# Finding Optimal Action Point for Multi-stage Spectrum Access in Cognitive Radio Networks

Bowen Li, Panlong Yang, *Member IEEE*
Jinlong Wang, Qihui Wu, *Member IEEE*
PLA University of Science and Technology, Nanjing, China
lbw1225@gmail.com, plyang@computer.org

Xiang-yang Li, *Senior Member*
Illinois Institute of Technology, USA
Department of Computer Science,
xli@iit.edu

***Abstract*—A critical challenge in Cognitive Radio Networks (CRN) is to make decision in real-time on accessing and releasing available channels, maximizing the spectrum utilization efficiency with primary users (PUs) protection. In this work, we make investigations on finding the optimal action point to explore and exploit the frequency-temporal diversity in addition to spectrum availability in multi-stage spectrum access. In tackling this difficulty, we model the Rayleigh fading channel probing process under PUs activity into a finite state Markov channel (FSMC) model with absorbing state, where a two-dimension (2D) optimal stopping problem is formulated for insightful analysis. Further, we prove that, the complexity of the 2D optimal stopping rule can be reduced to one threshold policy, where the optimal character still holds. After properly constructing multi-absorbing-states Markov chain, we can accurately achieve the expected throughput in our proposed scheme. Numerical and simulations results have also shown that, our threshold based access/switch strategy attains greater throughput than conventional idle/busy based access/switch strategy at the cost of minor additional delay in most cases.**

## I. INTRODUCTION

Conventional approaches for dynamic spectrum access are mainly focusing on finding idle channels and transmitting until a PU comes back. However, for a given node pair, the qualities of different channels are time dependent variables, which may lead to significant differences across all channels over time. Thus, it is possible for a CR to make multistage dynamic access decisions, which could further improve throughput in the potential opportunities during transmissions.

In this paper, we focus on the optimal multi-stage spectrum access strategy on providing a CR with the maximum possible throughput by exploring and exploiting frequency-temporal diversity in addition to spectrum availability. In our approach, a CR not only senses the busy/idle status of the channel, but also probes/monitors the instantaneous channel quality. Based on this additional information, a CR will make a real-time decision that when to access a channel and when to release it for another, which is considered as action point.

To tackle these difficulties, we adopt stochastic control framework. Our contributions are as follows: First, we derive the optimal action point for dynamic spectrum access, which is critical for a CR to decide whether to access/release a channel. Second, we consider a more general PUs activity model and realistic fading channel environment, which means our work is more applicable. Third, we develop an absorbing Markov analytical model for analyzing transmission process in FSMC channel under PUs protection, whose validity and accuracy is verified by simulation.

Compared to related works, this paper differs in one or more of the following aspects: 1). General unslotted PU idle time distribution. As depicted in [1], [2], [3], by assuming slotted PU activities, these papers focused on optimal channel access and selection. In comparison, we consider unslotted system and derive the optimal action point for channel access and release in a more general form. 2). Channel quality in addition to channel availability. In contrast to [1], [2], [3], and [4] that assume binary channel states, our project is jointly considering the channel quality and channel availability, thus achieving better spectrum utilization by choosing channel in good state. 3). Time-varying Rayleigh fading channel. The authors in [5] and [6] assume temporal independence of channel states in adjacent slots, which ignores the important time-dependent property of fading channel. Thus, they only make decision before accessing. While in our work, we model a more realistic Rayleigh fading environment by FSMC, and make decision during transmission as well as before accessing.

The remainder of the paper is organized as follows. In Section II, we present the system model. In Section III, we formally defined the problem, and derive the structure of the optimal strategy in Section IV, and obtain the optimal strategy in section V. We present the numeric results in Section VI, and conclude the paper in Section VII.

## II. SYSTEM MODEL

### A. System Description

We consider a cognitive wireless system consisting of $N$ non-overlapping licensed channels, and a single cognitive transmitter receiver pair that can use only one of the available channels at a time. Channels are indexed by the set $C = \{c_1, c_2, \dots, c_N\}$, each with a fixed bandwidth $B$Hz.

