ to denote the expected arrival vector on each channel. ∎

D. Load balance

Based on above developed optimal sensing strategy, we are interested in how the overall multi-user system works when the load of network changes. Intuitively there should be some balance among all the channels and we will show our results regarding load balance in this section.

First we prove a lemma.

Lemma 4.6: . $\frac{\partial G_i}{\partial G} \geq 0, \forall i \in \Omega$.

Proof. Poof is sketched here and we will prove this property by induction.

When $N = 1$, i.e., the system degenerates to a single channel case, we know $G_1 = G$, the claim holds obviously. And we have the induction basis.

Assume when $N = n-1$, the claim holds. Now consider the case with $N = n$. When there are n channels. Suppose under G, the stationary arrival vector is given by

$$\mathbf{G} = [G_1, G_2, ..., G_n]$$

We increase G to $G^{'}$ and without losing generality suppose channel 1's stationary arrival rate decreases, i.e., $G_1^{'} < G_1$. Then $G_{-1}^{'} > G_{-1}$. By induction hypothesis and $G_{-1}^{'} > G_{-1}$ we know

$$G_i^{'} > G_i, \forall i \in \Omega \backslash 1 \tag{31}$$

As we know under G and stationary distribution $\mathbf{G}$, the in flow of channel 1 $G_1(\text{IN})$ and out flow of channel 1 $G_1(\text{OUT})$ should be equal to each other, i.e., $G_1(\text{IN}) = G_1(\text{OUT})$. Consider the case with $G^{'}$. Let $G_1^{'} = G_1$ and $G_i^{'} > G_i, i \neq 1$. Based on the results from above sections we know

$$E\{V_i^{'}\} < E\{V_i\}, \forall i \in \Omega \backslash 1 \tag{32}$$

Denote the transmission the user can experience with the whole group as a combined virtual single channel and denote it as Channel "-1". Therefore we know For any user belongs to the "-1" group, the expected throughput decreases. Denote the expected value function as $E\{V_{-1}\}$. As

$$E\{V_{-1}^{'}\} < E\{V_{-1}\}, E\{V_1^{'}\} = E\{V_1\} \tag{33}$$

Based on derivation from last section we know the arrival into channel 1 should be non-decreased; also the threshold of skipping channel 1 is non-increased and therefore the out flow of channel 1 is non-increasing. Therefore the channel 1's arrival rate will go up and the case with $G_1^{'} < G_1$ cannot be stable, completing the proof. ∎

Based on the load balancing property above we now revisit the monotonicity of value functions. Note that for each stage $\frac{\partial G_i}{\partial G} \geq 0$, and combined with Lemma 4.4 we have the following lemma:

Lemma 4.7: $E\{V_i\}, i \in \Omega$ are all non-increasing functions of G.

E. Impact of data transmission T

In this subsection we analyze the impact of time reservation time T and will answer a general question whether by reserving more transmission time will bring users more benefits considering all the efforts we put on channel sensing.

Lemma 4.8: $E\{V_i\}, i = 1, 2, ..., N$ are all non-decreasing functions of T.

Proof. Proof can be found in Appendix-B. ∎

The results reflect the fact that once a user find a good enough transmission condition, it would like to reserve a longer time with the current channel setting provided it is within the channel's coherence time.

F. Monotonicity of threshold based policy

Here we will characterize the monotonicity of the threshold strategy w.r.t. G.

Lemma 4.9: G for $\frac{\int_{c_i}^{\infty} x f_{X_t^i}(x)dx + c_i \cdot P(X_t^i \leq c_i)}{1+t_i^r} \leq c_i$ is a one threshold region.

Proof. Proof can be found in Appendix-C. ∎

G. Iterative algorithm for computing the stopping rule

In this section we describe the process of calculating the threshold for each stage. Notice at the last stage there is no more channel to skip to, therefore the dynamic program reduces to a maximum rate of return problem.

$$\begin{aligned} V_{N,t}(x) &= \max\{X_t^N(x), \frac{T}{t_N^r + T} E\{V_{N,t+1}(y)|x\}\} \\ &= \max\{X_t^N(x), \frac{T}{t_N^r + T} E\{V_N\}\} \end{aligned} \tag{34}$$

This is a standard problem of maximum rate of return and we omit the details of solving this problem.

