Bursty Traffic in Energy-Constrained Opportunistic Spectrum Access

Yunxia Chen, Qing Zhao, and Ananthram Swami

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 compatible 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 (see [2]–[8] and references therein). In [2], the authors address the implementation of OSA in an ad hoc secondary network overlaying a GSM cellular network. In [3], [4], 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, [5] exploits the channel fading in the design of OSA for an efficient use of the secondary user's energy. In [6], 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 [7], access strategies for a slotted secondary user searching for opportunities in an un-slotted primary network is considered, where a round-robin single-channel sensing scheme is used. Modeling of spectrum

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

occupancy has been addressed in [8]. Measurements obtained from spectrum monitoring test-beds demonstrate the Makovian transition between busy and idle channel states in wireless LAN. For an overview on challenges and recent developments in OSA, readers are referred to [9].

Contributions This report extends [5] by incorporating the secondary user's traffic statistics into the energy-constrained OSA design. Taking into account the channel fading statistics and the secondary user's traffic and energy consumption characteristics, we formulate the sequential sensing and access decision-making problem as a POMDP. 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 develop monotonicity results for the optimal policies. In particular, we show for the one-channel case that there exists a threshold on the conditional probability that the channel is available, above which the secondary user should always choose to sense. There also exists a threshold on the channel fading condition, above which the secondary user should never transmit over this channel. These monotonicity results can help us accelerate the calculation of the optimal policies.

Numerical results show that the impact of the secondary users' residual energy and traffic statistics on its sensing and access decisions diminishes as the residual energy increases. We also investigate the cases where the secondary user may choose to sense channels even if its buffer is empty.

II. System Model

A. Primary Network Model

We consider a spectrum consisting of N channels (*e.g.*, different frequency bands or tones in an OFDM system), 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 occupancy state (SOS) $\mathbf{S}(t) \triangleq [S_1(t), \ldots, S_N(t)]$ follows a time-homogeneous discrete Markov process with state space $\mathcal{S}$ defined as

$$\mathcal{S} \triangleq \{0,1\}^N, \qquad \text{where } |\mathcal{S}| = 2^N. \tag{1}$$

The transition probabilities are denoted as

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

which represents the probability that the SOS transits from $\mathrm{s} \in \mathcal{S}$ to $\mathrm{s}' \in \mathcal{S}$ 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 independently and selfishly seek instantaneous spectrum opportunities in these N channels. At the beginning of each slot, the secondary user chooses M ($1 \leq M \leq N$) 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 Figure 1 and will be detailed in Section III-B.

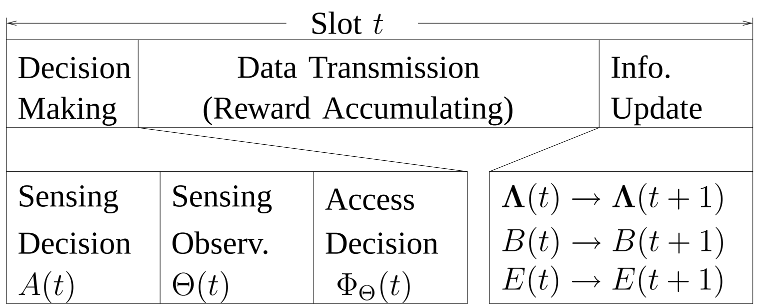

Fig. 1. The slot structure. The secondary user's knowledge of the SOS is characterized by $\mathbf{\Lambda}(t)$, and its buffer state and residual energy are denoted by $B(t)$ and $E(t)$, respectively.

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 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 The bursty traffic of the secondary user is modeled as a Poisson process with rate λ. That is, the probability of m packet arrivals in a slot is given by

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

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 (i.i.d.) across slots but not necessarily i.i.d. 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 current 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. Hence, the transmission energy consumption $E_{tx}(n)$ is a RV taking values from a finite set $\mathcal{E}_{tx}$:

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

where ε_k is the energy consumed in transmitting at the kth power level. The distribution of the transmission energy consumption $E_{tx}(n)$ is determined by the channel fading 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 a finite set $\mathcal{E}$:

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

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 $c_s \geq \sum_{k=1}^{L} c_k$.

