Joint Design for Opportunistic Spectrum Access with Multi-Channel Sensing

Yunxia Chen, Qing Zhao, and Ananthram Swami

I. Problem Statement

A. Network Model

Consider a spectrum of N orthogonal channels licensed to a slotted primary network. Let $S_n(t) \in \{0 \text{ (busy)}, 1 \text{ (idle)}\}$ denote the occupancy of channel n in slot t. We assume that the spectrum occupancy $\mathbf{S}(t) \stackrel{\Delta}{=} [S_1(t), \ldots, S_N(t)]$ follows a discrete-time homogeneous Markov process with finite state space $\mathbb{S} \stackrel{\Delta}{=} \{0,1\}^N$ and known transition probabilities $\{P_{\mathbf{s},\mathbf{s}'}\}_{\mathbf{s},\mathbf{s}' \in \mathbb{S}}$, where $P_{\mathbf{s},\mathbf{s}'} \stackrel{\Delta}{=} \Pr\{\mathbf{S}(t+1) = \mathbf{s}' \mid \mathbf{S}(t) = \mathbf{s}\}$ is the probability that the spectrum occupancy state transits from $\mathbf{s} \in \mathbb{S}$ to $\mathbf{s}' \in \mathbb{S}$. We assume that the transition probabilities remain unchanged for T slots.

We consider a secondary ad hoc network whose users independently and selfishly search for and exploit instantaneous spectrum opportunities in these N channels. Specifically, at the beginning of slot t, a secondary user with data to transmit

1) chooses a set $\mathcal{A}(t) \in \mathbb{A}_s^{(L)} \stackrel{\Delta}{=} \{\mathcal{A} : \mathcal{A} \subset \{1, \ldots, N\}, |\mathcal{A}| = L\}$ of channels to sense;
2) chooses an operating point $\mathcal{E}(t) \in \mathbb{A}_\delta^{(L)}(\mathcal{A}(t))$ for the spectrum sensor[1], which is used to detect the occupancy states $\mathbf{S}_\mathcal{A}(t) = \{S_n(t)\}_{n \in \mathcal{A}(t)}$ of the chosen channels $\mathcal{A}(t)$;
3) based on the sensing outcomes $\mathbf{\Theta}_\mathcal{A}(t) \stackrel{\Delta}{=} \{\Theta_n(t)\}_{n \in \mathcal{A}(t)} \in \{0 \text{ (busy)}, 1 \text{ (idle)}\}^L$, decides whether to access each sensed channel $\mathbf{\Phi}_\mathcal{A}(t) \stackrel{\Delta}{=} \{\Phi_n(t)\}_{n \in \mathcal{A}(t)} \in \{0 \text{ (no access)}, 1 \text{ (access)}\}^L$.

At the end of the slot, every successful transmission is acknowledged: $\mathbf{K}_\mathcal{A}(t) \stackrel{\Delta}{=} \{K_n(t)\}_{n \in \mathcal{A}(t)} \in \{0 \text{ (not successful)}, 1 \text{ (successful)}\}^L$. We assume that the acknowledgement is error-free and that secondary users do not collide with each other. That is, acknowledgement $K_n(t) = 1$ is received if and only if the secondary user accesses an idle channel, *i.e.*, $K_n(t) = S_n(t)\Phi_n(t)$. Our goal is to design an OSA strategy for secondary users, which sequentially specifies which channels in the spectrum to sense, which spectrum sensor to use, and which sensed channels to access.

[1] See [2] for a detailed description of the spectrum sensor.

B. A Constrained POMDP Formulation

As detailed in [1], [2], the optimal design of OSA can be formulated as a constrained partially observable Markov decision process (POMDP).

Belief Vector The secondary user's knowledge of the spectrum occupancy based on all past actions and observations can thus be summarized by a belief vector:

$$\mathbf{\Lambda}(t) \triangleq \{\lambda_{\mathbf{s}}(t)\}_{\mathbf{s}\in\mathbb{S}} \in \Pi(\mathbb{S}) \triangleq \left\{\{\lambda_{\mathbf{s}}\}_{\mathbf{s}\in\mathbb{S}} : \lambda_{\mathbf{s}} \in [0,1], \sum_{\mathbf{s}\in\mathbb{S}} \lambda_{\mathbf{s}} = 1\right\}, \tag{1}$$

where $\lambda_{\mathbf{s}}(t)$ is the conditional probability (given the decision and observation history) that the spectrum occupancy is in state s at the beginning of the slot t prior to the state transition. Hence, based on the belief vector $\mathbf{\Lambda}(t)$ at the beginning of slot t, we obtain the *a priori* distribution of current spectrum occupancy $\mathbf{S}(t)$ as

$$\Pr\{\mathbf{S}(t) = \mathbf{s}\} = \sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}}, \qquad \forall \mathbf{s} \in \mathbb{S}, \tag{2}$$

which can be in the design of spectrum sensor.

