Bursty Traffic in Energy-Constrained Opportunistic Spectrum Access

Yunxia Chen, Qing Zhao
Department of Electrical and Computer Engineering
University of California, Davis, CA 94536

Ananthram Swami
Army Research Laboratory
Adelphi, MD 20783

***Abstract*—We design opportunistic spectrum access strategies for improving spectrum efficiency. In each slot, a secondary user chooses a subset of channels to sense and decides whether to access based on the sensing outcomes. Incorporating the secondary user's residual energy and buffer state, we formulate this sequential decision-making problem as a partially observable Markov decision process (POMDP). Within the POMDP framework, we obtain stationary optimal sensing and access policies. By exploiting the rich structure of the underlying problem, we develop monotonicity results for the optimal policies, which accelerate the computations. Numerical results are provided to study the impact of the secondary user's packet arrival rate and residual energy on the optimal sensing and access decisions.**

I. Introduction

Opportunistic spectrum access (OSA) is one of the approaches envisioned for dynamic spectrum management [1]. It has received increasing attention due to its compatibility with the current spectrum management policy and legacy wireless systems. The basic idea of OSA is to allow secondary users to search for and exploit local and instantaneous spectrum availability in a non-intrusive manner. Correspondingly, basic design components of OSA include 1) a sensing strategy that specifies whether to sense and where in the spectrum to sense and 2) an access strategy that determines whether to access based on the sensing outcomes.

Related Work The design and implementation of OSA have been addressed in the literature [2]–[7]. In [2], the authors address the implementation of OSA in an ad hoc secondary network overlaying a GSM cellular network. In [3], optimal distributed MAC protocols are proposed within the framework of partially observable Markov decision process (POMDP). The proposed protocols ensure synchronous hopping of the secondary transmitter and receiver in the spectrum without introducing extra control message exchange. More recently, [4] exploits the channel fading in the design of OSA for an efficient use of secondary users' energy. In [5], a separation principle is established for the optimal joint design of the physical layer spectrum sensor and the MAC layer sensing and access policies. In [6], access strategies for a slotted secondary user searching for opportunities in an unslotted primary network is considered, where a round-robin single-channel sensing scheme is used. Modeling of spectrum occupancy has been addressed in [7]. Measurements obtained from spectrum monitoring test-beds demonstrate the Makovian transition between busy and idle channel states in wireless LAN. For an overview on recent developments in OSA and a survey of other dynamic spectrum access approaches, readers are referred to [8].

Contributions This paper extends [4] by incorporating both the bursty traffic and the energy constraint of secondary users into OSA design. We consider a secondary network whose users independently and selfishly seek spectrum opportunities in a slotted primary network. We formulate the sequential sensing and access decision-making of a secondary user as a POMDP problem, which takes into account the channel fading, the residual energy as well as the buffer state of the secondary user. We show that this POMDP terminates in a finite but random time. The optimal sensing and access strategies are thus given by the stationary optimal policies of this POMDP.

By exploiting the rich structure of the underlying problem, we then develop monotonicity results for the optimal policies. In particular, we show that for the one-channel case, the optimal sensing policy is a threshold policy: the secondary user with packets to transmit should sense a channel if and only if (iff) the conditional probability that this channel is available is above a certain threshold. Moreover, the optimal access policy is also a threshold policy: the secondary user should transmit over an idle channel iff the channel fading level is below a certain threshold. These monotonicity results can help us accelerate the calculation of the optimal sensing and access policies.

Finally, we provide numerical results to study different factors that affect the optimal sensing and access decisions. We find that the impact of the secondary user's residual energy and buffer state on the optimal decisions diminishes as the residual energy increases. We also see the benefit of sensing a channel even if the secondary user does not have any packets to send.

II. System Model

A. Primary Network Model

We consider a spectrum consisting of N channels, each with bandwidth W_n $(n = 1, \cdots, N)$. These N channels are licensed to a slotted primary network. Let $S_n(t) \in \{0 \text{ (busy)}, 1 \text{ (idle)}\}$ denote the occupancy of channel n by the primary network in slot t. We assume that the spectrum

[0]This work was supported in part by the Army Research Laboratory Collaborative Technology Alliance on Communication and Networks under Grant DAAD19-01-2-0011 and by the National Science Foundation under Grants CNS-0627090 and ECS-0622200.

occupancy state (SOS) $\mathbf{S}(t) \stackrel{\Delta}{=} [S_1(t), \ldots, S_N(t)]$ follows a time-homogeneous discrete Markov process with state space $\mathcal{S}$ defined as

$$\mathcal{S} \stackrel{\Delta}{=} \{0,1\}^N, \quad \text{where } |\mathcal{S}| = 2^N. \tag{1}$$

The transition probabilities are denoted as

$$P_{\mathbf{S}}(\mathbf{s}'|\mathbf{s}) \stackrel{\Delta}{=} \Pr\{\mathbf{S}(t) = \mathbf{s}' \,|\, \mathbf{S}(t-1) = \mathbf{s}\}, \quad \mathbf{s}, \mathbf{s}' \in \mathcal{S}, \tag{2}$$

which represents the probability that the SOS transits from $\mathbf{s}$ to $\mathbf{s}'$ at the beginning of slot t. The transition probabilities of the SOS are determined by the statistics of the primary traffic. We assume that they are known or have been learned.