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 Figure 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 underlying 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:

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

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 $U(k|\mathrm{s}, a)$ of the sensing outcome $\Theta(t)$ given current SOS and chosen channel $A(t) = a > 0$ is obtained as:

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

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 $\mathbf{\Phi}(t) \stackrel{\Delta}{=} [\Phi_0(t), \Phi_1(t), \ldots, \Phi_L(t)]$ 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 $E(t)$:

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

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 consumptions in this slot. Specifically, since the packet arrival process is assumed to be Poisson, the number of arrivals is i.i.d. across slots. Hence, the evolution of the buffer state is a Markov process whose transition probabilities are given by

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

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

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

where $1_{[\Phi_\Theta=1]}$ 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 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 conditions, ***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. 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 $\mathbb{S}$:

$$\mathbb{S} \triangleq \{(\mathbf{S}, B, E) : \mathbf{S} \in \mathcal{S}, B \in \mathcal{B}, E \in \mathcal{E}\}, \quad (15)$$

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[1] $Y(t) \triangleq \{A(\tau), \Theta(\tau)\}_{\tau=1}^{t-1}$. As shown in [10],

[1]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.

the statistical information on the SOS can be encapsulated in a belief vector $\mathbf{\Lambda}(t) \stackrel{\Delta}{=} \{\Lambda_{\mathbf{s}}(t)\}_{\mathbf{s}\in\mathcal{S}}$, 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.*,

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

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 $\mathbf{\Lambda}(t+1) \stackrel{\Delta}{=} \{\Lambda_{\mathbf{s}}(t+1)\}_{\mathbf{s}\in\mathcal{S}}$ as

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

Note that $\mathcal{T}_\Lambda(\mathbf{\Lambda}(t)|A(t),\Theta(t))$ will be written 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 sufficient statistics for making optimal sensing and access decisions. A policy $\boldsymbol{\pi}$ of the POMDP is given by a sequence of functions:

$$\boldsymbol{\pi} \stackrel{\Delta}{=} [\pi_1, \pi_2, \ldots], \tag{18}$$

where each function π_t maps from the current information state $\{\mathbf{\Lambda}(t), B(t), E(t)\}$ to a sensing decision $A(t)$ and a set of allowable access decisions $\mathbf{\Phi}(t) \in \mathbb{A}_c(B(t), E(t))$ 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, given buffer state $B(t)$ and residual energy $E(t)$, the immediate reward $R^{(A,\mathbf{\Phi})}_{B,E,\Theta}(t)$ can be defined as a function of action $\{A(t), \mathbf{\Phi}(t)\}$ and sensing outcome $\Theta(t)$:

$$R^{(A,\mathbf{\Phi})}_{B,E,\Theta}(t) \stackrel{\Delta}{=} 1_{[A(t)>0]} 1_{[\mathbf{\Phi}(t)\in\mathbb{A}_c(B(t),E(t)),\Phi_\Theta(t)=1]} B_a, \tag{19}$$

where $1_{[x]}$ is the indicator function of x. Note that $1_{[A(t)>0]} = 1$ iff the secondary user has sensed a channel, and $1_{[\mathbf{\Phi}(t)\in\mathbb{A}_c(B(t),E(t)),\Phi_\Theta(t)=1]} = 1$ 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

$$\pi^* = \arg\max_{\pi} \mathbb{E}_\pi \left[\sum_{t=1}^{\infty} R_{B,E,\Theta}^{(A,\mathbf{\Phi})}(t) \;\middle|\; \mathbf{\Lambda}(1), E(1) = \mathcal{E}_0 \right], \tag{20}$$

where $\mathbf{\Lambda}(1)$ 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 (20). We also develop structural results for efficiently calculating the optimal policies.

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 (20) has a stationary optimal policy.