Policy Within the POMDP framework, a sensing policy π_s that sequentially determines which channels to sense is given by a sequence of functions:

$$\pi_s = [\mu_1, \ldots, \mu_T], \quad \text{ where } \mu_t : \Pi(\mathbb{S}) \to \mathbb{A}_s^{(L)}, \tag{3}$$

where μ_t maps the belief vector $\mathbf{\Lambda}(t) \in \Pi(\mathbb{S})$ at the beginning of slot t to a set $\mathcal{A}(t) \in \mathbb{A}_s^{(L)}$ of channel to be sensed in slot t. Similarly, a sensor operating policy π_δ that specifies, in each slot t, an achievable sensor operating point $\mathcal{E}(t) \in \mathbb{A}_\delta^{(L)}$ is given by a sequence of functions:

$$\pi_\delta = [\omega_1, \ldots, \omega_T], \quad \text{ where } \omega_t : \Pi(\mathbb{S}) \to \mathbb{A}_\delta^{(L)}(\mathcal{A}(t)), \tag{4}$$

where $\mathcal{A}(t) = \mu_t(\mathbf{\Lambda}(t))$ is the set of channels chosen by the sensing policy in slot t.

As shown in [2], randomized access policies are required to achieve optimality of this constrained POMDP. A randomized access policy specifies, in each slot t, a set of transmission probabilities $\mathcal{F}(t) \triangleq \{f_n(\boldsymbol{\theta}_{\mathcal{A}}, t)\}_{n\in\mathcal{A}(t), \boldsymbol{\theta}_{\mathcal{A}}\in\{0,1\}^L}$, where

$$f_n(\boldsymbol{\theta}_{\mathcal{A}}, t) \triangleq \Pr\{\Phi_n(t) = 1 \,|\, \mathbf{\Theta}_{\mathcal{A}}(t) = \boldsymbol{\theta}_{\mathcal{A}}\}$$

is the probability of accessing sensed channel $n \in \mathcal{A}(t)$ when sensing outcomes are given by $\boldsymbol{\theta}_{\mathcal{A}} \in \{0,1\}^L$. For notation convenience, we let $\mathbb{A}_c^{(L)} \triangleq [0,1]^{L2^L}$ include all possible sets of transmission probabilities $\mathcal{F}(t)$. Hence, a randomized access policy is given by

$$\pi_c = [\nu_1, \ldots, \nu_T], \qquad \text{where } \nu_t : \Pi(\mathbb{S}) \rightarrow \mathbb{A}_c^{(L)}. \tag{5}$$

Objective Our goal is to develop the optimal OSA strategy by jointly optimizing the sensor operating policy π_δ and the spectrum sensing/access policies $\{\pi_s, \pi_c\}$. The objective is to maximize the total expected number of information bits that can be delivered by the secondary user (*i.e.*, the total expected reward of the POMDP) in T slots under the constraint that the probability of collision perceived by the primary network in any channel and any slot is capped below ζ, *i.e.*,

$$\begin{aligned} \{\pi_\delta^*, \pi_s^*, \pi_c^*\} &= \arg\max_{\pi_\delta, \pi_s, \pi_c} \mathbb{E}_{\{\pi_\delta, \pi_s, \pi_c\}} \left[\sum_{t=1}^{T} R(t) \middle| \mathbf{\Lambda}(1) \right] \\ \text{s.t. } P_n(t) &\triangleq \Pr\{\Phi_n(t) = 1 \,|\, S_n(t) = 0\} \leq \zeta, \quad \forall t, n, \end{aligned} \tag{6}$$

where $\mathbb{E}_{\{\pi_\delta, \pi_s, \pi_c\}}$ is the expectation given sensor operating policy π_δ and sensing and access policies $\{\pi_s, \pi_c\}$, $P_n(t)$ is the probability of collision perceived by the primary network in channel n in slot t, and $\mathbf{\Lambda}(1)$ is the initial belief vector given by the stationary distribution of the spectrum occupancy. Note that when $\Pr\{S_n(t) = 0\} = 0$, no collision will occur and the optimal access decision is straightforward $\Phi_n(t) = 1$.