Recall that we have

$$V_i(x) = \max\{\hat{X}_t^i(x), \frac{T}{t_i^r + T} E\{V_i\}\} \tag{35}$$

We can thus solve dynamic equations backward. After calculating stage $i+1$'s threshold and $E\{V_{i+1}\}$, we can proceed to the previous stage and calculate. Consider stage i now. Based on the arrival rate G_i, constant transmission time T, we can calculate $c_i = \frac{T}{T+t_{i+1}^c} E\{V_{i+1}\}$ and t_{i+1}^r. Next step we proceed to calculate

$$\frac{\int_{c_i}^{\infty} x f_{X_t^i}(x)dx + c_i \cdot P(X_t^i \leq c_i)}{1 + \bar{t}_i^r}$$

If it is less than c_i, we are done and claim this is the threshold. Otherwise, we proceed to a fixed-point equation

$$\lambda^* = \frac{\int_{\lambda^*}^{\infty} x f_{X_t^i}(x)dx}{P(X_t^i \geq \lambda^*) + \bar{t}_i^r}$$

We solve this iteratively as following.

- Starting with any value λ_0^*, i.e., λ_0^* is the initial guess. Set $n = 0$.
- Set $\lambda_{n+1}^* := \frac{\int_{\lambda_n^*}^{\infty} x f_{X_t^i}(x)dx}{P(X_t^i \geq \lambda_n^*) + \bar{t}_i^r}$.
- $n := n + 1$.
- Repeat until converge.

H. A case study

For a packet joining a new channel, very likely it has to wait an extra amount of time. This is the major cost for enabling spectral diversity, i.e., contention time. This extra waiting time consists of two parts.

$$t_i^c = E_i^{c1} + E_i^{c2} \tag{36}$$

Here E_i^{c1} indicates the waiting time spent on an "unforunate" arrival during another user's transmission; while E_i^{c2} stands for the time for competition contention. Next we show how to compute E_i^{c1} and E_i^{c2}.

Consider a random access system. Let ζ denotes the random back-off time; y denotes the transmission duration of a CTS packet; $f_Y(Y \leq y)$ be the probability that a packet arrives within the last y transmission duration of a current going transmission. Following the results from [8], we have

$$f_Y(y) = G_s^i e^{-G_s^i y} \tag{37}$$

here G_s^i is the success rate of users' competition which is given by

$$G_s^i = \frac{G_i e^{-2G_i}}{1 + (1+T)G_i e^{-2G_i}} \tag{38}$$

Denote W as the contention window. Therefore we have

$$W_i = \frac{e^{2G_i}}{G_i} - 1 \tag{39}$$

Then E_i^{c1} can be calculated as following.

$$\begin{aligned} E_i^{c1} &= \int_0^{1+T} f_Y(y)(1/\zeta + y)dy \\ &= \frac{1}{G_s^i} + \frac{1}{\zeta} - (T + 1 + \frac{1}{G_s^i} + \frac{1}{\zeta})e^{-(T+1)G_s^i} \end{aligned} \tag{40}$$

and

$$E_i^{c2} = (e^{2G_i} - 1) \cdot (1/\zeta + 2) \cdot (1 + T/W_i) + 2 \tag{41}$$

Meanwhile, we can see for a current user to release a channel, it will take him E_i^{c2} amount of time to get back the transmission right again, i.e.,

$$t_i^r = E_i^{c2} = (e^{2G_i} - 1) \cdot (1/\zeta + 2) \cdot (1 + T/W_i) + 2$$

Not hard to verify that(by taking derivatives) the assumptions in Assumption 1 and 2 are all met. The results in this section will be used in simulations.

Another point in calculating t^r and t^c is to help the users capture the arrival rate information. For some systems, the arrival rate information is not revealed to all users. Under this case, by collecting empirical data of t^r and t^c we can get an estimate of each, which the users can then use to estimate the arrival information at each channel.

V. Numerical Results

In this section we show a number of simulation results. Here we assume there are 5 independent channels with their transmission rates exponentially distributed within a finite range. Their average transmission rate are as following.

$$\{1/0.4, 1/0.6, 1/0.5, 1/0.3, 1/0.2\} \tag{42}$$

We set the back-off parameter as $1/\zeta = 10$ time units; transmission time $T = 50$ time units. The simulation results are as following.

From Fig.1 it is clear that opportunistic optimal strategy indeed can bring in benefits but only under mild arrival or competition rate from cognitive users. At the same time, according to Fig.1, with the increase of arrival rate, the average throughput per time unit decreases; meanwhile as the value functions simply will decide the stopping threshold at each channel, we can see with larger arrival rate, the stopping threshold gets smaller and smaller, i.e., the stopping region is enlarged. From Fig.2, the average throughput performance increases along with increase of reserved data transmission time T. But through simulation we also observe that this increase gets slower along with the increase of T; this is due to the increase of T also will increase cost or penalty for channel releasing and switching in our opportunistic channel access policy. Overall, these results are consistent with our analysis.