When a CR wants to transmit, it starts scanning the channels sequentially, as shown in Figure 1. Specifically, it randomly picks a channel, say $c_1$, and samples it for $\tau_s$ time. If it is idle, the CR further performs channel measurement using CPP/PFP exchanged mechanism (see in [6]). The channel probing takes $\tau_p$ time and obtains the exact channel quality information $\phi$. After each channel sensing and probing, the user needs to make decision immediately, on whether to access current

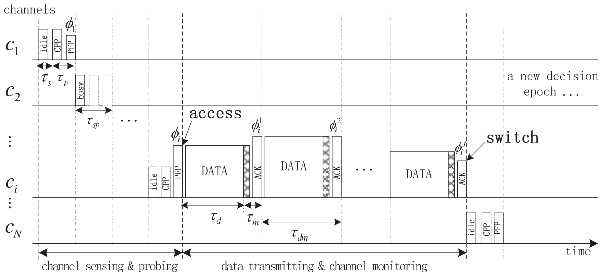


Fig. 1. Optimal action point for dynamic spectrum access and release

channel or skip to sense/probe another channel (for the sake of PUs protection, recalling a channel which is previously skipped is not allowed). Note that if the channel is busy during the sensing phase, no probing packet should be exchanged to avoiding interfering with PUs. However, the CR still has to wait for $\tau_p$ time to skip to another channel, thus to guarantee the tx-rx synchronization under the case that the channel availability of CR transmitter and receiver is different [6]. Therefore, whether or not the channel is idle, the time cost for one channel sensing/probing is $\tau_{sp} = \tau_s + \tau_p$.

When the user decides to access channel $c_i$ after observing channel quality state $\phi_i$, it transmits data packet using the maximum achievable data rate supported by current channel quality. By embedding training sequence into the tail of the DATA and inserting the channel quality information into ACK, the user obtains the latest channel state information (ACK brings channel quality information back and NACK means channel is occupied [4]). Such channel monitoring costs $\tau_m$ time for each packet duration $\tau_{dm}$. Similarly, at each ending instance of a packet transmission, the user has to make real-time decision whether to continue transmitting over current channel or switch to another channel for a new channel searching process.

We assume that the clock is synchronized among the transmitter and the receiver [6]. Hence, under the same decision rule, the transmitter and receiver will always sense, probe, access and switch to the same channel at the same time.

### B. Channel Availability Model

The availability status of a channel is modeled as a continuous-time random process that alternates between two states: idle and busy. A busy (idle) state indicates that some (no) PR user is transmitting over the channel. The sojourn times of idle and busy states are independent. They follow distribution function $F_i(\cdot)$ and $F_b(\cdot)$ with means $\alpha$ and $\beta$, respectively. The distribution functions can be estimated through extensive measurement during the deployment of secondary network, through various methods as in [7]. Here we focus on the homogeneous channel utilization scenario, which may correspond to the scenario that all channels belong to the same licensed network.

When a channel is observed at an arbitrary time, its idle and busy probabilities are given by $\theta_i = \frac{\alpha}{\alpha+\beta}$ and $\theta_b = \frac{\beta}{\alpha+\beta}$, respectively. Further, given that the channel is idle at time $t$, the probability that the it will remain idle or become busy during the transmission time $\tau_{dm}$ is respectively given by

$$\rho_{i,i} = \frac{1 - F_i(t + \tau_{dm})}{1 - F_i(t)}, \quad \rho_{i,b} = \frac{F_i(t + \tau_{dm})}{1 - F_i(t)}$$

### C. Wireless Channel Model and Multi-rate Link

A Rayleigh fading channel can be accurately modeled as a finite-state Markov channel (FSMC) in a slow fading environment [8] [9]. With an $K$ state FSMC: $\{0, 1, \ldots, K-1\}$, the signal-to-noise ratio (SNR) $\gamma$ at the receiver can be partitioned into multiple non-overlapping intervals by thresholds $\Gamma_k$ ($k \in \{0, 1, \ldots K\}$), where $0 = \Gamma_0 < \Gamma_1 < \Gamma_2 < \cdots < \Gamma_K = \infty$. The channel is said to be in state $k$ if $\Gamma_k \leq \gamma < \Gamma_{k+1}$ and thus can achieve data rate $R_k = B\ln(1 + \Gamma_k)$ nats/s without error by applying proper transmission mode. Without loss of generality, we set the rate increment between adjacent channel states to be $\eta$. Then, we have $\Gamma_k = e^{\frac{k\eta}{B}} - 1$, $k \in \{0, 1, \ldots K-1\}$.

As in a typical multipath propagation environment, the received SNR is distributed exponentially. Given the average SNR $\gamma_0$, the steady-state probability is thus given by [8]

$$\pi_k = \int_{\Gamma_k}^{\Gamma_{k+1}} p(\gamma)\, d\gamma = e^{-\frac{\Gamma_k}{\gamma_0}} - e^{-\frac{\Gamma_{k+1}}{\gamma_0}}, \quad k = 0, 1, \ldots, K-1$$

The state transition probability can be obtained as follows:

$$\begin{cases} \mu_{k,k+1} = \frac{\Lambda(\Gamma_{k+1})}{\pi_k}\tau_{dm}, & k = 0, 1, \ldots, K-2 \\ \mu_{k,k-1} = \frac{\Lambda(\Gamma_k)}{\pi_k}\tau_{dm}, & k = 1, 2, \ldots, K-1 \end{cases} \tag{1}$$