B. Secondary Network Model

Consider an overlay ad hoc secondary network whose users seek instantaneous spectrum opportunities in these N channels. At the beginning of each slot, a secondary user chooses M ($1 \le M \le N$) channels to sense and determines whether to access based on the sensing outcomes. The secondary user can also turn to the sleeping mode in which no channel will be sensed or accessed in this slot. The sequence of operations performed by the secondary user in each slot is illustrated in Fig. 1 and will be detailed in Section III-B.

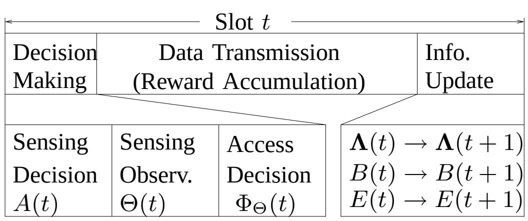

Fig. 1. The slot structure. The secondary user's knowledge of the SOS is characterized by $\mathbf{\Lambda}(t)$, and its buffer state and residual energy are denoted by $B(t)$ and $E(t)$, respectively.

Our goal is to design the optimal OSA strategy for the secondary user, which sequentially specifies which channels to sense and whether to access. The design objective is to maximize the throughput of the secondary user during its battery lifetime. For ease of presentation, we assume $M = 1$ (*e.g.*, the case of single carrier communications). Our formulation can be generalized to $M > 1$.

Traffic Model We model the bursty traffic of the secondary user as a Poisson process[1] with rate λ. That is, the probability of m packet arrivals in a slot is given by

$$q_m \stackrel{\Delta}{=} \frac{e^{-\lambda}\lambda^m}{m!}, \quad m = 0, 1, \ldots. \tag{3}$$

The transmission time of a packet is assumed equal to the slot length. We assume that the secondary user has a finite buffer with maximum size l. Packets are dropped when the buffer overflows. Let $B(t) \in \mathcal{B}$ denote the number of packets in the secondary user's buffer at the beginning of slot t, where $\mathcal{B}$ contains all possible buffer states:

$$\mathcal{B} \stackrel{\Delta}{=} \{0 \text{ (empty)}, 1, \ldots, l\}. \tag{4}$$

Channel Model We adopt a block channel fading model. Specifically, we assume that the channel gain between the secondary user and its destination is a random variable (rv) identically and independently distributed (iid) across slots but not necessarily iid across channels.

Energy Model The secondary user is powered by a battery with initial energy $\mathcal{E}_0$. We consider three types of energy consumption by the secondary user in a slot. Let e_p denote the energy consumed in the sleeping mode and e_s the energy consumed in sensing the occupancy of a channel. The energy consumed in transmitting over channel n is denoted by $E_{tx}(n)$.

We assume that the secondary user only has a finite number L of transmission power levels due to hardware and power limitations. According to the fading condition, the secondary user adjusts its transmission power to ensure successful reception at its destination. In general, the better the fading condition, the lower the transmission power level. The transmission energy consumption $E_{tx}(n)$ is a rv taking values from a finite set $\mathcal{E}_{tx}$:

$$\mathcal{E}_{tx} \stackrel{\Delta}{=} \{\varepsilon_k\}_{k=1}^L, \quad 0 < \varepsilon_1 < \ldots < \varepsilon_L \le \infty, \tag{5}$$

where ε_k is the energy consumed in transmitting at the kth power level. By setting $\varepsilon_L = \infty$, we can include the case where the channel is so badly faded that no transmission is allowed. The distribution of the transmission energy consumption $E_{tx}(n)$ is determined by the channel distribution, and is denoted by

$$p_n(k) \stackrel{\Delta}{=} \Pr\{E_{tx}(n) = \varepsilon_k\}, \quad k = 1, \ldots, L, \tag{6}$$

where $\sum_{k=1}^L p_n(k) = 1$.

Let $E(t)$ denote the secondary user's residual energy at the beginning of slot t. Note that $E(t)$ is a rv depending on the fading conditions and the secondary user's actions in all previous slots. Since the transmission energy consumption is restricted to the set $\mathcal{E}_{tx}$, the residual energy $E(t)$ belongs to

$$\begin{aligned} \mathcal{E} \stackrel{\Delta}{=} \{E : E = \mathcal{E}_0 - \sum_{k=1}^L c_k\varepsilon_k - c_s e_s - c_p e_p \ge 0; \\ c_s \ge \sum_{k=1}^L c_k, c_k, c_s, c_p \in \mathbb{Z};\ c_k, c_s, c_p \ge 0\} \cup \{0\}, \end{aligned} \tag{7}$$

where c_k is the number of slots when the secondary user transmits at the kth power level, c_s is the number of slots when the secondary user senses a channel, and c_p is the number of slots when the secondary user operates in the sleeping mode. Note that the secondary user must sense the channel before accessing it in order to avoid collisions with the primary users. We thus have $c_s \ge \sum_{k=1}^L c_k$.

III. A Decision-Theoretic Framework

In this section, we formulate the energy-constrained OSA design as an unconstrained POMDP.

A. Sequential Decision-Making

We illustrate in Fig. 1 the sequence of operations in each slot. At the beginning of slot t, the SOS transits to $\mathbf{S}(t) \in \mathcal{S}$ according to the Markovian primary traffic model $P_{\mathbf{S}}(\mathbf{s}'|\mathbf{s})$.