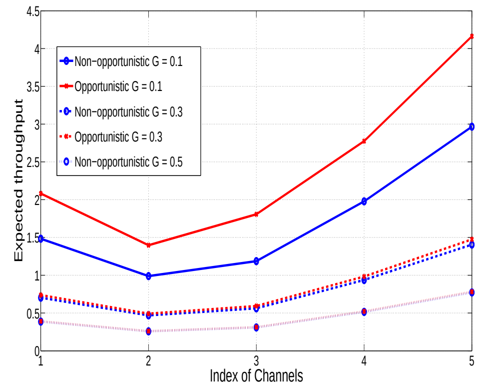

Fig. 1. Throughput under Opportunistic & non-opportunistic strategies under different G

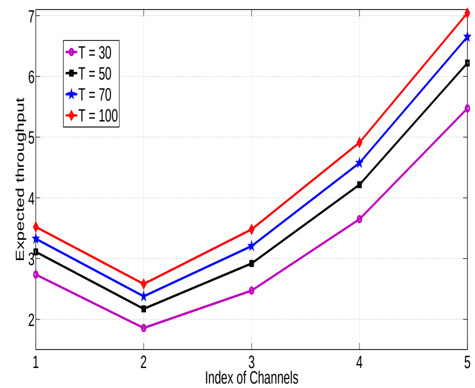

Fig. 2. Throughput under Opportunistic strategies with different T

Meanwhile these observations follow the intuition as when the arrival rate is high the risk of skipping the current transmission opportunity is dominating over the potential gain from the change of channel conditions. Therefore each cognitive user would rather transmit immediately compared to seeking for diversity gains. On the other hand, essentially if a user can reserve a longer transmission time, the potential benefits from opportunistic strategy increase.

Next we show the decision table conditioning on continuation on each channel (here in our experiment we consider a user starts channel sensing and decision process from channel 1). The intuition reflected from the table is that as the channels'

Arrival	Ch 1	Ch 2	Ch 3	Ch 4	Ch 5
0.05	STAY	SWITCH	SWITCH	SWITCH	STAY
0.1	STAY	SWITCH	SWITCH	STAY	STAY
0.3	STAY	SWITCH	SWITCH	STAY	STAY
0.5	STAY	SWITCH	SWITCH	STAY	STAY

TABLE I
DECISION OF I.I.D CHANNELS WITH DIFFERENT ARRIVAL RATE

conditions change fast, users on most channels, as long as either the current channel is a not bad one or there will not be much better channels later, will stay in the current one instead of joining an brand new channel to avoid future collisions especially when the collision probability (arrival rate) is high.

From the above table we can also see with the change of arrival rate, the continuation decision on certain channel may change. Take channel 4 for example. When the arrival rate is low, users on channel 4 is encouraged to switch as the penalty for switching to channel 5 is light and meanwhile channel 5 can provides better average throughput according to our experiment's parameter setting; however, when the arrival rate goes large, the penalty or risk of switching to channel 5 dominates the potential benefits by channel switching. Therefore users on channel 4 will chose to stay on the current channel instead of skip.

Another observation from the decision table is the clear advantages of our strategy compared to the isolated single channel opportunistic strategy by making use of only temporal diversity. Take channels 2 & 3 for example. As channels 2 & 3's expected throughput (with the lowest expected transmission rate according to our experiment setting) is limited even by taking temporal diversity on the single channel. Therefore, the continuation decision is shown to be SWITCH, i.e., the optimal strategy provides users on channel 3 more rewards(which is essentially brought in by spectral diversity) than temporal diversity by sticking with the single channel 3 which follows the strategy in [15].

VI. CONCLUSION

In this paper, OSA problems for multi-user multi-channel wireless network have been investigated and addressed. Optimal spectrum access policies have been developed and examined by taking use of multi-user, spectral and temporal diversities. The impacts of different factors such as arrival rate from cognitive users, transmission reservation times, etc have been investigated. At the same time, through theoretical and simulation results, benefits from opportunistic spectrum access under different network conditions and settings are shown clearly.

REFERENCES

[1] Sahand Haji Ali Ahmad, Mingyan Liu, Tara Javidi, Qing Zhao, and Bhaskar Krishnamachari. Optimality of myopic sensing in multichannel opportunistic access. *IEEE Trans. Inf. Theor.*, 55(9):4040–4050, September 2009.

[2] N.B. Chang and Mingyan Liu. Optimal competitive algorithms for opportunistic spectrum access. *Selected Areas in Communications, IEEE Journal on*, 26(7):1183 –1192, september 2008.