Proposition 1: For the energy-constrained OSA design given by (20)*, there exist stationary optimal sensing and access policies.*

Proof: The proof is based on the fact that the POMDP given in (20) terminates in a finite but random stopping time. See Appendix A 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 (20) is to express the objective explicitly as a function of the information state and the actions. Given current information state $(\mathbf{\Lambda}, B, E)$, we let $Q(\mathbf{\Lambda}, B, E|0)$ and $Q(\mathbf{\Lambda}, B, E|A, \mathbf{\Phi})$ be the maximum expected total reward that can be obtained by taking actions $A=0$ and $\{A>0, \mathbf{\Phi} \in \mathbb{A}(B,E)\}$, respectively. The value function $V(\mathbf{\Lambda}, B, E)$, defined as the maximum expected total reward that can be accumulated starting from information state $(\mathbf{\Lambda}, B, E)$, can be written in terms of the Q-functions:

$$\begin{aligned} & V(\mathbf{\Lambda}, B, E) \\ & = \max\{Q(\mathbf{\Lambda}, B, E|0), \max_{\substack{A\in\{1,\ldots,N\} \\ \mathbf{\Phi}\in\mathbb{A}(B,E)}} Q(\mathbf{\Lambda}, B, E|A, \mathbf{\Phi})\}, \quad (21a) \\ & V(\mathbf{\Lambda}, B, E) = 0, \qquad \text{if } E < e_s + \varepsilon_1. \quad (21b) \end{aligned}$$

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 $Q(\mathbf{\Lambda}, B, E|0)$ is given by the future reward $V(\mathbf{\Lambda}', B', E')$, where $\{\mathbf{\Lambda}', B', E'\}$ represents the updated information state. Specifically, we obtain that

$$Q(\mathbf{\Lambda}, B, E|0) = \sum_{B'\in\mathcal{B}} P_B^0(B'|B) V(\mathcal{T}_\Lambda(\mathbf{\Lambda}|0), B', \mathcal{T}_E(E|0)), \quad (22)$$

where $P_B^0(B'|B)$ governs the buffer state transition and is given by (4), the updated belief vector $\mathcal{T}_\Lambda(\mathbf{\Lambda}|0)$ 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 $Q(\mathbf{\Lambda}, B, E|A, \mathbf{\Phi})$ consists of two parts: the immediate reward $R_{B,E,\Theta}^{(A,\mathbf{\Phi})}$ defined as (19) and the future reward $V(\mathbf{\Lambda}', B', E')$. Averaging over all possible SOS, observations, and packet arrivals, we obtain that

$$\begin{aligned} & Q(\mathbf{\Lambda}, B, E|A, \mathbf{\Phi}) \\ & = \sum_{\mathbf{s},\mathbf{s}'\in\mathcal{S}} \Lambda_{\mathbf{s}'} P_{\mathbf{S}}(\mathbf{s}|\mathbf{s}') \sum_{k\in\mathbb{O}} U(k|\mathbf{s}, A) \Big[R_{B,E,k}^{(A,\mathbf{\Phi})} \\ & + \sum_{B'\in\mathcal{B}} P_B^{\Phi_k}(B'|B) V(\mathcal{T}_\Lambda(\mathbf{\Lambda}|A,k), B', \mathcal{T}_E(E|A, \mathbf{\Phi}, k)) \Big]. \end{aligned} \quad (23)$$

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

Using (21) – (23), 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 (21). Algorithms for solving POMDPs exist in the literature [10] 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 [12]. By exploiting the rich structure of the energy-constrained OSA problem, we develop monotonicity results for the optimal sensing and access policies.

Proposition 2: **Threshold Optimal Sensing Policy**

Consider the one-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

$$A^* = \begin{cases} 1 & \textit{if } \Lambda_0 P_S(1|0) + \Lambda_1 P_S(1|1) \geq r_{th} \\ 0 & \textit{otherwise}, \end{cases} \tag{24}$$

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 $\mathbf{\Lambda} = [\Lambda_0, \Lambda_1]$.

Proof: See Appendix B. □

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 $\Lambda_0 = 1 - \Lambda_1$). Proposition 3 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-increasing on 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

$$\Phi_k^* = \begin{cases} 1 & \textit{if } k \leq k_{th} \\ 0 & \textit{otherwise}. \end{cases} \tag{25}$$

Furthermore, the threshold k_{th} is independent of the belief vector when $N = 1$ (i.e., the one-channel case).