Sensing Decision Based on its knowledge of the SOS and its local buffer state $B(t)$ and residual energy $E(t)$, the secondary user first chooses a channel $A(t)$ to sense:

$$A(t) \in \{0 \text{ (sleeping mode)}, 1, \ldots, N\}, \tag{8}$$

[1] Our formulation can be readily extended to the case where the secondary user's packet arrivals follow a Markov-modulated Poisson process (MMPP). See Section III-B for details.

where $A(t)=0$ represents the sleeping mode.

Sensing Observation If a channel $A(t)=a>0$ is sensed, the secondary user observes the channel occupancy and fading condition. The sensing outcome is denoted by

$$\Theta(t)\in\{0\text{ (busy)},1,\ldots,L\}, \tag{9}$$

where $\Theta(t)=0$ indicates that the chosen channel is busy, and $\Theta(t)=k>0$ indicates that the chosen channel is idle and the fading condition requires the secondary user to transmit at the kth power level. We assume perfect spectrum sensing. Hence, the distribution $U(k|\mathbf{s},a)$ of the sensing outcome $\Theta(t)$ given current SOS and chosen channel $A(t)=a>0$ is obtained as:

$$\begin{aligned}U(k|\mathbf{s},a)&\overset{\Delta}{=}\Pr\{\Theta(t)=k\,|\,\mathbf{S}(t)=\mathbf{s},A(t)=a\}\\&=\begin{cases}p_a(k), & \text{if } s_a=1,k\neq 0,\\ 1, & \text{if } s_a=0,k=0.\end{cases}\end{aligned} \tag{10}$$

Access Decision Based on the sensing outcome $\Theta(t)$, the secondary user determines whether to transmit over the chosen channel $A(t)>0$:

$$\Phi_\Theta(t)\in\{0\text{ (no access), }1\text{ (access)}\}. \tag{11}$$

Let $\mathbf{\Phi}(t)\overset{\Delta}{=}[\Phi_0(t),\Phi_1(t),\ldots,\Phi_L(t)]$ denote the set of access decisions, one for each possible sensing outcome $\Theta(t)\in\{0,\ldots,L\}$. Clearly, when $\Theta(t)=0$ (busy), the secondary user should refrain from transmission, *i.e.,* $\Phi_0(t)=0$. We also note that the secondary user should not transmit (*i.e.,* $\Phi_\Theta(t)=0$) when it does not have enough energy to combat the current channel fading (*i.e.,* $E(t)<e_s+\varepsilon_\Theta$) or its buffer is empty (*i.e.,* $B(t)=0$). Let $\mathbb{A}_c(B(t),E(t))$ denote the access action space, which includes all allowable access decisions $\mathbf{\Phi}(t)$ given current buffer state $B(t)$ and residual energy $E(t)$:

$$\begin{aligned}\mathbb{A}_c(B(t),E(t))\overset{\Delta}{=}\{&\mathbf{\Phi}=[\Phi_0,\ldots,\Phi_L]\in\{0,1\}^{L+1}:\Phi_0=0;\\&\Phi_k=0\text{ if }E(t)<e_s+\varepsilon_k\text{ or }B(t)=0\}.\end{aligned} \tag{12}$$

Information Update At the end of each slot, the secondary user can update its knowledge of the SOS by incorporating the decisions and observations made in this slot (see Section III-B for details). The secondary user's local state $(B(t),E(t))$ also changes due to the packet arrivals and energy consumption in this slot. Specifically, since the packet arrival process is assumed to be Poisson, the number of arrivals is iid across slots. Hence, the evolution of the buffer state is a Markov process whose transition probabilities are given by

$$\begin{aligned}P_B^i(b'|b)&\overset{\Delta}{=}\Pr\{B(t+1)=b'|B(t)=b,i\text{ packet was sent}\}\\&=\sum_{m=0}^{\infty}q_m 1_{[b'=\max\{b-i+m,l\}]},\quad b,b'\in\mathcal{B},\end{aligned} \tag{13}$$

where $i=0,1$ is the number of packets delivered in this slot, and l is the maximum buffer size. The residual energy reduces from $E(t)$ to

$$\begin{aligned}E(t+1)&=\mathcal{T}_E(E(t)|A(t),\mathbf{\Phi}(t),\Theta(t))\\&\overset{\Delta}{=}\begin{cases}E(t)-e_p, & \text{if } A(t)=0,\\ E(t)-e_s-1_{[\Phi_\Theta=1]}\varepsilon_\Theta & \text{otherwise,}\end{cases}\end{aligned} \tag{14}$$

where $1_{[X]}$ is an indicator function, $1_{[\Phi_\Theta=1]}$ indicates whether the secondary user has accessed the chosen channel, and ε_Θ is the energy required for a successful transmission. Note that no observations and access decisions are made when the secondary user is in the sleeping mode. For simplicity, we will write $\mathcal{T}_E(E(t)|A(t),\mathbf{\Phi}(t),\Theta(t))$ as $\mathcal{T}_E(E(t)|0)$ when $A(t)=0$.