[3] Nicholas B. Chang and Mingyan Liu. Optimal channel probing and transmission scheduling for opportunistic spectrum access. *IEEE/ACM Trans. Netw.*, 17(6):1805–1818, 2009.

[4] Thomas S. Ferguson. Optimal stopping and applications. *Mathematics Department, UCLA*, 2006.

[5] V. Kanodia, A. Sabharwal, and E. Knightly. Moar: A multi-channel opportunistic auto-rate media access protocol for ad hoc networks. In *Proceedings of the First International Conference on Broadband Networks*, BROADNETS '04, pages 600–610, Washington, DC, USA, 2004. IEEE Computer Society.

[6] Ying C. Liang, Yonghong Zeng, E. C. Y. Peh, and Anh T. Hoang. Sensing-Throughput tradeoff for cognitive radio networks. *IEEE Transactions on Wireless Communications*, 7(4):1326–1337, 2008.

[7] Keqin Liu and Qing Zhao. Indexability of restless bandit problems and optimality of whittle index for dynamic multichannel access. *IEEE Trans. Inf. Theor.*, 56(11):5547–5567, November 2010.

[8] Y. Liu, M. Liu, and J. Deng. Is diversity gain worth the pain: a delay comparison between opportunistic multi-channel mac and single-channel mac. *Proc. of the 31st IEEE Conference of Computer Communications (INFOCOM '12), Minisymposium*, March 25-30 2012.

[9] Sean P. Meyn and R. L. Tweedie. Stability of Markovian Processes III: Foster-Lyapunov Criteria for Continuous-Time Processes. *Advances in Applied Probability*, 25(3), 1993.

[10] Xiangping Qin and Randall Berry. Exploiting multiuser diversity for medium access control in wireless networks. In *Proc. of IEEE INFOCOM*, pages 1084–1094, 2003.

[11] Tao Shu and Marwan Krunz. Throughput-efficient sequential channel sensing and probing in cognitive radio networks under sensing errors. In *Proceedings of the 15th annual international conference on Mobile computing and networking*, MobiCom '09, pages 37–48, New York, NY, USA, 2009. ACM.

[12] Sheu-Sheu Tan, Dong Zheng, Junshan Zhang, and James Zeidler. Distributed opportunistic scheduling for ad-hoc communications under delay constraints. In *Proceedings of the 29th conference on Information communications*, INFOCOM'10, pages 2874–2882, Piscataway, NJ, USA, 2010. IEEE Press.

[13] Qing Zhao, B. Krishnamachari, and Keqin Liu. On myopic sensing for multi-channel opportunistic access: structure, optimality, and performance. *Wireless Communications, IEEE Transactions on*, 7(12):5431 – 5440, december 2008.

[14] Qing Zhao, Lang Tong, Ananthram Swami, and Yunxia Chen. Decentralized cognitive mac for opportunistic spectrum access in ad hoc networks: A pomdp framework. *IEEE Journal on Selected Areas in Communications*, 25(3):589–600, 2007.

[15] Dong Zheng, Weiyan Ge, and Junshan Zhang. Distributed opportunistic scheduling for ad hoc networks with random access: an optimal stopping approach. *IEEE Trans. Inf. Theor.*, 55(1):205–222, January 2009.

APPENDIX A
PROOF OF MONOTONICITY OF VALUE FUNCTION W.R.T. $\mathbf{G}$

We prove this by induction.

When $i = N$, i.e., the last stage, we have

$$\lambda_N^* \bar{t}_N^r = \int_{\lambda_N^*}^{+\infty} (x - \lambda_N^*) f_{X^N}(x) dx \tag{43}$$

As t_N^r is a non-decreasing function w.r.t G_N, it is also increasing with $\mathbf{G}$. And with the increase of t_N^r, the solution λ_N^* can not increase. Thus we proved that λ_N^* is a non-increasing function w.r.t. $\mathbf{G}$.

Next let's assume the non-decreasing property holds for $i = n+1, n+1 \leq N$. Consider $i = n$. We prove both case with $\lambda_n^* < c_n$ and $\lambda_n^* \geq c_n$.

For the second case $\lambda_n^* \geq c_n$, we have

$$\lambda_n^* \bar{t}_N^r = \int_{\lambda_n^*}^{+\infty} (x - \lambda_n^*) f_{X_t^n}(x) dx \tag{44}$$

Similarly we know λ_n^* is a decreasing function w.r.t λ.