Proof: See Appendix C. □

Proposition 3 extends [5] by considering the buffer state in the energy-constrained OSA design. It enables us to reduce the access action space $\mathbb{A}_c(B, E)$ to

$$\begin{aligned} \mathbb{A}_c(B, E) = \{\mathbf{\Phi} : \Phi_0 = 0; 1 \geq \Phi_1 \geq \ldots \geq \Phi_L \geq 0; \\ \Phi_k = 0 \text{ if } E(t) < e_s + \varepsilon_k \text{ or } B(t) = 0\}. \end{aligned} \tag{26}$$

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.

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 (21) 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

We first consider the case where the secondary user's buffer is non-empty $B > 0$. For simplicity, we focus on 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 proven in Propositions 2 and 3.

In Figure 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 $P_S(1|1) = 0.7$

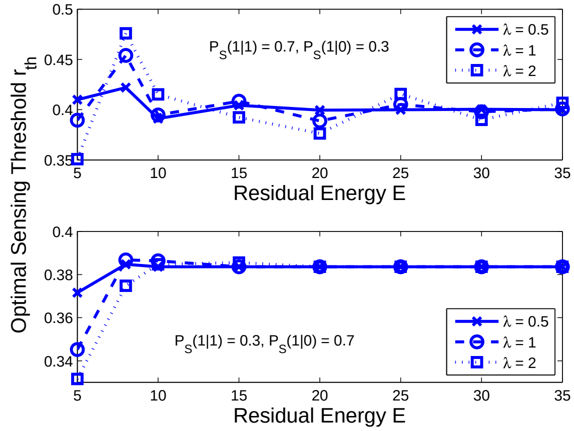

Fig. 2. Optimal thresholds r_{th} for making sensing decisions A^* when the buffer is non-empty. $e_s = 0.5$, $e_p = 0.1$, $\mathcal{E}_{tx} = \{1, 2, 3, 4\}$, $[p_n(1), p_n(2), p_n(3), p_n(4)] = [0.4, 0.3, 0.2, 0.1]$.

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 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.

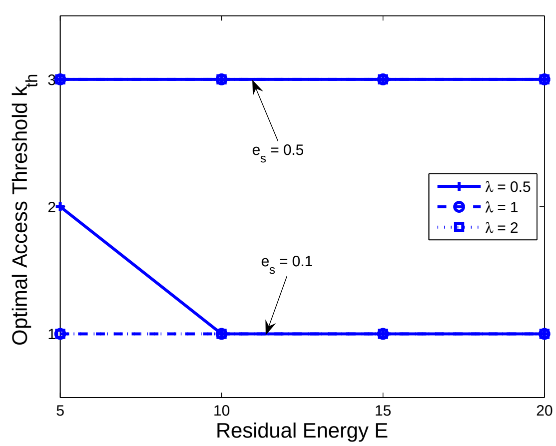

Fig. 3. Optimal thresholds k_{th} for making access decisions Φ_k^* when the buffer is non-empty. $P_S(1|1) = 0.7$, $P_S(1|0) = 0.3$, $e_p = 0.1$, $\mathcal{E}_{tx} = \{1, 2, 3, 4\}$, $[p_n(1), p_n(2), p_n(3), p_n(4)] = [0.2, 0.3, 0.3, 0.2]$.

In Figure 3, we plot the optimal access threshold k_{th} for different packet arrival rates λ. As expected, the optimal threshold k_{th} with the sensing energy consumption e_s (see [5] 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 Figures 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 leads to a complexity-reduced OSA strategy, which can achieve nearly optimal performance. Specifically, the secondary user only needs to calculate the optimal thresholds r_{th} and k_{th} for small residual energies $E \leq E^*$. When $E > E^*$, the secondary user can simply adopt the optimal sensing and access thresholds computed for $E = E^*$.