Where $\tau_d$ is the length of packet duration time and $\Lambda(\cdot)$ is the level crossing rate function given by [8]

$$\Lambda(\Gamma) = \sqrt{\frac{2\pi\Gamma}{\gamma_0}} f_d e^{-\frac{\Gamma}{\gamma_0}}$$

## III. Problem Formulation

In the sensing/probing phase, each channel sensing and probing step costs $\tau_{sp}$ and obtains the channel state $\phi$, where $\phi \in \{s_0, s_1, \ldots, s_K\}$. The channel state is $s_0$ means that the channel is busy or idle but with quality state 0, while $s_k$ means channel is idle and is in quality state $k$. We define $\phi_i$ as the outcome of the $i$th sensing and probing step. The distribution of $\phi_i$ can be calculated as follows.

$$\Pr\{\phi_i = s_k\} = p_k = \begin{cases} \theta_b + \theta_i\pi_0 & k = 0 \\ \theta_i\pi_k & 1 \leq k \leq K-1 \end{cases} \tag{2}$$

In the data transmission and channel monitoring phase, channel state is observed periodically with a cost $\tau_m$ for each $\tau_{dm}$. We define $\phi_i^j$ as the outcome of the $j$th channel monitoring step over channel $i$. Under the constraints of PUs protection, CR should terminate the current transmission immediately once observing channel state $s_0$. Thus, we using a FSMC with an absorbing state $s_0$ to characterize the channel temporal property. The conditional probability $\Pr\{\phi_i^{j+1} = s_l | \phi_i^j = s_k\}$ ($k > 0$) is then given by

$$q_{k,l} = \begin{cases} \rho_{i,b} + \rho_{i,i}\mu_{k,0} & l = 0 \\ \rho_{i,i}\mu_{k,l} & 1 \leq l \leq K-1 \end{cases} \tag{3}$$

Suppose that the user decides to access a channel with state $\phi_i$ after the $i$th sensing and probing step, and then terminates data transmitting (switches) when monitoring the current channel state turns to $\phi_i^j$ in the $j$th channel monitoring step, then the observed sequence is $\left\{\phi_1, \phi_2, \cdots, \phi_i; \phi_i^1, \phi_i^2, \cdots, \phi_i^j\right\}$, and the throughput reward is given by

$$\xi_{(i,j)} = \frac{B_{(i,j)}}{T_{(i,j)}} = \frac{\tau_d \sum_{t=0}^{j-1} R\left(\phi_i^t\right)}{i\tau_{sp} + j\tau_{dm}} \tag{4}$$

Where $\phi_i^0 = \phi_i$. $B_{(i,j)}$ and $T_{(i,j)}$ is respectively the number of transmitted data and the corresponding consumed time in one decision epoch.

Define $\psi = \{(I, J) : 0 < I < \infty, 0 < J < \infty\}$ as the set of all possible two-dimension stopping rules. As the decision process is repeated independently, following the results of [10], to maximize the true expected throughput $E\left[B_{(I,J)}/T_{(I,J)}\right]$ is equivalent to maximize the ratio $E\left[B_{(I,J)}\right]/E\left[T_{(I,J)}\right]$. Hence, our goal is to find an optimal joint rule $(I^*, J^*) \in \psi$ that maximizes the following rate of return objective function:

$$\max_{(I,J)\in\psi} \left\{E\left[B_{(I,J)}\right]/E\left[T_{(I,J)}\right]\right\} \tag{5}$$

## IV. Structural Properties of the Two-Dimension Optimal Stopping Rule

In this section, we drive the basic properties of the optimal rule. We first introduce the following lemma.

*Lemma 1:* if for some $\lambda$, $\max_{(I,J)\in\psi} E[B_{(I,J)} - \lambda T_{(I,J)}] = 0$, then $\max_{(I,J)\in\psi} \frac{E[B_{(I,J)}]}{E[T_{(I,J)}]} = \lambda$ . And, $(I^*, J^*)$ is optimal for maximizing $\max_{(I,J)\in\psi} \frac{E[B_{(I,J)}]}{E[T_{(I,J)}]}$ if $\max_{(I,J)\in\psi} E[B_{(I,J)} - \lambda T_{(I,J)}] = 0$ is attained at $(I^*, J^*) \in \psi$.