C. Optimality Equation

The value function $V_t(\mathbf{\Lambda}(t))$, defined as the maximum expected remaining reward that can be obtained starting from slot t ($1 \leq t \leq T$) given that the current belief vector is $\mathbf{\Lambda}(t) \in \Pi(\mathbb{S})$, is the key to solving the POMDP. It can be shown that the value function is the unique solution to the following optimality equation:

$$\begin{aligned} V_t(\mathbf{\Lambda}(t)) = \max_{A = \{\mathcal{A}, \mathcal{E}, \mathcal{F}\} \in \mathbb{A}^{(L)}} \sum_{\mathbf{s} \in \mathbb{S}} \sum_{\mathbf{s}' \in \mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}', \mathbf{s}} \sum_{\mathbf{k}_\mathcal{A} \in \{0,1\}^L} U_{\mathbf{s}, \mathbf{k}_\mathcal{A}}^{(L)}(A) \\ \times \left[R_{\mathbf{k}_\mathcal{A}}^{(\mathcal{A})} + V_{t+1}(\mathcal{T}(\mathbf{\Lambda}(t) \,|\, A, \mathbf{k}_\mathcal{A})) \right], \quad 1 \leq t < T, \end{aligned} \tag{7a}$$

$$V_T(\mathbf{\Lambda}(T)) = \max_{A = \{\mathcal{A}, \mathcal{E}, \mathcal{F}\} \in \mathbb{A}^{(L)}} \sum_{\mathbf{s} \in \mathbb{S}} \sum_{\mathbf{s}' \in \mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}', \mathbf{s}} \sum_{\mathbf{k}_\mathcal{A} \in \{0,1\}^L} U_{\mathbf{s}, \mathbf{k}_\mathcal{A}}^{(L)}(A) R_{\mathbf{k}_\mathcal{A}}^{(\mathcal{A})}, \tag{7b}$$

$$\text{s.t. } P_n(t) = \sum_{\boldsymbol{\theta}_\mathcal{A}, \mathbf{s}_\mathcal{A} \in \{0,1\}^L} h_{\mathbf{S}_\mathcal{A} | S_n}(\mathbf{s}_\mathcal{A} \,|\, 0) \, l_{\mathbf{\Theta}_\mathcal{A} | \mathbf{S}_\mathcal{A}}(\boldsymbol{\theta}_\mathcal{A} \,|\, \mathbf{s}_\mathcal{A}) \, f_n(\boldsymbol{\theta}_\mathcal{A}) \leq \zeta, \quad \forall n, t, \tag{7c}$$

where $\mathbb{A}^{(L)} \stackrel{\Delta}{=} \{\{\mathcal{A},\mathcal{E},\mathcal{F}\} : \mathcal{A} \in, \mathcal{E} \in \mathbb{A}_\delta^{(L)}(\mathcal{A}), \mathcal{F} \in \mathbb{A}_c^{(L)}\}$ include all possible decisions in a slot, and

$$U_{\mathbf{s},\mathbf{k}_\mathcal{A}}^{(L)}(A) \stackrel{\Delta}{=} \Pr\{\mathbf{K}_\mathcal{A} = \mathbf{k}_\mathcal{A} \,|\, \mathbf{S} = \mathbf{s}\}, \tag{8a}$$

$$h_{\mathbf{S}_\mathcal{A}|S_n}(\mathbf{s}_\mathcal{A} \,|\, i) \stackrel{\Delta}{=} \Pr\{\mathbf{S}_\mathcal{A} = \mathbf{s}_\mathcal{A} \,|\, S_n = i\}, \tag{8b}$$

$$l_{\mathbf{\Theta}_\mathcal{A}|\mathbf{S}_\mathcal{A}}(\boldsymbol{\theta}_\mathcal{A} \,|\, \mathbf{s}_\mathcal{A}) \stackrel{\Delta}{=} \Pr\{\mathbf{\Theta}_\mathcal{A} = \boldsymbol{\theta}_\mathcal{A} \,|\, \mathbf{S}_\mathcal{A} = \mathbf{s}_\mathcal{A}\}, \tag{8c}$$

$$\mathcal{T}(\mathbf{\Lambda}(t) \,|\, A, \mathbf{k}_\mathcal{A}) \stackrel{\Delta}{=} \mathbf{\Lambda}(t+1), \quad \text{where } \lambda_\mathbf{s}(t+1) = \frac{\sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} U_{\mathbf{s},\mathbf{k}_\mathcal{A}}^{(L)}(A)}{\sum_{\mathbf{s}\in\mathbb{S}} \sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} U_{\mathbf{s},\mathbf{k}_\mathcal{A}}^{(L)}(A)}. \tag{8d}$$

Specifically, $h_{\mathbf{S}_\mathcal{A}|S_n}(\mathbf{s}_\mathcal{A} \,|\, i)$ is the conditional distribution of channel occupancies $\mathbf{S}_\mathcal{A} \stackrel{\Delta}{=} \{S_n\}_{n\in\mathcal{A}}$, which can be calculated via (2):

$$h_{\mathbf{S}_\mathcal{A}|S_n}(\mathbf{s}_\mathcal{A} \,|\, i) = \frac{\sum_{\bar{\mathbf{s}}\in\mathbb{S}} \Pr(\mathbf{S} = \bar{\mathbf{s}}) 1_{\bar{\mathbf{s}}_\mathcal{A} = \mathbf{s}_\mathcal{A}, \bar{s}_n = i}}{\sum_{\bar{\mathbf{s}}\in\mathbb{S}} \Pr(\mathbf{S} = \bar{\mathbf{s}}) 1_{\bar{s}_n = i}}. \tag{9}$$