B. Optimal Decisions for Empty Buffer

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

Consider two coupled channels $N = 2$ 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 on the SOS correlation $|\rho|$. Specifically, given the secondary user's residual energy E, there exists a threshold $\rho_{th} \in [0, 1]$ such that

$$1_{[A^*>0]} = \begin{cases} 1, & \text{if } |\rho| \geq \rho_{th}, \\ 0, & \text{otherwise.} \end{cases} \tag{27}$$

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

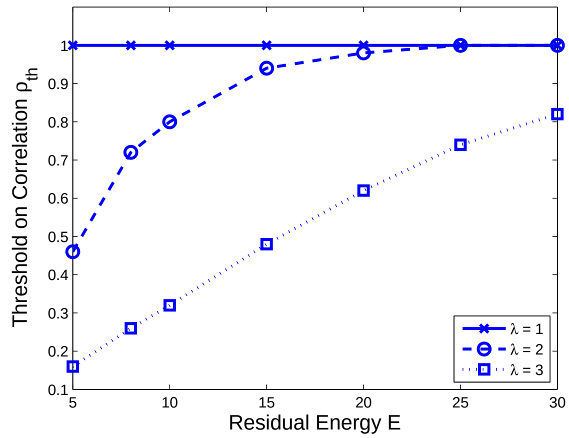

Fig. 4. Thresholds ρ_{th} on the SOS correlation for making optimal sensing decisions $1_{[A^*>0]}$ when the buffer is empty. The belief vector $\mathbf{\Lambda}$ is given by the stationary distribution. $e_p = 0.1$, $e_s = 0.2$, $\mathcal{E}_{tx} = \{1, 2\}$, $[p_n(1), p_n(2)] = [0.6, 0.4]$.

with the residual energy E. In the extreme case of $E = \infty$, the secondary user should always choose to sense.

VI. Conclusion

Within the framework of POMDP, we incorporated the secondary user's traffic statistics into the energy-constrained OSA design. By exploiting the underlying structure of the problem, we developed monotonicity results on the optimal sensing and access policies. 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.

Appendix A: Proof of Proposition 1

The belief update given in (17) reveals that the evolution of the belief vector is a Markov process. Hence, the POMDP given in (20) can be viewed as a fully observable Markov decision process (MDP) with uncountable states $\{\mathbf{\Lambda}(t), B(t), E(t)\}$.

The fact that the secondary user has a finite amount of initial energy and consumes at least $\min\{e_p, e_s\}$ energy in each slot implies that the evolution of the system state $\{\mathbf{\Lambda}(t), B(t), E(t)\}$ terminates in a finite but random time. The inevitable termination makes the POMDP given in (20) an instance of stochastic shortest path MDP problem, which has stationary optimal policies as shown in [13].

APPENDIX B: PROOF OF PROPOSITION 2

When $N = 1$, the belief vector $\mathbf{\Lambda} = [\Lambda_0, \Lambda_1]$ can be characterized by a scalar Λ_1 since $\Lambda_0 = 1 - \Lambda_1$. Let $Q(\Lambda_1, B, E|A)$ denote the maximum expected reward that can be obtained if the secondary user chooses sensing action A given that the current information state is (Λ_1, B, E). For the single-buffer case $l = 1$, the Q-function can be obtained as

$$B = 1:$$

$$Q(\Lambda_1, 1, E|0) = V(\mathcal{T}(\Lambda_1), 1, E - e_p), \tag{28a}$$

$$Q(\Lambda_1, 1, E|1) = (1 - \mathcal{T}(\Lambda_1))V(0, 1, E - e_s) + \mathcal{T}(\Lambda_1)Y(E), \tag{28b}$$

$$B = 0:$$

$$Q(\Lambda_1, 0, E|0) = q_0 V(\mathcal{T}(\Lambda_1), 0, E - e_p) + (1 - q_0)V(\mathcal{T}(\Lambda_1), 1, E - e_p), \tag{28c}$$

$$\begin{aligned} Q(\Lambda_1, 0, E|1) = &(1 - \mathcal{T}(\Lambda_1))[q_0 V(0, 0, E - e_s) + (1 - q_0)V(0, 1, E - e_s)], \\ &+ \mathcal{T}(\Lambda_1)[q_0 V(1, 0, E - e_s) + (1 - q_0)V(1, 1, E - e_s)], \end{aligned} \tag{28d}$$