The result of Lemma 1 is evident. Strict proof can be found in [10]. According to this lemma, we consider a transformed version of the problem, whose reward is defined by

$$w_{(i,j,\lambda)} = B_{i,j)} - \lambda T_{(i,j)} = \tau_d \sum_{t=0}^{j-1} R(\phi_i^t) - i\lambda\tau_{sp} - j\lambda\tau_{dm} \tag{6}$$

The corresponding objective function is given by

$$V(\lambda) = \max_{(I,J)\in\psi} E\left[w_{(I,J,\lambda)}\right] \tag{7}$$

Here, $\lambda$ can be seemed as the supposed throughput. For a given $\lambda$, there exists a expected return $V(\lambda)$ under corresponding optimal rule $(I_\lambda^*, J_\lambda^*)$ for the transformed problem. Our goal is to find $\lambda^*$ that makes $V(\lambda^*) = 0$. Then the corresponding optimal rule $(I_{\lambda^*}^*, J_{\lambda^*}^*)$ is the exact solution to the original problem (5).

### A. Property of the Temporal Optimal Stopping Rule

For a fixed supposed throughput $\lambda$ and frequency-domain stopping rule $I$, the reward of stopping transmission after monitoring $\phi_I^j$ can be derived from (6).

$$\begin{aligned} w_{(I,j,\lambda)} &= \tau_d \sum_{t=0}^{j-1} R\left(\phi_I^t\right) - j\lambda\tau_{dm} - I\lambda\tau_{sp} \\ &= \sum_{t=0}^{j-1} \left[R\left(\phi_I^t\right)\tau_d - \lambda\tau_{dm}\right] - I\lambda\tau_{sp} \end{aligned} \tag{8}$$

The temporal optimal stopping rule under a given $\lambda$ and rule $I$ is then given by

$$J_{(\lambda,I)}^* = \arg\max_J E\left[\sum_{t=0}^{j-1} \left[R\left(\phi_I^t\right)\tau_d - \lambda\tau_{dm}\right]\right] \tag{9}$$

Before exploring the optimal rule, we introduce following two lemmas, which is respectively derived from principle of optimality and property of FSMC.

*Lemma 2:* Let $S_{t_1}^{t_2} = \sum_{t=t_1}^{t_2} \left[R\left(\phi_I^t\right)\tau_d - \lambda\tau_{dm}\right]$. For a given $(\lambda, I)$ and current channel state $\phi_I^j$: 1). If there exists $k \geq 0$ that making $E[S_j^{j+k}|\phi_I^j] > 0$, then the optimal strategy is to continue transmitting; 2). If $E[S_j^{j+k}|\phi_I^j] \leq 0$ come into existence for all $k \geq 0$, then the optimal strategy is to switch.

*Lemma 3:* Define $f(x, s) = E[R(s^{(x)})|s^{(0)} = s]$, $x = 0, 1, 2\ldots$ and $s = 1, 2, \ldots, K-1$. Then $f(x, s)$ is a decreasing function of $x$ when $s \in \{s : R(s) \geq \bar{R}\}$ and increasing function when $s \in \{s : R(s) \geq \bar{R}\}$. Where $\bar{R} = \sum_{k=1}^{K-1} \pi_k R_k$.

Based on these two lemmas, we get the following theorem.

*Theorem 1:* For a given $(\lambda, I)$, the optimal strategy $J_\lambda^*$ has a threshold structure:

1). For $\lambda \geq \frac{\rho_{i,i}\tau_d}{\tau_{dm}}\bar{R}$, denote $\phi_\lambda^s = \max\{\phi : R(\phi) \leq \frac{\lambda\tau_{dm}}{\rho_{i,i}\tau_d}\}$, then $J_\lambda^*$ is to switch if $\phi_I^j \leq \phi_\lambda^s$ and to transmit otherwise.

2). For $\lambda < \frac{\rho_{i,i}\tau_d}{\tau_{dm}}\bar{R}$, there exists $\phi_\lambda^s \in \{\phi : 0 \leq R(\phi) \leq \frac{\lambda\tau_{dm}}{\rho_{i,i}\tau_d}\}$ that satisfying: $J_\lambda^*$ is to switch if $\phi_I^j \leq \phi_\lambda^s$ and to continue transmitting otherwise.

*Proof:* Let $y_t = \rho_{i,i}\tau_d E[R(s^{(t)})|s^{(0)} = \phi_I^j] - \lambda\tau_{dm}$,from Lemma3, we have:

for all $t \geq 0$, $\begin{cases} y_t \leq y_0, & if \quad R(\phi_I^j) \geq \bar{R} \\ y_t \leq y_\infty, & if \quad R(\phi_I^j) < \bar{R} \end{cases}$

As CR is only allowed to transmit when $\phi_I^j \neq s_0$ (for the sake of PUs protection), we have $E[S_j^{j+k}|\phi_I^j] = \sum_{t=0}^{k} y_t\rho_{i,i}^t$.