Clearly, $h_{\mathbf{S}_\mathcal{A}|S_n}(\mathbf{s}_\mathcal{A} \,|\, i) = 0$ if $s_n \neq i$. $l_{\mathbf{\Theta}_\mathcal{A}|\mathbf{S}_\mathcal{A}}(\boldsymbol{\theta}_\mathcal{A} \,|\, \mathbf{s}_\mathcal{A})$ is an error probability of the spectrum sensor, which is determined by the current operating point $\mathcal{E}$. $U_{\mathbf{s},\mathbf{k}_\mathcal{A}}^{(L)}(A)$ is the probability of observing $\mathbf{K}_\mathcal{A} = \mathbf{k}_\mathcal{A}$ given action $A = \{\mathcal{A},\mathcal{E},\mathcal{F}\} \in \mathbb{A}^{(L)}$ and current spectrum occupancy $\mathbf{S} = \mathbf{s}$, which can be derived as

$$\begin{aligned} U_{\mathbf{s},\mathbf{k}_\mathcal{A}}^{(L)}(A) &= \sum_{\boldsymbol{\theta}_\mathcal{A}\in\{0,1\}^L} \Pr\{\mathbf{\Theta}_\mathcal{A} = \boldsymbol{\theta}_\mathcal{A} \,|\, \mathbf{S} = \mathbf{s}\} \Pr\{\mathbf{K}_\mathcal{A} = \mathbf{k}_\mathcal{A} \,|\, \mathbf{\Theta}_\mathcal{A} = \boldsymbol{\theta}_\mathcal{A}, \mathbf{S} = \mathbf{s}\} \\ &= \sum_{\boldsymbol{\theta}_\mathcal{A}\in\{0,1\}^L} l_{\mathbf{\Theta}_\mathcal{A}|\mathbf{S}_\mathcal{A}}(\boldsymbol{\theta}_\mathcal{A} \,|\, \mathbf{s}_\mathcal{A}) \prod_{n\in\mathcal{A}} \Pr\{K_n = k_n \,|\, \mathbf{\Theta}_\mathcal{A} = \boldsymbol{\theta}_\mathcal{A}, \mathbf{S}_\mathcal{A} = \mathbf{s}_\mathcal{A}\} \\ &= \sum_{\boldsymbol{\theta}_\mathcal{A}\in\{0,1\}^L} l_{\mathbf{\Theta}_\mathcal{A}|\mathbf{S}_\mathcal{A}}(\boldsymbol{\theta}_\mathcal{A} \,|\, \mathbf{s}_\mathcal{A}) \prod_{n\in\mathcal{A}} [k_n s_n f_n(\boldsymbol{\theta}_\mathcal{A}) + (1-k_n)(1 - s_n f_n(\boldsymbol{\theta}_\mathcal{A}))]. \end{aligned} \tag{10}$$

In principle, the optimal decision $\{\mathcal{A}^*,\mathcal{E}^*,\mathcal{F}^*\}$ in each slot can be obtained by solving (7) recursively. However, without any structural results on this constrained POMDP, (7) is computationally prohibitive. A nature question here is when there exists a separation principle similar to [1] that can be used for the optimal design of OSA with multi-channel sensing.

II. Separation Principle

Theorem 1: **The Separation Principle**

When the spectrum sensor and the access strategy are designed independently across channels, the optimal joint design of OSA can be carried out in two steps:

Step 1: *Choose the sensor operating points and the transmission probabilities to maximize the expected immediate reward subject to collision constraint.*

Step 2: *Choose the sensing strategy to maximize the expected total reward.*

In this case, the optimal spectrum sensor should detect the occupancy of a chosen channel by applying the measurement of this channel to the optimal NP detector with probability of miss detection equal to ζ, and the optimal access decision is to trust the sensing outcome.

Proof: See Appendix A. □□□

We call the spectrum sensor and the access policy given in Theorem 1 the II approach, indicating the conditions of independent spectrum opportunity identification and independent access decision-making. We have shown in Theorem 1 that the II approach is optimal under these conditions. In Proposition 1, we will further show that this II approach is locally optimal (*i.e.*, maximizes the instantaneous throughput) under certain relaxed conditions.

Proposition 1: When the spectrum sensor is designed independently across channels, the II approach given in 1 is locally optimal if the spectrum occupancy evolves independently across channels.

Proof: See Appendix B. □□□

III. Heuristic Approaches

We propose two heuristic approaches to exploiting the channel correlation: the PHY layer and the MAC layer approaches.