The updated SOS knowledge, buffer state $B(t+1)$, and residual energy $E(t+1)$ are then used to make optimal decisions in the next slot $t+1$. The above procedure repeats until the secondary user is incapable of successful transmission under any channel fading condition, *i.e.,* its residual energy $E(t)$ drops below the minimum energy required to sense and access a channel: $e_s+\min\mathcal{E}_{tx}=e_s+\varepsilon_1$.

B. A POMDP Formulation

The sequential decision-making process described in Section III-A can be cast in the framework of POMDP. Specifically, the system state can be characterized by the following three components: 1) the SOS of the primary network $\mathbf{S}(t)\in\mathcal{S}$; 2) the buffer state $B(t)\in\mathcal{B}$ of the secondary user; and 3) the residual energy $E(t)\in\mathcal{E}$ of the secondary user[2]. While the buffer state and the residual energy are fully observable to the secondary user, the current SOS cannot be directly observed due to partial spectrum monitoring. We thus have a POMDP with composite system state space $\mathbb{S}$:

$$\mathbb{S}\overset{\Delta}{=}\{(\mathbf{S},B,E):\mathbf{S}\in\mathcal{S},B\in\mathcal{B},E\in\mathcal{E}\}, \tag{15}$$

where $\mathcal{S},\mathcal{B},\mathcal{E}$ are defined in (1), (4), and (7) respectively.

Sufficient Statistics At the beginning of each slot t, the secondary user's knowledge of the SOS is provided by its decision and observation history[3] $Y(t)\overset{\Delta}{=}\{A(\tau),\Theta(\tau)\}_{\tau=1}^{t-1}$. As shown in [9], the statistical information on the SOS can be encapsulated in a belief vector $\mathbf{\Lambda}(t)\overset{\Delta}{=}\{\Lambda_{\mathbf{s}}(t)\}_{\mathbf{s}\in\mathcal{S}}$, where $\Lambda_{\mathbf{s}}(t)\in[0,1]$ and $\sum_{\mathbf{s}\in\mathcal{S}}\Lambda_{\mathbf{s}}(t)=1$. Each element $\Lambda_{\mathbf{s}}(t)$ represents the conditional probability (given the decision and observation history) that the SOS is $\mathbf{s}$ at the beginning of this slot prior to the state transition, *i.e.,*

$$\Lambda_{\mathbf{s}}(t)\overset{\Delta}{=}\Pr\{\mathbf{S}(t-1)=\mathbf{s}|Y(t)\}. \tag{16}$$

The belief vector can be updated at the end of slot t by incorporating the sensing decision $A(t)$ and the observation $\Theta(t)$ in this slot. Specifically, applying Bayes rule, we obtain the updated belief vector $\mathbf{\Lambda}(t+1)\overset{\Delta}{=}\{\Lambda_{\mathbf{s}}(t+1)\}_{\mathbf{s}\in\mathcal{S}}$ as

$$\begin{aligned}&\mathbf{\Lambda}(t+1)=\mathcal{T}_\Lambda(\mathbf{\Lambda}(t)|A(t),\Theta(t)),\quad\text{where}\\&\Lambda_{\mathbf{s}}(t+1)=\\&\begin{cases}\sum_{\mathbf{s}'}\Lambda_{\mathbf{s}'}(t)P_{\mathbf{S}}(\mathbf{s}|\mathbf{s}'), & \text{if } A(t)=0,\\ \dfrac{\sum_{\mathbf{s}'}\Lambda_{\mathbf{s}'}(t)P_{\mathbf{S}}(\mathbf{s}|\mathbf{s}')U(k|\mathbf{s},a)}{\sum_{\mathbf{s},\mathbf{s}'}\Lambda_{\mathbf{s}'}(t)P_{\mathbf{S}}(\mathbf{s}|\mathbf{s}')U(k|\mathbf{s},a)}, & \text{otherwise.}\end{cases}\end{aligned} \tag{17}$$

For simplicity, we will write $\mathcal{T}_\Lambda(\mathbf{\Lambda}(t)|A(t),\Theta(t))$ as $\mathcal{T}_\Lambda(\mathbf{\Lambda}(t))|0)$ when $A(t)=0$.

The belief vector $\mathbf{\Lambda}(t)$ together with the fully observable buffer state $B(t)$ and residual energy $E(t)$ are thus the

[2]If packet arrivals are modeled as an MMPP, then the system state should also include the state of the underlying MMPP in addition to these three components $(\mathbf{S}(t),B(t),E(t))$.

[3]Since we have assumed perfect spectrum sensing, the current SOS information provided by the secondary user's access decisions is contained in the sensing outcome. The incorporation of the sensing decisions and the observations suffices.

sufficient statistics for making optimal sensing and access decisions. A policy $\boldsymbol{\pi}$ of the POMDP is given by a sequence of functions: $\boldsymbol{\pi} \stackrel{\Delta}{=} [\pi_1, \pi_2, \ldots]$, where each function π_t maps from the current information state $\{\boldsymbol{\Lambda}(t), B(t), E(t)\}$ to a sensing decision $A(t)$ and a set of allowable access decisions $\boldsymbol{\Phi}(t) \in \mathbb{A}_c(B(t), E(t))$ in slot t. If π_t is identical for all t, $\boldsymbol{\pi}$ is called a stationary policy.