For all $\phi_I^j \in \{\phi : R(\phi) > \frac{\lambda\tau_{dm}}{\rho_{i,i}\tau_d}\}$, $E[S_j^j|\phi_I^j] = \rho_{i,i}R(\phi_I^j)\tau_d - \lambda\tau_{dm} > 0$. According to part 1 of Lemma 2, it is optimal to continue transmitting. On the other hand, as $y_t \leq \max\{y_0, y_\infty\}$ a.s. for all $t \geq 0$, for $\phi_I^j \in \{\phi : R(\phi) \leq \frac{\lambda\tau_{dm}}{\rho_{i,i}\tau_d}\}$, if $\lambda \geq \frac{\rho_{i,i}\tau_d}{\tau_{dm}}\bar{R}$, then $y_0 \leq 0$, $y_\infty \leq 0$. Thus for all $t \geq 0$, $y_t \leq 0$. Consequently, $E[S_j^{j+k}|\phi_I^j] = \sum_{t=0}^{k} y_t\rho_{i,i}^t \leq 0$ a.s. for all $k \geq 0$. Applying Lemma 2, we get theorem 1-1.

For $\phi_I^j \in \{\phi : R(\phi) \leq \frac{\lambda\tau_{dm}}{\rho_{i,i}\tau_d}\}$, if $\lambda < \frac{\rho_{i,i}\tau_d}{\tau_{dm}}\bar{R}$, then $R(\phi_I^j) < \bar{R}$, thus we have: $y_0 < y_1 < \cdots < y_\infty$ and $y_0 < 0$, $y_\infty > 0$. Thus, if there exists some $k$ that makes $E[S_j^{j+k}|\phi_I^j] > 0$, it must be in $k = \infty$. Define $g(\phi_I^j) = E[S_j^\infty|\phi_I^j]$. It is a strictly increasing function, and $g(0) \leq 0$, $g(\frac{\lambda\tau_{dm}}{\rho_{i,i}\tau_d}) > 0$. Hence, there exists a state

$\phi_\lambda^s \in \{\phi : 0 \leq R(\phi) \leq \frac{\lambda\tau_{dm}}{\rho_{i,i}\tau_d}\}$ satisfying that: for all $\phi_I^j \leq \phi_\lambda^s$, $E[S_j^{j+k}|\phi_I^j] \leq 0$ a.s. for all $k \geq 0$; and for all $\phi_I^j > \phi_\lambda^s$, $E[S_j^\infty|\phi_I^j] > 0$. Thus, from Lemma 2, we get theorem 1-2. ∎

*B. Structural Property of the Two-Dimension Optimal Stopping Rule*

For a given $\lambda$, the optimal strategy $J_\lambda^*$ is derived from Theorem 1. Then, under this temporal optimal rule, the expected reward of accessing $c_i$ after observing channel state $\phi_i$ is

$$w_{(i,J_\lambda^*,\lambda)} = E[b(\phi_i,\phi_\lambda^s)] - \lambda E[t_{dm}(\phi_i,\phi_\lambda^s)] - i\lambda\tau_{sp} \quad (10)$$

Where $E[b(\phi_i,\phi_\lambda^s)]$ and $E[t_{dm}(\phi_i,\phi_\lambda^s)]$ respectively denote the expected number of transmitted bits and consumed time in each transmission phase under the strategy that accessing channel $c_i$ with state $\phi_i$ and switching when state become worse than or equal to $\phi_\lambda^s$.

*Theorem 2:* For a given $\lambda$, the optimal rule $I_\lambda^*$ has a threshold structure. Let $\phi_\lambda^a = \min\{\phi : E[t(\phi,\phi_\lambda^s)] - \lambda E[t_{dm}(\phi,\phi_\lambda^s)] \geq V(\lambda)\}$ ($\phi_\lambda^a > \phi_\lambda^s$). Then $I_\lambda^*$ is to access if $\phi_i \geq \phi_\lambda^a$ and to sense/probe another channel otherwise.

*Proof:* let $w_{(i,J_\lambda^*,\lambda)} = Y_i(\lambda)$, $E[b(\phi_i,\phi_\lambda^s)] - \lambda E[t_{dm}(\phi_i,\phi_\lambda^s)] = X_i(\lambda)$, $\lambda\tau_{sp} = \delta(\lambda)$, then the reward can be rewritten as $Y_i(\lambda) = X_i(\lambda) - i\delta(\lambda)$. For a given $\lambda$, applying the result in [10](chapter 4.2), the optimal rule is

$$I_\lambda^* = \min\{i \geq 1 : E[b(\phi_i,\phi_\lambda^s)] - \lambda E[t_{dm}(\phi_i,\phi_\lambda^s)] \geq V(\lambda)\}$$