where q_0 is the probability that no packet arrives in the current slot (see (3)), $\mathcal{T}(\Lambda_1) \stackrel{\Delta}{=} (1 - \Lambda_1)P_S(1|0) + \Lambda_1 P_S(1|1)$ is the probability that the channel is available given the current belief vector Λ_1, and $Y(E)$ is the maximum expected reward that can be obtained if the channel is sensed to be available in the current slot:

$$\begin{aligned} Y(E) \stackrel{\Delta}{=} \sum_{k=1}^{L} p(k) \max\{&V(1, 1, E - e_s); \\ &1_{[E \geq e_s + \varepsilon_k]}[1 + q_0 V(1, 0, E - e_s - \varepsilon_k) + (1 - q_0)V(1, 1, E - e_s - \varepsilon_k)]\}. \end{aligned} \tag{29}$$

Note that $Y(E)$ is independent of the belief vector. The value function and the optimal sensing action can be obtained in terms of the Q-functions, respectively, as

$$V(\Lambda_1, B, E) = \max_{A \in \{0,1\}} Q(\Lambda_1, B, E|A), \tag{30}$$

and

$$A^*(\Lambda_1, B, E) = \arg \max_{A \in \{0,1\}} Q(\Lambda_1, B, E|A). \tag{31}$$

Lemma 1: Given current buffer state $B = 1$ and residual energy E, we have that for any belief vector Λ_1,

$$Y(E) \geq V(\Lambda_1, 1, E - e_p), \tag{32}$$

where $Y(E)$ is defined in (29).

Proof: We first show by induction that for any belief vector Λ_1,

$$Y(E) \geq V(\Lambda_1, 1, E - e_s). \tag{33}$$

When residual energy $E < e_s + \varepsilon_1$, the above inequality (33) holds since both sides are equal to 0. Suppose that this inequality holds for all residual energies $E < \tilde{E}$. Then, applying (28a) and (28b) to (30), we obtain that for any belief vector Λ_1,

$$\begin{aligned} &V(\Lambda_1, 1, \tilde{E} - e_s) \\ &= \max\{V(\mathcal{T}(\Lambda_1), 1, \tilde{E} - e_s - e_p); (1 - \mathcal{T}(\Lambda_1))V(0, 1, \tilde{E} - 2e_s) + \mathcal{T}(\Lambda_1)Y(\tilde{E} - e_s)\} \\ &\leq \max\{Y(\tilde{E} - e_p); (1 - \mathcal{T}(\Lambda_1))Y(\tilde{E} - e_s) + \mathcal{T}(\Lambda_1)Y(\tilde{E} - e_s)\} \\ &= \max\{Y(\tilde{E} - e_p); Y(\tilde{E} - e_s)\} \leq Y(\tilde{E}), \end{aligned} \tag{34}$$

where the last inequality follows since $Y(E)$ given in (29) is monotonically non-decreasing, which follows from the fact the value function increases with the residual energy. This completes the proof of (33).

To show Lemma 1, we again use induction. Specifically, when $E < e_s + \varepsilon_1$, the inequality (32) holds since both sides are equal to 0. Suppose that (32) holds for all residual energies $E < \tilde{E}$. Then, applying (28a) and (28b) to (30), we obtain that for any belief vector Λ_1,

$$\begin{aligned} &V(\Lambda_1, 1, \tilde{E} - e_p) \\ &= \max\{V(\mathcal{T}(\Lambda_1), 1, \tilde{E} - 2e_p); (1 - \mathcal{T}(\Lambda_1))V(0, 1, \tilde{E} - e_s - e_p) + \mathcal{T}(\Lambda_1)Y(\tilde{E} - e_p)\} \\ &\leq \max\{Y(\tilde{E} - e_p); (1 - \mathcal{T}(\Lambda_1))Y(\tilde{E} - e_p) + \mathcal{T}(\Lambda_1)Y(\tilde{E} - e_p)\} \\ &\leq Y(\tilde{E} - e_p) \leq Y(\tilde{E}). \end{aligned} \tag{35}$$

Note that $V(0, 1, \tilde{E} - e_s - e_p) \leq Y(\tilde{E} - e_p)$ follows from (33). We point out that the proof of (33) is not necessary in the case of $e_s \geq e_p$ since $V(0, 1, \tilde{E} - e_s - e_p) \leq Y(\tilde{E} - e_s) \leq Y(\tilde{E} - e_p)$ follows from directly from the monotonicity of $Y(E)$ in this case. □