Reward and Objective A nature definition of the reward is the number of bits delivered by the secondary user in a slot, which is assumed to be proportional to the channel bandwidth. Specifically, we define the immediate reward $R_{B,E,\Theta}^{(A,\boldsymbol{\Phi})}(t)$ as

$$R_{B,E,\Theta}^{(A,\boldsymbol{\Phi})}(t) \stackrel{\Delta}{=} 1_{[A(t)>0]} 1_{[\boldsymbol{\Phi}(t) \in \mathbb{A}_c(B(t),E(t)), \Phi_\Theta(t)=1]} B_a. \quad (18)$$

Note that $1_{[A(t)>0]} = 1$ iff the secondary user has sensed a channel, and $1_{[\boldsymbol{\Phi}(t) \in \mathbb{A}_c(B(t),E(t)), \Phi_\Theta(t)=1]} = 1$ iff the secondary user has successfully transmitted a packet.

As noted in Section III-A, the POMDP terminates, *i.e.*, no reward will be accumulated, once the residual energy $E(t)$ drops below $e_s + \varepsilon_1$. Hence, the total expected reward represents the total expected number of bits delivered by the secondary user during its battery lifetime. The objective of the POMDP can thus be written as

$$\pi^* = \arg\max_{\pi} \mathbb{E}_\pi \left[\sum_{t=1}^{\infty} R_{B,E,\Theta}^{(A,\boldsymbol{\Phi})}(t) \;\middle|\; \boldsymbol{\Lambda}(1), E(1) = \mathcal{E}_0 \right], \quad (19)$$

where $\boldsymbol{\Lambda}(1)$ is the initial belief vector, which can be set to the stationary distribution of the SOS if no information on the initial SOS is available.

IV. Optimal Energy-Constrained OSA Design

In this section, we derive recursive formulas for calculating the optimal policies of the POMDP given in (19). We also develop structural results for an efficient calculation.

A. Stationary Optimal Policy

Stationary policies are usually preferred due to reduced memory requirements and low complexity in implementation. We show that the POMDP given in (19) has a stationary optimal policy.

Proposition 1: There exist stationary optimal sensing and access policies for optimal energy-constrained OSA design.

Proof: The proof is based on the fact that the POMDP given in (19) terminates in a finite but random stopping time. See [10] for details. □

Proposition 1 enables us to focus on stationary policies without losing optimality. For brevity, we omit the time index in subsequent sections.

B. Optimality Equation

The next step to solving (19) is to express the objective explicitly as a function of the information state and the actions. Given current information state $(\boldsymbol{\Lambda}, B, E)$, we let $Q(\boldsymbol{\Lambda}, B, E|0)$ and $Q(\boldsymbol{\Lambda}, B, E|A, \boldsymbol{\Phi})$ be the maximum expected total reward that can be obtained by taking actions $A = 0$ and $\{A > 0, \boldsymbol{\Phi} \in \mathbb{A}(B, E)\}$, respectively. The value function $V(\boldsymbol{\Lambda}, B, E)$, defined as the maximum expected total reward that can be accumulated starting from information state $(\boldsymbol{\Lambda}, B, E)$, can be written in terms of the Q-functions:

$$V(\boldsymbol{\Lambda}, B, E) = \max\{Q(\boldsymbol{\Lambda}, B, E|0), \max_{\substack{A \in \{1,\ldots,N\} \\ \boldsymbol{\Phi} \in \mathbb{A}(B,E)}} Q(\boldsymbol{\Lambda}, B, E|A, \boldsymbol{\Phi})\},$$
$$V(\boldsymbol{\Lambda}, B, E) = 0, \qquad \text{if } E < e_s + \varepsilon_1. \quad (20)$$

We derive below iterative formulas for calculating the value function and the Q-functions. In the sleeping mode $A = 0$, no immediate reward will be obtained. Hence, the maximum expected total reward $Q(\boldsymbol{\Lambda}, B, E|0)$ is given by the future reward $V(\boldsymbol{\Lambda}', B', E')$, where $\{\boldsymbol{\Lambda}', B', E'\}$ represents the updated information state. Specifically, we obtain that

$$Q(\boldsymbol{\Lambda}, B, E|0) = \sum_{B' \in \mathcal{B}} P_B^0(B'|B) V(\mathcal{T}_\Lambda(\boldsymbol{\Lambda}|0), B', \mathcal{T}_E(E|0)), \quad (21)$$

where $P_B^0(B'|B)$ governs the transition of the buffer state and is given by (4), the updated belief vector $\mathcal{T}_\Lambda(\boldsymbol{\Lambda}|0)$ and residual energy $\mathcal{T}_E(E|0)$ are given by (17) and (14), respectively.