Let $h_\lambda(\phi_i) = E[b(\phi_i,\phi_\lambda^s)] - \lambda E[t_{dm}(\phi_i,\phi_\lambda^s)]$. Under the optimal rule $J_\lambda^*$, $h_\lambda(\phi_i)$ is strictly increasing with $\phi_i$. Thus, the optimal accessing rule can be rewritten as $I_\lambda^* = \min\{i \geq 1 : \phi_i \geq \phi_\lambda^a\}$. ∎

Through Theorem 1 and Theorem 2, we conclude that for arbitrary given $\lambda$, the joint optimal strategy $(I_\lambda^*, J_\lambda^*)$ must be threshold-based. Thus for the maximum achieved throughput $\lambda^*$, the optimal strategy $((I_{\lambda^*}^*, J_{\lambda*}^*))$ is also threshold-based. Recalling that $V(\lambda^*) = 0$, we have the following theorem:

*Theorem 3:* The joint optimal strategy $(I_{\lambda^*}^*, J_{\lambda^*}^*)$ can be simplified to a single threshold form $A_{s_{k^*}}$:

1). Senses and probes channels sequentially in frequency-domain until finding an idle channel with quality better than $s_{k^*}$, then access it for data transmission;

2). Transmits data and monitors current channel in temporal-domain until observing channel state is worse than or equal to $s_{k^*}$ (NACK is included since channel state is treated as $s_0$ in that case), then switch to another channel for a new round of channel searching.

*Proof:* Let $E[b(\phi_1,\phi_2)] - \lambda^* E[t_{dm}(\phi_1,\phi_2)] = h(\phi_1,\phi_2)$. From theorem 1, for $\lambda^* < \frac{\rho_{i,i}\tau_d}{\tau_{dm}}\bar{R}$, if $\phi > \phi_{\lambda^*}^s$, then $h(\phi,\phi_{\lambda^*}^s) \geq h(\phi,s_0) = E[S^\infty|\phi] > 0$; for $\lambda^* \geq \frac{\rho_{i,i}\tau_d}{\tau_{dm}}\bar{R}$, $\phi_{\lambda^*}^s = \max\{\phi : R(\phi) \leq \frac{\lambda^*\tau_{dm}}{\rho_{i,i}\tau_d}\}$, which means user only transmitting on a channel with $R > \frac{\lambda^*\tau_{dm}}{\rho_{i,i}\tau_d}$. Thus, $h(\phi,\phi_{\lambda^*}^s) > 0$ a.s. for all $\phi > \phi_{\lambda^*}^s$. On the other side, as $V(\lambda^*) = 0$, from theorem 2, $\phi_{\lambda^*}^a = \min\{\phi : h(\phi,\phi_{\lambda^*}^s) \geq 0\}$ ($\phi_{\lambda^*}^a > \phi_{\lambda^*}^s$). Thus, for all $\phi > \phi_{\lambda^*}^s$, we have $\phi \geq \phi_{\lambda^*}^a$. By combining theorem 1 with theorem 2 and setting $s_{k^*} = \phi_{\lambda^*}^s$, we get theorem 3. ∎

## V. Optimal Strategy and Maximum Throughput

We use ergodic approach to obtain the optimal strategy and corresponding achievable throughput. Specifically, we calculate the expected throughput $\lambda_{s_{k^*}} = \frac{E[B(A_{s_{k^*}})]}{E[T(A_{s_{k^*}})]}$ of each candidate strategy $A_{s_{k^*}}$ ($s_{k^*} \in \{s_0, s_1, \ldots, s_{K-2}\}$), and then choose the best strategy which brings maximum expected throughput.

Under the strategy $A_{s_{k^*}}$, the transmission process can be modeled as an multi-absorbing-states Markov chain. By setting the states $s_k \leq s_{k^*}$ as absorbing states, we get the transition matrix of the remainder transit states as follows.

$$Q_{k^*} = \begin{bmatrix} q_{k^*+1,k^*+1} & q_{k^*+1,k^*+2} & \cdots & 0 \\ q_{k^*+2,k^*+1} & q_{k^*+2,k^*+2} & q_{k^*+2,k^*+3} & \vdots \\ 0 & \ddots & \ddots & 0 \\ \vdots & q_{K-2,K-3} & q_{K-2,K-2} & q_{K-2,K-1} \\ 0 & \cdots & q_{K-1,K-2} & q_{K-1,K-1} \end{bmatrix}$$

Where $q_{k,l}$ is defined in 3. Then, we have follow lemma.

*Lemma 4:* For an absorbing Markov chain, the fundamental matrix $\mathbf{O}_{k^*} = \mathbf{I} + \sum_{n=1}^{\infty} \mathbf{Q}_{k^*}^{(n)} = (\mathbf{I} - \mathbf{Q}_{k^*})^{-1}$. The entry $o_{ij}$ of $\mathbf{O}_{k^*}$ gives the expected number of times that the process is in the transient state $s_j$, if it is started in the transient state $s_i$.