Next, we prove Proposition 2 based on Lemma 1 and the convexity of the value function in the belief vector. Suppose that the optimal sensing action $A^*(\Lambda_1, 1, E) = 1$ given current belief vector Λ_1, *i.e.*, $Q(\Lambda_1, 1, E|1) \geq Q(\Lambda_1, 1, E|0)$. Consider any belief vector $\tilde{\Lambda}_1$ such that $\mathcal{T}(\tilde{\Lambda}_1) \geq \mathcal{T}(\Lambda_1)$, where $\mathcal{T}(\Lambda_1) = (1 - \Lambda_1)P_S(1|0) + \Lambda_1 P_S(1|1)$. Since $Y(E)$ defined in (29) is independent of the belief vector, we obtain from (28b) that

$$\begin{aligned} Q(\tilde{\Lambda}_1,1,E|1) &= (1-\mathcal{T}(\tilde{\Lambda}_1))V(0,1,E-e_s)+\mathcal{T}(\tilde{\Lambda}_1)Y(E) \\ &= (1-\mathcal{T}(\tilde{\Lambda}_1))\frac{Q(\Lambda_1,1,E|1)-\Lambda_1 Y(E)}{1-\mathcal{T}(\Lambda_1)}+\mathcal{T}(\tilde{\Lambda}_1)Y(E) \\ &= \frac{1-\mathcal{T}(\tilde{\Lambda}_1)}{1-\mathcal{T}(\Lambda_1)}Q(\Lambda_1,1,E|1)+\frac{\mathcal{T}(\tilde{\Lambda}_1)-\mathcal{T}(\Lambda_1)}{1-\mathcal{T}(\Lambda_1)}Y(E) \\ &\geq \frac{1-\mathcal{T}(\tilde{\Lambda}_1)}{1-\mathcal{T}(\Lambda_1)}Q(\Lambda_1,1,E|0)+\frac{\mathcal{T}(\tilde{\Lambda}_1)-\mathcal{T}(\Lambda_1)}{1-\mathcal{T}(\Lambda_1)}Y(E). \end{aligned} \tag{36}$$

Applying (28a) and Lemma 1 to (36), we obtain that

$$Q(\tilde{\Lambda}_1,1,E|1) \geq \frac{1-\mathcal{T}(\tilde{\Lambda}_1)}{1-\mathcal{T}(\Lambda_1)}V(\mathcal{T}(\Lambda_1),1,E-e_p)+\frac{\mathcal{T}(\tilde{\Lambda}_1)-\mathcal{T}(\Lambda_1)}{1-\mathcal{T}(\Lambda_1)}V(\mathcal{T}(\Lambda_1),1,E-e_p). \tag{37}$$

Since the value function V is convex in belief vector, we obtain from (37) that

$$\begin{aligned} Q(\tilde{\Lambda}_1,1,E|1) &\geq V\left(\frac{1-\mathcal{T}(\tilde{\Lambda}_1)}{1-\mathcal{T}(\Lambda_1)}\mathcal{T}(\Lambda_1)+\frac{\mathcal{T}(\tilde{\Lambda}_1)-\mathcal{T}(\Lambda_1)}{1-\mathcal{T}(\Lambda_1)},1,E-e_p\right) \\ &= V(\mathcal{T}(\tilde{\Lambda}_1),1,E-e_p)=Q(\tilde{\Lambda}_1,1,E|0). \end{aligned} \tag{38}$$

Hence, the optimal sensing action is $A^*(\tilde{\Lambda}_1,1,E)=1$. This completes the proof of Proposition 2.