In the sensing mode $A > 0$, the maximum expected total reward $Q(\boldsymbol{\Lambda}, B, E|A, \boldsymbol{\Phi})$ consists of two parts: the immediate reward $R_{B,E,\Theta}^{(A,\boldsymbol{\Phi})}$ defined as (18) and the future reward $V(\boldsymbol{\Lambda}', B', E')$. Averaging over all possible SOS, observations, and packet arrivals, we obtain that

$$\begin{aligned} &Q(\boldsymbol{\Lambda}, B, E|A, \boldsymbol{\Phi}) \\ &= \sum_{\mathbf{s}, \mathbf{s}' \in \mathcal{S}} \Lambda_{\mathbf{s}'} P_{\mathbf{S}}(\mathbf{s}|\mathbf{s}') \sum_{k \in \mathbb{O}} U(k|\mathbf{s}, A) \Big[R_{B,E,k}^{(A,\boldsymbol{\Phi})} \\ &\quad + \sum_{B' \in \mathcal{B}} P_B^{\Phi_k}(B'|B) V(\mathcal{T}_\Lambda(\boldsymbol{\Lambda}|A, k), B', \mathcal{T}_E(E|A, \boldsymbol{\Phi}, k)) \Big]. \end{aligned} \quad (22)$$

where $P_B^{\Phi_\Theta}(B'|B)$, $\mathcal{T}_\Lambda(\boldsymbol{\Lambda}|A, \Theta)$, and $\mathcal{T}_E(E|A, \boldsymbol{\Phi}, \Theta)$ are given in (4), (17) and (7), respectively.

Using (20) – (22), we can solve the value function and the Q-functions recursively in an increasing order of the residual energy E. The optimal sensing and access decisions are then given by the maximizers of (20). Algorithms for solving POMDPs exist in the literature [9] and are applicable here.

C. Monotonicity Results on Optimal Design

While powerful in problem modeling, POMDPs are generally computationally expensive. Structural results are thus desirable since they can provide insights into the underlying problem and accelerate computations [11]. By exploiting the rich structure of the energy-constrained OSA problem, we develop monotonicity results for the optimal sensing and access policies in Propositions 2 and 3.

Proposition 2: **Threshold Optimal Sensing Policy**

Consider the single-channel ($N = 1$) and single-buffer ($l = 1$) case. The optimal decision on whether to sense is a threshold policy in terms of the conditional probability that the channel is available. Specifically, given buffer state $B = 1$ and residual energy E, there exists a threshold $r_{th} \in [\min\{P_S(1|0), P_S(1|1)\}, \max\{P_S(1|0), P_S(1|1)\}]$ such that the optimal sensing decision A^* is given by

$$A^* = \begin{cases} 1 & \text{if } \Lambda_0 P_S(1|0) + \Lambda_1 P_S(1|1) \geq r_{th} \\ 0 & \text{otherwise}, \end{cases} \quad (23)$$

where $\Lambda_0 P_S(1|0) + \Lambda_1 P_S(1|1)$ is the probability that the channel is available given current belief vector $\boldsymbol{\Lambda} = [\Lambda_0, \Lambda_1]$.

Proof: See [10]. □

Recall that a stationary sensing policy is given by a function that specifies a sensing decision A for each possible information state $\{\mathbf{\Lambda}, B, E\}$ (or equivalently $\{\Lambda_1, B, E\}$ since $\Lambda_0 = 1 - \Lambda_1$ when $N = 1$). Proposition 2 indicates that the optimal sensing policy can also be represented by a function mapping from the secondary user's local state (B, E) to a threshold r_{th} on the sensing decisions. Since the threshold $r_{th} \in [\min\{P_S(1|0), P_S(1|1)\}, \max\{P_S(1|0), P_S(1|1)\}]$ belongs to a subset of the belief space $\Lambda_1 \in [0, 1]$, the search for the optimal threshold r_{th} is less complex than finding the optimal decision for each belief vector.

Proposition 3: **Threshold Optimal Access Policy**

For a given sensing action $A > 0$, the optimal access decision is non-decreasing in the channel fading condition. Specifically, given belief vector $\mathbf{\Lambda}$ and residual energy E, there exists a threshold $k_{th} \in \{1, \ldots, L\}$ such that the optimal access decision Φ_k^* is given by

$$\Phi_k^* = \begin{cases} 1 & \text{if } k \le k_{th} \\ 0 & \text{otherwise.} \end{cases} \tag{24}$$

Furthermore, the threshold k_{th} is independent of the belief vector in the single-channel case ($N = 1$).

Proof: See [10]. □

Proposition 3 extends [4] by considering the buffer state in the design of energy-constrained OSA. It enables us to reduce the access action space $\mathbb{A}_c(B, E)$ to

$$\begin{aligned} \mathbb{A}_c(B, E) = \{&\mathbf{\Phi} : \Phi_0 = 0; 1 \ge \Phi_1 \ge \ldots \ge \Phi_L \ge 0; \\ &\Phi_k = 0 \text{ if } E(t) < e_s + \varepsilon_k \text{ or } B(t) = 0\}. \end{aligned} \tag{25}$$

Hence, the size of the access action space is reduced from exponential 2^L as given in (12) to linear L in the number of power levels, leading to a more efficient search for the optimal access policy.

Furthermore, Proposition 3 indicates that the optimal access policy is independent of the belief vector when $N = 1$. That is, the optimal access policy can be specified by a function mapping from the secondary user's local state (B, E) to a threshold k_{th} for the access decisions. Since there are only finitely many local states (B, E), the complexity of calculating the optimal access policy can be significantly reduced.

D. Distributed Implementation

As seen from (20), the information state $\{\mathbf{\Lambda}, B, E\}$ governs the channel selection. Hence, to ensure synchronous hopping in the spectrum without introducing extra control message exchange, the secondary transmitter and its desired receiver must maintain the same information state in each slot. In [4], we have described how to achieve synchronous update of the belief vector $\mathbf{\Lambda}$ and the residual energy E at the transmitter and the receiver. Below we briefly comment on the synchronous update of the buffer state B in an ad hoc secondary network where there is no central coordinator or dedicated communication/control channel. For a detailed description of distributed implementation, readers are referred to [3].