The proof of Lemma 4 can be found in [11] (theorem 11.4 of chapter 11.2). Following Lemma 4, we have

$$E[b(s_k,s_{k^*})] = \tau_d e_{k-k^*}^T (\mathbf{I} - \mathbf{Q}_{k^*})^{-1} r_{k^*}$$
$$E[t_{dm}(s_k,s_{k^*})] = \tau_{dm} e_{k-k^*}^T (\mathbf{I} - \mathbf{Q}_{k^*})^{-1} z_{k^*}$$

Where $e_{k-k^*}$ is the $k-k^*$th column of the identity matrix $\mathbf{I}_{K-1-k^*}$, $r_{k^*} = [R_{k^*+1}, R_{k^*+2}, \ldots, R_{K-1}]$ is the data rate vector. $z_{k^*}$ is a $(K-1-k^*) \times 1$ vector all of whose entries are 1.

Then, the expected number of transmitted bits and the expected cost time in transmission phase under the strategy $A_{s_{k^*}}$ are respectively given by

$$\begin{aligned} E[B(A_{s_{k^*}})] &= E\left[E[b(s_k,s_{k^*})]|k > k^*\right] \\ &= \tfrac{\tau_d}{1-P_{k^*}} \left[\sum_{k>k^*}^{K-1} p_k e_{k-k^*}^T\right] (\mathbf{I} - \mathbf{Q}_{k^*})^{-1} r_{k^*} \end{aligned} \quad (11)$$

$$\begin{aligned} E[T_{dm}(A_{s_{k^*}})] &= E\left[E[t_{dm}(s_k,s_{k^*})]|k > k^*\right] \\ &= \tfrac{\tau_{dm}}{1-P_{k^*}} \left[\sum_{k>k^*}^{K-1} p_k e_{k-k^*}^T\right] (\mathbf{I} - \mathbf{Q}_{k^*})^{-1} z_{k^*} \end{aligned} \quad (12)$$

Where $P_{k^*} = Pr(k \leq k^*) = \sum_{k=0}^{k^*} p_k$.

For sensing/probing phase, the expected cost is given by

$$E[T_{sp}(A_{s_{k^*}})] = \tau_{sp} \sum_{i=1}^{\infty} [i P_{k^*}^{i-1}(1 - P_{k^*})] = \frac{\tau_{sp}}{1 - P_{k^*}} \quad (13)$$

Thus, from (11)(12)(13), the expected throughput is

$$\lambda_{s_{k^*}} = \frac{\tau_d \left[\sum_{k>k^*}^{K-1} p_k e_{k-k^*}^T\right] (\mathbf{I} - \mathbf{Q}_{k^*})^{-1} r_{k^*}}{\tau_{dm} \left[\sum_{k>k^*}^{K-1} p_k e_{k-k^*}^T\right] (\mathbf{I} - \mathbf{Q}_{k^*})^{-1} z_{k^*} + \tau_{sp}}$$

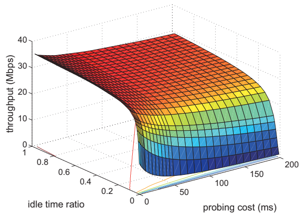


Fig. 2. theoretic throughput

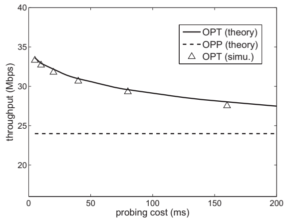


Fig. 3. throughput vs. probing cost

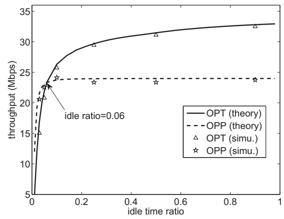


Fig. 4. throughput vs. idle time ratio

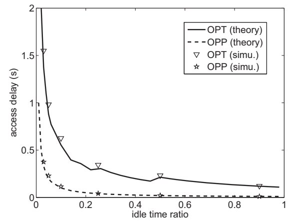


Fig. 5. access delay vs. idle time ratio

## VI. NUMERICAL RESULTS

We consider a typical scenario here that the sojourn time of idle and busy are exponential distributed and channel quality are slowly varied. The parameters for the simulation are as following: carrier frequency $f_c = 100MHz$, channel bandwidth $B = 6MHz$, average received SNR $\gamma_0 = 15dB$, terminal mobile speed $v = 1m/s$, average idle plus busy time $\alpha + \beta = 100s$, channel sensing time $\tau_s = 10ms$, training sequence time $\tau_m = 1ms$ and packet length $\tau_{dm} = 50ms$. By setting the rate interval $\eta = 3Mbps$, we build a 16 states Markov channel model to characterize the time-varying behavior of such wireless fading environment.