1) The PHY Layer Approach: When the occupancy states are correlated across channels, we have correlated channel measurements at the PHY layer. Hence, the measurements of all chosen channels should be jointly exploited in spectrum opportunity identification. With this in mind, we propose a heuristic design of the spectrum sensor: it detects the occupancy of a channel by applying the measurements of all chosen channels to the optimal NP detector with probability of miss detection equal to ζ. Different from the II approach, this heuristic approach incorporates all channel measurements into the occupancy detection of each individual channel. Hence, the structures of the optimal NP detector used in the heuristic approach and the II approach are different, leading to different error performance of the spectrum sensor.

By combining the above heuristic design of spectrum sensor with the access strategy of the II approach, we have the JI approach that performs joint spectrum opportunity identification and independent access decision-making. Proposition (2) provides a sufficient condition under which this JI approach is locally optimal.

Proposition 2: When the access policy is designed independently across channels, the JI approach is locally optimal and it reduces to the II approach if the spectrum occupancy evolves independently across channels.

Proof: See Appendix C. □□□

2) The MAC Layer Approach: When channel occupancies are correlated, so are the sensing outcomes given by the spectrum sensor. Hence, the channel correlation can also be exploited at the MAC layer by making access decisions jointly across channels. A heuristic approach is to adopt the myopic access policy that maximizes the instantaneous throughput. Specifically, for given chosen channels $\mathcal{A} \in \mathbb{A}_s$ and belief vector $\mathbf{\Lambda}(t) \in \Pi(\mathbb{S})$ in slot t, we choose transmission probabilities $\hat{\mathcal{F}} = \{f_n(\boldsymbol{\theta}_\mathcal{A})\}_{\substack{n\in\mathcal{A}\\ \boldsymbol{\theta}_\mathcal{A}\in\{0,1\}^L}} \in [0,1]^{L2^L}$ as follows

$$\begin{aligned}
\hat{\mathcal{F}} &= \arg\max_{\mathcal{F}\in\mathbb{A}_c^{(L)}} \mathbb{E}\left[R_{\mathbf{K}_\mathcal{A}}^{(\mathcal{A})} \,\middle|\, \mathbf{\Lambda}(t)\right] \\
&= \arg\max_{\mathcal{F}\in\mathbb{A}_c^{(L)}} \sum_{n\in\mathcal{A}} B_n \Pr\{K_n = 1\} = \arg\max_{\mathcal{F}\in\mathbb{A}_c^{(L)}} \sum_{n\in\mathcal{A}} B_n \Pr\{\Phi_n S_n = 1\} \\
&= \arg\max_{\mathcal{F}\in\mathbb{A}_c^{(L)}} \sum_{n\in\mathcal{A}} B_n \Pr\{S_n = 1\} \sum_{\boldsymbol{\theta}_\mathcal{A},\mathbf{s}_\mathcal{A}\in\{0,1\}^L} h_{\mathbf{S}_\mathcal{A}|S_n}(\mathbf{s}_\mathcal{A} \,|\, 1)\, l_{\mathbf{\Theta}_\mathcal{A}|\mathbf{S}_\mathcal{A}}(\boldsymbol{\theta}_\mathcal{A} \,|\, \mathbf{s}_\mathcal{A})\, f_n(\boldsymbol{\theta}_\mathcal{A}) \qquad \text{(11a)}
\end{aligned}$$

$$\text{s.t.} \quad P_n(t) = \sum_{\boldsymbol{\theta}_\mathcal{A},\mathbf{s}_\mathcal{A}\in\{0,1\}^L} h_{\mathbf{S}_\mathcal{A}|S_n}(\mathbf{s}_\mathcal{A} \,|\, 0)\, l_{\mathbf{\Theta}_\mathcal{A}|\mathbf{S}_\mathcal{A}}(\boldsymbol{\theta}_\mathcal{A} \,|\, \mathbf{s}_\mathcal{A})\, f_n(\boldsymbol{\theta}_\mathcal{A}) \leq \zeta, \quad \forall n, t, \qquad \text{(11b)}$$

where the conditional probability $h_{\mathbf{S}_\mathcal{A}|S_n}(\mathbf{s}_\mathcal{A} \,|\, i)$ $(i = 0, 1)$ of the current channel occupancies $\mathbf{S}_\mathcal{A}$ and the sensing error probability $l_{\mathbf{\Theta}_\mathcal{A}|\mathbf{S}_\mathcal{A}}(\boldsymbol{\theta}_\mathcal{A} \,|\, \mathbf{s}_\mathcal{A})$ are defined in (8).