For the first case $\lambda_n^* < c_n$,

$$\lambda_n^* = \frac{\int_{c_n}^{\infty} x f_{X_t^n}(x)dx + c_n \cdot P(X_t^n \leq c_n)}{1 + \bar{t}_n^r} \tag{45}$$

We can easily get $E\{V_n\} = \int_{c_n}^{\infty} x f_{X_t^n}(x)dx + c_n \cdot P(X_t^n \leq c_n)$; take derivative of $E\{V_n\}$ with respect to $\mathbf{G}$ we get

$$\frac{\partial E\{V_n\}}{\partial \mathbf{G}} = \frac{\partial [E(X_t^n) - \int_0^{c_n} x f_{X_t^n}(x)dx + c_n P(X_t^n \leq c_n)]}{\partial c_n} \cdot \frac{\partial c_n}{\partial \mathbf{G}} \tag{46}$$

$$\frac{\partial c_n}{\partial \mathbf{G}} = \frac{\frac{\partial E\{V_{n+1}\}}{\partial \mathbf{G}}(T + t_{n+1}^c) - E\{V_{n+1}\}\frac{\partial t_{n+1}^c}{\partial G_{n+1}}}{(T + t_{n+1}^c)^2} \tag{47}$$

By induction hypothesis we know $\frac{\partial E\{V_{n+1}\}}{\partial \mathbf{G}} \leq 0$ and $\frac{\partial t_{n+1}^c}{\partial G_{n+1}} \geq 0$. Therefore we conclude that $\frac{\partial c_n}{\partial G_n} \leq 0, \frac{\partial E\{V_n\}}{\partial \mathbf{G}} \leq 0$. Induction step is thus completed. ∎

Appendix B
Monotonicity of value functions w.r.t. T

When $i = N$, i.e. the last stage, we have

$$\lambda_N^* \bar{t}_N^r = \int_{\lambda_N^*}^{+\infty} (x - \lambda_N^*) f_{X_t^N}(x)dx \tag{48}$$

As $\bar{t}_N^r$ is a non-increasing function w.r.t T. And with the decrease of t_N^r, the solution λ_N^* increases. Thus we proved that λ_N^* is a non-decreasing function w.r.t $\mathbf{G}$. Assume now the claim holds for $i = n + 1$. When $i = n$, consider two cases. For the second case $\lambda_n^* \geq c_n$, we have

$$\lambda_n^* \bar{t}_n^r = \int_{\lambda_n^*}^{+\infty} (x - \lambda_n^*) f_{X_t^n}(x)dx \tag{49}$$

Similar with the $i = N$ case we know λ_n^* is an increasing function w.r.t T.

For the first case $\lambda_n^* < c_n$,

$$\lambda_n^* = \frac{\int_{c_n}^{\infty} x f_{X_t^n}(x)dx + c_n \cdot P(X_t^n \leq c_n)}{1 + \bar{t}_n^r} \tag{50}$$

We can easily get

$$E\{V_n\} = \int_{c_n}^{\infty} x f_{X_t^n}(x)dx + c_n \cdot P(X_t^n \leq c_n) \tag{51}$$

Take derivative of $E\{V_n\}$ with respect to T

$$\frac{\partial E\{V_n\}}{\partial T} = \frac{\partial [E(X_t^n) - \int_0^{c_n} x f_{X_t^n}(x)dx + c_n P(X_t^n \leq c_n)]}{\partial c_n} \cdot \frac{\partial c_n}{\partial T} \tag{52}$$

With basic algebra (we will omit here) and combine with the fact $\frac{\partial E\{V_{n+1}\}}{\partial T} \geq 0$ (induction hypothesis) and $\frac{\partial t_{n+1}^c}{\partial T} \geq 0$, we conclude that $\frac{\partial c_n}{\partial T} > 0, \frac{\partial E\{V_n\}}{\partial T} > 0$. Induction step is thus completed. ∎

Appendix C
Monotonicity of threshold strategy w.r.t. G

Consider $\int_{c_i}^{\infty} x f_{X_t^i}(x)dx + c_i \cdot P(X_t^i \leq c_i) - c_i(1 + \bar{t}_i^r)$. Taking derivative gives us

$$\frac{\partial \int_{c_i}^{\infty} x f_{X_t^i}(x)dx + c_i \cdot P(X_t^i \leq c_i) - c_i(1 + \bar{t}_i^r)}{\partial G} = -\frac{\partial c_i}{\partial G} - \bar{t}_i^r \frac{\partial c_i}{\partial G} - c_i \frac{\partial \bar{t}_i^r}{\partial G} < 0 \ . \tag{53}$$

∎