Due to the random packet arrival process, the receiver does not know the exact buffer state B of the transmitter. Hence, to ensure synchronous hopping, the transmitter should use the receiver's knowledge of the buffer state $\tilde{B}$ for decision-making in each slot. Meanwhile, the transmitter should inform the receiver of its true buffer state B so that the receiver can update its knowledge. We propose the use of the request-to-send (RTS) and clear-to-send (CTS) messages to synchronize the buffer states at the transmitter and the receiver. Specifically, the transmitter piggybacks its true buffer state B to every RTS message in the opportunity identifying stage (see [3], [4] for details). The receiver will confirm the reception of the buffer state in its clear-to-send (CTS) message. The buffer state used for decision-making is then updated $\tilde{B} = B$ at both the transmitter and the receiver after the successful RTS-CTS exchange. In the case when the transmitter fails to update the receiver's knowledge of the buffer state $\tilde{B}$ for a long period of time, we can reset the buffer state $\tilde{B}$ used for decision-making according to the transmitter's traffic statistics.

V. Numerical Results

In this section, we provide numerical results to study the impact of the secondary user's traffic statistics λ and residual energy E on the optimal energy-constrained OSA design. In all figures, the optimal sensing and access decisions are determined by solving (20) recursively for the information state $\{\mathbf{\Lambda}, B, E\}$ of interest. We assume that the secondary user has a single-size buffer *i.e.*, $l = 1$.

A. Optimal Decisions for Non-Empty Buffer

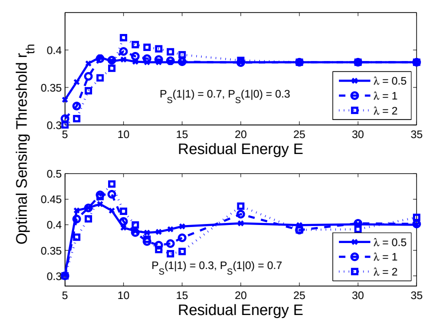

Fig. 2. Optimal thresholds r_{th} for making sensing decisions A^* when the buffer is non-empty. $e_s = 0.5$, $e_p = 0.1$, $\mathcal{E}_{tx} = \{1, 2, 3, 4\}$, $[p_n(1), p_n(2), p_n(3), p_n(4)] = [0.2, 0.3, 0.3, 0.2]$.

We first consider the case where the secondary user's buffer is non-empty. We consider the single-channel case $N = 1$ in which the SOS transition is characterized by $P_S(i|j) = \Pr\{S(t+1) = i | S(t) = j\}$, $i, j = 0, 1$. The optimal sensing and access policies are thus given by the thresholds r_{th} and k_{th} as stated in Propositions 2 and 3.

In Fig. 2, we plot the optimal sensing threshold r_{th} as a function of the residual energy E for different packet arrival rates λ. In the upper plot, we consider the cases where $P_S(1|1) = 0.7$ and $P_S(1|0) = 0.3$, *i.e.*, the channel occupancy state remains unchanged with a large probability. The opposite case where $P_S(1|1) = 0.3$ and $P_S(1|0) = 0.7$ is considered in the lower plot. We see that when the residual energy E is small, the optimal threshold r_{th} is highly dependent on the packet arrival rate λ. As residual energy E increases, the impact of λ and E on the optimal threshold r_{th} diminishes.

We notice that the optimal thresholds r_{th} for different packet arrival rates converge to a common steady value when the residual energy E is large.

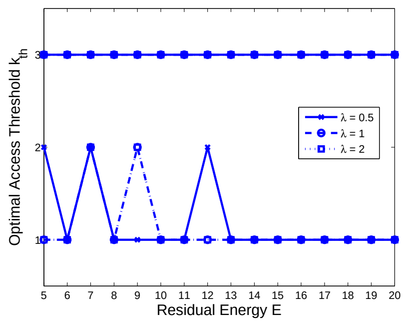

Fig. 3. Optimal thresholds k_{th} for making access decisions Φ_k^* when the buffer is non-empty. $P_S(1|1) = 0.7$, $P_S(1|0) = 0.3$, $e_p = 0.1$, $\mathcal{E}_{tx} = \{1, 2, 3, 4\}$, $[p_n(1), p_n(2), p_n(3), p_n(4)] = [0.2, 0.3, 0.3, 0.2]$.

In Fig. 3, we plot the optimal access threshold k_{th} for different packet arrival rates λ. As expected, the optimal threshold k_{th} increases with the sensing energy consumption e_s (see [4] for explanation). Similar to the optimal sensing thresholds r_{th}, the optimal access thresholds k_{th} for different packet arrival rates λ may differ from each other when the residual energy E is small but a common steady value will be reached when E is large.

Combining Figs. 2 and 3, we see that the impact of the residual energy E and the traffic statistics λ on the optimal sensing and access decisions is negligible when the residual energy E is sufficiently large. This observation suggests a complexity-reduced OSA strategy. Specifically, the secondary user only needs to calculate and store the optimal policies for small residual energies $E \leq E^*$. When $E > E^*$, the secondary user can simply adopt the optimal decisions for $E = E^*$.