Combining this myopic access policy with the spectrum sensor of the II approach, we have the IJ approach that performs independent spectrum opportunity identification and joint access decision-making. In this case, the myopic access policy given in (11) can be obtained via linear programming. Proposition 3 shows that this IJ approach is equivalent to the II approach when the spectrum occupancy evolves independently across channels.

Proposition 3: When the spectrum sensor is designed independently across channels, the IJ approach reduces to the II approach, and it is locally optimal if the spectrum occupancy evolves independently across channels.

Proof: See Appendix B. □□□

APPENDIX A: PROOF OF THEOREM 1

When the spectrum sensor and the access policy are designed independently across channels, the secondary user chooses, in each slot, a set of channels $\mathcal{A} \in \mathbb{A}_s^{(L)}$ and then determines for each chosen channel $n \in \mathcal{A}$ a sensor operating point $(\epsilon_n, \delta_n) \in \mathbb{A}_\delta^{(1)}(n)$ and a pair of transmission probabilities $(f_n(0), f_n(1)) \in \mathbb{A}_c^{(1)} = [0,1]^2$. Let $A^{(L)} \triangleq \{\mathcal{A}, \{(\epsilon_n, \delta_n)\}_{n\in\mathcal{A}}, \{(f_n(0), f_n(1))\}_{n\in\mathcal{A}}\}$ and $A_n \triangleq \{n, (\epsilon_n, \delta_n), (f_n(0), f_n(1))\} \in \mathbb{A}$, where A_n corresponds to the action taken on chosen channel $n \in \mathcal{A}$.

When the spectrum sensor is designed independently across channels, we can write $l_{\boldsymbol{\Theta}_\mathcal{A}|\mathbf{S}_\mathcal{A}}(\boldsymbol{\theta}_\mathcal{A} \,|\, \mathbf{s}_\mathcal{A}) = \Pr\{\boldsymbol{\Theta}_\mathcal{A} = \boldsymbol{\theta}_\mathcal{A} \,|\, \mathbf{S}_\mathcal{A} = \mathbf{s}_\mathcal{A}\} = \prod_{n\in\mathcal{A}} \Pr\{\Theta_n = \theta_n \,|\, S_n = s_n\}$ in a product form since the occupancy detection of a channel is independent of the measurement of other chosen channels. When the access policy is designed independently across channels, we have $f_n(\boldsymbol{\theta}_\mathcal{A}) = f_n(\theta_n)$ for all sensing outcomes $\boldsymbol{\theta}_\mathcal{A} \in \{0,1\}^L$. Therefore, we can write the observation probability $U^{(L)}_{\mathbf{s},\mathbf{k}_\mathcal{A}}(A^{(L)})$ given action $A^{(L)}$ as (10)

$$
\begin{aligned}
U^{(L)}_{\mathbf{s},\mathbf{k}_\mathcal{A}}(A^{(L)}) &= \sum_{\boldsymbol{\theta}_\mathcal{A}\in\{0,1\}^L} \prod_{n\in\mathcal{A}} \Pr\{\Theta_n = \theta_n \,|\, S_n = s_n\}[k_n s_n f_n(\theta_n) + (1-k_n)(1-s_n f_n(\theta_n))] \\
&= \prod_{n\in\mathcal{A}} \sum_{\theta_n=0}^{1} \Pr\{\Theta_n = \theta_n \,|\, S_n = s_n\}[k_n s_n f_n(\theta_n) + (1-k_n)(1-s_n f_n(\theta_n))] \\
&= \prod_{n\in\mathcal{A}} U_{\mathbf{s},k_n}(A_n), \qquad (12)
\end{aligned}
$$

where $U_{\mathbf{s},k_n}(A_n) = \Pr\{K_n = k_n | \mathbf{S} = \mathbf{s}\}$ is the observation probability in the design of OSA with single-channel sensing (see [3]). Similarly, after some algebras, the design constraint in (7c) can be written as

$$
\begin{aligned}
P_n(t) = \sum_{\mathbf{s}_\mathcal{A}\in\{0,1\}^L} \sum_{\boldsymbol{\theta}_{\bar{\mathcal{A}}_n}\in\{0,1\}^{L-1}} & \Pr\{\mathbf{S}_\mathcal{A} = \mathbf{s}_\mathcal{A} \,|\, S_n = 0\} \Pr\{\boldsymbol{\Theta}_{\bar{\mathcal{A}}_n} = \boldsymbol{\theta}_{\bar{\mathcal{A}}_n} \,|\, \mathbf{S}_{\bar{\mathcal{A}}_n} = \mathbf{s}_{\bar{\mathcal{A}}_n}\} \\
& \times \sum_{\theta_n=0}^{1} \Pr\{\Theta_n = \theta_n \,|\, S_n = 0\} f_n(\theta_n) \qquad (13) \\
= \sum_{\theta_n=0}^{1} \Pr\{\Theta_n = \theta_n \,|\, S_n = 0\} f_n(\theta_n) &= (1-\delta_n) f_n(0) + \delta_n f_n(1) \leq \zeta, \quad \forall n \in \mathcal{A}.
\end{aligned}
$$