We first derive the theoretic achievable throughput of our proposed scheme in Fig.2. We observe that the throughput is slowly declined with probing cost increasing and/or idle time ratio decreasing in most of the cases except that the idle time ratio is lower than 0.1.

For performance comparison, we consider two strategies: OPT and OPP. OPT stands for our optimal threshold-based accessing/switching strategy. OPP is the conventional access strategy that CR accessing the first idle channel and transmitting until a PU comes back. The throughput performance under these two strategies is shown in Fig.3 and Fig.4.

In Fig.3, we consider a moderate spectrum opportunity environment ($\alpha = \beta = 50s$). Both theoretic and simulated results shows that OPT always perform better than OPP, even when probing cost is up to 200ms. This result indicates that to further explore channel quality is highly beneficial for system throughput in usual dynamic spectrum environment.

In Fig.4, we fixed probing cost $\tau_p = 30ms$ to study the throughput performance in different environments. By adjusting the idle time ratio from 0.01 (scarcity of spectrum opportunities) to 0.99 (abundance of spectrum opportunities), we get the theoretic and simulated throughput curves of OPT and OPP respectively. It is clear that OPT is much better than OPP in most cases. However, when facing severe shortage of spectrum opportunities ($\frac{\alpha}{\alpha+\beta} < 0.06$), channel quality exploration seems no longer a good choice. The result is reasonable for that when spectrum opportunity is extremely exiguous: to find a good idle channel leading high cost of searching time and waste of spectrum opportunities, thus brings on poor throughput.

We further investigate the access delay of OPT and OPP in different environments. As shown in Fig.5, our high throughput gain is achieved with delay expense, thus our scheme is more likely applied to delay tolerant applications.

## VII. CONCLUSION

In this work, we use two-dimension optimal stopping to achieve optimal dynamic spectrum access problem in CRN. In addition to spectrum availability, the frequency-temporal diversity of fading channels are effectively explored and exploited under our optimal rule, which is simple threshold based. Further, we get the accurate throughput by formulating the transmission process under the optimal rule to be an multi-absorbing-states Markov chain. Simulation works and numerical results shows that, in typical cognitive wireless networks, the proposed scheme significantly improve network throughput.

## VIII. ACKNOWLEDGMENT

This work was supported by NSF CNS-0832120, CNS-1035894, 973 Program of China (grant No.2009CB320400), 863 Program of China (grant No.2009AA01Z243), NSFC (grant No.60932002, No.61003277), program for Zhejiang Provincial Key Innovative Research Team, program for Zhejiang Provincial Overseas High-Level Talents and Jiangsu Natural Science Foundation (grant No. BK2010102).

## REFERENCES

[1] Y. Gai, B. Krishnamachari, and R. Jain, "Learning multiplayer channel allocations in cognitive radio networks: A combinatorial multi-armed bandit formulation," in *IEEE DySPAN*, 2010.

[2] A. Anandkumar, N. Michael, A. K. Tang, and A. Swami, "Distributed algorithms for learning and cognitive medium access with logarithmic regret," *IEEE Journal on Selected Areas in Communications*, 2010.

[3] K. Liu and Q. Zhao, "Decentralized multi-armed bandit with multiple distributed players," *CoRR*, vol. abs/0910.2065, 2009.

[4] S. Huang, X. Liu, and Z. Ding, "Optimal sensing-transmission structure for dynamic spectrum access," in *IEEE INFOCOM*, 2009.

[5] N. B. Chang and M. Liu, "Optimal channel probing and transmission scheduling for opportunistic spectrum access," *IEEE/ACM TRANSACTIONS ON NETWORKING*, vol. 17, pp. 1805–1818, 2009.

[6] T. Shu and M. Krunz, "Throughput-efficient sequential channel sensing and probing in cognitive radio networks under sensing errors," in *ACM MOBICOM*, 2009, pp. 37–48.

[7] D. Chen, S. Yin, Q. Zhang, M. Liu, and S. Li, "Mining spectrum usage data: a large-scale spectrum measurement study," in *ACM MOBICOM*, 2009, pp. 13–24.

[8] Q. Zhang and S. A. Kassam, "Finite-state markov model for rayleigh fading channels," *IEEE Transactions on Communications*, vol. 47, pp. 1688–1692, 1999.

[9] J. Razavilar, K. J. R. Liu, and S. I. Marcus, "Jointly optimized bit-rate/delay control policy for wireless packet networks with fading channels," *IEEE Transactions on Communications*, vol. 50, pp. 484–494, 2002.

[10] T. Ferguson, "Optimal stopping and applications."

[11] C. M. Grinstead and J. L. Snell, "Introduction to probability."