B. Optimal Decisions for Empty Buffer

We note that even if the buffer is empty, the secondary user may want to sense a channel in order to gain information on the SOS for future use. Next, we study the optimal decision $1_{[A^*>0]}$ on whether to sense when $B = 0$.

Consider two coupled channels where the SOS is either $\mathbf{S}(t) = [0, 1]$ (*i.e.*, only channel 2 is idle) or $\mathbf{S}(t) = [1, 0]$ (*i.e.*, only channel 1 is idle). We assume that $P_S([1, 0]|[0, 1]) = P_S([0, 1]|[1, 0]) = \alpha$ so that the correlation between the SOS in two successive slots can be characterized by a single parameter $\rho = 1 - 2\alpha$. Extensive numerical results show that the optimal decision $1_{[A^*>0]}$ on whether to sense is non-decreasing in the SOS correlation $|\rho|$. Specifically, given the secondary user's residual energy E, there exists a threshold $\rho_{th} \in [0, 1]$ such that

$$1_{[A^*>0]} = \begin{cases} 1, & \text{if } |\rho| \geq \rho_{th}, \\ 0, & \text{otherwise.} \end{cases} \tag{26}$$

In Fig. 4, we plot the threshold ρ_{th} on the SOS correlation as a function of the residual energy E for different packet arrival rates λ. We see that the threshold ρ_{th} decreases with λ. Intuitively, when λ is large, there is a high probability that packets will arrive in this slot, and hence the secondary user

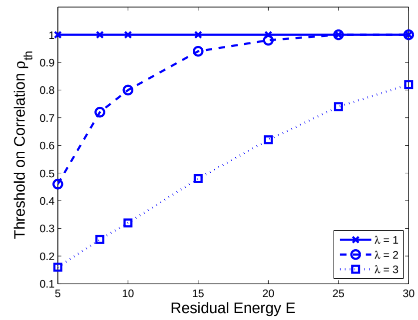

Fig. 4. Thresholds ρ_{th} on the SOS correlation for making optimal sensing decisions $1_{[A^*>0]}$ when the buffer is empty. The initial belief vector $\mathbf{\Lambda}(1)$ is given by the stationary distribution of the underlying Markov process. $e_p = 0.1$, $e_s = 0.2$, $\mathcal{E}_{tx} = \{1, 2\}$, $[p_n(1), p_n(2)] = [0.6, 0.4]$.

should be more active in collecting information on the SOS for better channel selection in the next slot. We also observe that the threshold ρ_{th} increases with the residual energy E.

VI. Conclusion

Within the framework of POMDP, we incorporated the bursty traffic of secondary users in the design of energy-constrained OSA. We developed monotonicity results on the optimal sensing and access policies for efficient computation. Numerical results revealed that the impact of the secondary user's traffic statistics and residual energy on the optimal sensing and access decisions diminishes when the residual energy is large.

References

[1] "DARPA: The Next Generation (XG) Program." http://www.darpa.mil/sto/smallunitops/xg.html.

[2] P. Papadimitratos, S. Sankaranarayanan, and A. Mishra, "A bandwidth sharing approach to improve licensed spectrum utilization," *IEEE Communications Magazine*, vol. 43, pp. 10-14, Dec. 2005.

[3] Q. Zhao, L. Tong, A. Swami, and Y. Chen "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 3, pp. 589-600, Apr. 2007.

[4] Y. Chen, Q. Zhao, and A. Swami, "Distributed cognitive MAC for energy-constrained opportunistic spectrum access," in *Proceedings of MILCOM*, Washington, DC, pp. 1 - 7, Oct. 2006.

[5] Y. Chen, Q. Zhao, and A. Swami, " Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors," *submitted to IEEE Transactions on Information Theory*, Feb. 2007; available at http://www.ece.ucdavis.edu/∼qzhao/Journal.html.

[6] Q. Zhao, S. Geirhofer, L. Tong, and B. M. Sadler, "Optimal dynamic spectrum access via periodic channel sensing," in *Proc. of Wireless Communications and Networking Conference (WCNC)*, Hong Kong, pp. 33-37, Mar. 2007.

[7] S. Geirhofer, L. Tong, and B. M. Sadler, "Dynamic spectrum access in WLAN channels: empirical model and its stochastic analysis," in *Proceedings of the First International Workshop on Technology and Policy in Accessing Spectrum (TAPAS), Boston, MA, Aug. 2006.*

[8] Q. Zhao and B. Sadler, "A survey of dynamic spectrum access: signal processing, networking, and regulatory policy," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 79-89, May 2007.

[9] E. J. Sondik, *The optimal control of partially observable Markov processes*, PhD thesis, Stanford University, Stanford, California, 1971.

[10] Y. Chen, Q. Zhao, and A. Swami, "Bursty traffic in energy-constrained opportunistic spectrum access," technical report, TR-07-03, University of California, Davis, Mar. 2007. http://www.ece.ucdavis.edu/∼qzhao/Report.html

[11] W. S. Lovejoy, "Some monotonicity results for partially observed Markov decision processes," *Operations Research*, vol. 35, no. 5, pp. 736 - 743, Sep. - Oct. 1987.