Substituting (12) into (7), we can see that the sensor operating point (ϵ_n, δ_n) and transmission probabilities $(f_n(0), f_n(1))$ of a chosen channel $n \in \mathcal{A}$ affect the maximum remaining reward only through $U_{\mathbf{s},1}(A_n) = s_n[\epsilon_n f_n(0) + (1-\epsilon_n) f_n(1)]$, which is independent of the actions

$\{A_m\}_{m\in\mathcal{A}\backslash\{n\}}$ taken on other channels. Moreover, the simplified constraint (13) reveals that the collision probability of a channel n is also independent of the actions $\{A_m\}_{m\in\mathcal{A}\backslash\{n\}}$ taken on other channels. Therefore, the design of the spectrum sensor and access policy can be decoupled across channels. Following the same proof as given in [3], we can show that the expected remaining reward increases with $\epsilon_n f_n(0)+(1-\epsilon_n)f_n(1)$ for each channel $n\in\mathcal{A}$.

On the other hand, the expected immediate reward $\mathbb{E}[R^{\mathcal{A}}_{\mathbf{K}_{\mathcal{A}}}\,|\,\mathbf{\Lambda}(t)]$ is given by

$$\mathbb{E}[R^{\mathcal{A}(t)}_{\mathbf{K}_{\mathcal{A}(t)}}\,|\,\mathbf{\Lambda}(t)]=\sum_{n\in\mathcal{A}}B_n\Pr\{K_n=1\}=\sum_{n\in\mathcal{A}}B_n\Pr\{S_n=1\}[\epsilon_n f_n(0)+(1-\epsilon_n)f_n(1)], \quad (14)$$

which also increases with each $\epsilon_n f_n(0)+(1-\epsilon_n)f_n(1)$. Therefore, the separation principle developed in [1] can be used for the optimal design of OSA with multi-channel $L>1$ sensing.

APPENDIX B: PROOF OF PROPOSITIONS 1 AND 3

To show Propositions 1 and 3, we derive the myopic spectrum sensor and the myopic access policy below. Let $\mathcal{A}\in\mathbb{A}_s^{(L)}$ denote a set of chosen channels and $\bar{\mathcal{A}}_n=\mathcal{A}\backslash\{n\}$ be all chosen channels except for n. Since spectrum opportunities are identified independently across channels, we have

$$h_{\mathbf{S}_{\mathcal{A}}|S_n}(\mathbf{s}_{\mathcal{A}}\,|\,0)=1_{[s_n=0]}h_{\mathbf{S}_{\bar{\mathcal{A}}_n}|S_n}(\mathbf{s}_{\bar{\mathcal{A}}_n}\,|\,0). \quad (15)$$

Since spectrum opportunities evolve independently across channels, we have

$$h_{\mathbf{S}_{\bar{\mathcal{A}}_n}|S_n}(\mathbf{s}_{\bar{\mathcal{A}}_n}\,|\,0)=h_{\mathbf{S}_{\bar{\mathcal{A}}_n}|S_n}(\mathbf{s}_{\bar{\mathcal{A}}_n}\,|\,1)=\Pr\{\mathbf{S}_{\bar{\mathcal{A}}_n}=\mathbf{s}_{\bar{\mathcal{A}}_n}\}. \quad (16)$$

Therefore, given belief vector $\mathbf{\Lambda}(t)$ and chosen channels $\mathcal{A}$ in slot t, the myopic (*i.e.*, locally optimal) sensor operating point $(\hat{\epsilon}_n,\hat{\delta}_n)$ and transmission probabilities $\hat{\mathcal{F}}=\{\hat{f}_n(\boldsymbol{\theta}_{\mathcal{A}})\}$ are given by (11)

$$\{(\hat{\epsilon}_n, \hat{\delta}_n), \hat{\mathcal{F}}\} = \arg \max_{\substack{(\epsilon_n,\delta_n)\in\mathbb{A}_\delta^{(1)} \\ \mathcal{F}\in\mathbb{A}_c^{(L)}}} \mathbb{E}\left[R_{\mathbf{K}_\mathcal{A}}^{(\mathcal{A})} \mid \mathbf{\Lambda}(t)\right]$$

$$= \arg \max_{\substack{(\epsilon_n,\delta_n)\in\mathbb{A}_\delta^{(1)} \\ \mathcal{F}\in\mathbb{A}_c^{(L)}}} \sum_{n\in\mathcal{A}} B_n \Pr\{S_n = 1\} \sum_{\theta_n=0}^{1} \Pr\{\Theta_n = \theta_n \mid S_n = 1\} g_n(\theta_n)$$