APPENDIX C: PROOF OF PROPOSITION 3

Let $G^{\phi}(\mathbf{\Lambda},B,E|A,k)$ denote the maximum expected reward that can be obtained given current information state $(\mathbf{\Lambda},B,E)$, sensing action $A>0$, sensing outcome $\Theta=k\in\{1,\ldots,L\}$, and access decision $\Phi_k=\phi\in\{0,1\}$. Then, the optimal access decision Φ_k^* under sensing outcome $\Theta=k$ is given by

$$\Phi_k^*=\max_{\phi\in\{0,1\}} G^{\phi}(\mathbf{\Lambda},B,E|A,k), \quad k=1,\ldots,L. \tag{39}$$

We obtain $G^{\phi}(\mathbf{\Lambda},B,E|A,k)$ in terms of the value function V as

$$G^0(\mathbf{\Lambda},B,E|A,k)=\sum_{\mathbf{s},\mathbf{s}'\in\mathcal{S}}\Lambda_{\mathbf{s}'}P_{\mathbf{S}}(\mathbf{s}|\mathbf{s}')\sum_{B'\in\mathcal{B}}P_B^0(B'|B)V(\mathbf{\Lambda}',B',E-e_s), \tag{40a}$$

$$G^1(\mathbf{\Lambda},B,E|A,k)= 1_{[E\geq e_s+\varepsilon_k]}\sum_{\mathbf{s},\mathbf{s}'\in\mathcal{S}}\Lambda_{\mathbf{s}'}P_{\mathbf{S}}(\mathbf{s}|\mathbf{s}')\left[1_{[s_A=1]}+\sum_{B'\in\mathcal{B}}P_B^1(B'|B)V(\mathbf{\Lambda}',B',E-e_s-\varepsilon_k)\right], \tag{40b}$$

where $\mathbf{\Lambda}'\triangleq\mathcal{T}_\Lambda(\mathbf{\Lambda}|A,k)$ denotes the updated belief vector as given in (17), which is the same for all sensing outcomes $k\in\{1,\ldots,L\}$. From (40a), we find that $G^0(\mathbf{\Lambda},B,E|A,k)$ is identical for all $k \in \{1, \ldots, L\}$. To this end, it suffices to show that $G^1(\mathbf{\Lambda}, B, E|A, k)$ given in (40b) is monotonically non-increasing on sensing outcomes $k \in \{1, \ldots, L\}$ since the optimal access decision is given by (39).

As the sensing outcome k increases, the required energy consumption ε_k as defined in (5) increases and hence $E - e_s - \varepsilon_k$ decreases. Hence, the indicator function $1_{[E \geq e_s + \varepsilon_k]}$ is non-increasing on sensing outcomes k. Furthermore, since the value function V is non-decreasing in residual energies, the future reward $V(\mathbf{\Lambda}', B', E - e_s - \varepsilon_k)$ is also non-increasing on sensing outcomes. Therefore, $G^1(\mathbf{\Lambda}, B, E|A, k)$ given in (40b) is non-increasing on sensing outcomes, and Proposition 3 follows.

When $N = 1$, the optimal access decision $\Phi_k^* = 1$ iff

$$\mathbb{E}_{B' \sim P_B^0}[V(1, B', E - e_s)] < 1_{[E \geq e_s + \varepsilon_k]}[1 + \mathbb{E}_{B' \sim P_B^1} V(1, B', E - e_s - \varepsilon_k)], \tag{41}$$

where B is the current buffer state, E is the residual energy, and $\mathbb{E}_{B' \sim P_B^i}$ denotes the expectation taken over the random variable B' according to the buffer transition probabilities $P_B^i(B'|B)$ given in (13). Clearly, the condition given in (41) and hence the optimal access decision are independent of the current belief vector.

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, and A. Swami, "Decentralized cognitive MAC for dynamic spectrum access," in *Proceedings of the First IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 224 - 232, Nov. 2005.

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

[5] 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.

[6] 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.

[7] Q. Zhao, S. Geirhofer, L. Tong, and B. M. Sadler, "Optimal dynamic spectrum access via periodic channel sensing," to appear in *Proceedings of WCNC*, 2007.

[8] 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.*

[9] Q. Zhao and B. Sadler, "A survey of dynamic spectrum access: signal processing, networking, and regulatory policy," to appear in *IEEE Signal Processing Magazine*, May 2007; available at http://www.ece.ucdavis.edu/∼qzhao/Journal.html.

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

[11] 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

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

[13] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 1. Belmont, Massachusetts: Athena Scientific, 1995.