$$= \arg \max_{\substack{(\epsilon_n,\delta_n)\in\mathbb{A}_\delta^{(1)} \\ \mathcal{F}\in\mathbb{A}_c^{(L)}}} \sum_{n\in\mathcal{A}} B_n \Pr\{S_n = 1\}[\epsilon_n g_n(0) + (1-\epsilon_n) g_n(1)] \tag{17a}$$

$$\text{s.t.} \quad P_n(t) = \sum_{\theta_n=0}^{1} \Pr\{\Theta_n = \theta_n \mid S_n = 0\} g_n(\theta_n) = (1-\delta_n) g_n(0) + \delta_n g_n(1) \leq \zeta, \quad \forall n \in \mathcal{A}, \tag{17b}$$

where $g(\theta_n) \in [0,1]$ is defined as

$$g_n(\theta_n) \stackrel{\Delta}{=} \sum_{\boldsymbol{\theta}_{\bar{\mathcal{A}}_n}\in\{0,1\}^{L-1}} f_n(\boldsymbol{\theta}_{\bar{\mathcal{A}}_n}, \theta_n) \sum_{\mathbf{s}_{\bar{\mathcal{A}}_n}\in\{0,1\}^{L-1}} \Pr\{\mathbf{S}_{\bar{\mathcal{A}}_n} = \mathbf{s}_{\bar{\mathcal{A}}_n}\} \prod_{m\in\bar{\mathcal{A}}_n} \Pr\{\Theta_m = \theta_m \mid S_m = s_m\}. \tag{18}$$

We see from (17) that the myopic approach should maximize $\epsilon_n g_n(0) + (1-\epsilon_n) g_n(1)$ under the constraint $(1-\delta_n) g_n(0) + \delta_n g_n(1) \leq \zeta$ for every chosen channel $n \in \mathcal{A}$, leading to the same optimization problem as [3, (11)]. Hence, $\hat{\delta}_n = \zeta$ and $(\hat{g}_n(0), \hat{g}_n(1)) = (0,1)$ are the solution to (17). That is, the II sensor is locally optimal. Furthermore, since $(\hat{g}_n(0), \hat{g}_n(1)) = (0,1)$ is achieved by choosing $\hat{f}_n(\boldsymbol{\theta}_{\bar{\mathcal{A}}_n}, \theta_n) = 1_{[\theta_n=1]}$ in (18), transmission probabilities $\hat{f}_n(\boldsymbol{\theta}_\mathcal{A}) = \theta_n$ are locally optimal, which completes the proof of Proposition 1. Proposition 3 follows directly from the fact that the IJ approach employs the myopic access policy and the II sensor, which has been proven to be locally optimal.

Appendix C: Proof of Proposition 2

When the access policy is designed independently across channels, we have $f_n(\boldsymbol{\theta}_\mathcal{A}) = f_n(\theta_n)$ for any sensing outcome $\mathbf{\Theta}_\mathcal{A} = \boldsymbol{\theta}_\mathcal{A}$ from the chosen channels $\mathcal{A}$. Hence, given belief vector $\mathbf{\Lambda}(t)$ and chosen channels $\mathcal{A}$ in slot t, the myopic spectrum sensor $\hat{\mathcal{E}}$ and access decision

$\{(\hat{f}_n(0), \hat{f}_n(1))\}_{n\in\mathcal{A}}$ are given by

$$\{\hat{\mathcal{E}}, \{(\hat{f}_n(0), \hat{f}_n(1))\}_{n\in\mathcal{A}}\} = \arg \max_{\substack{\mathcal{E}\in\mathbb{A}_\delta^{(L)} \\ f_n(0), f_n(1)\in[0,1]}} \sum_{n\in\mathcal{A}} B_n \Pr\{S_n = 1\}[\Pr\{\Theta_n = 1 \,|\, S_n = 1\} \times f_n(1) + \Pr\{\Theta_n = 0 \,|\, S_n = 1\} f_n(0)] \tag{19a}$$

$$\text{s.t.} \quad P_n(t) = \Pr\{\Theta_n = 1 \,|\, S_n = 0\} f_n(1) + \Pr\{\Theta_n = 0 \,|\, S_n = 0\} f_n(0) \leq \zeta, \quad \forall n \in \mathcal{A}, \tag{19b}$$

where $\Pr\{\Theta_n = \theta_n \,|\, S_n = s_n\}$ is determined by the sensor operating point $\mathcal{E} \in \mathbb{A}_\delta^{(L)}$:

$$\Pr\{\Theta_n = \theta_n \,|\, S_n = s_n\} = \sum_{\boldsymbol{\theta}_{\bar{\mathcal{A}}_n}, \mathbf{s}_{\bar{\mathcal{A}}_n} \in\{0,1\}^{L-1}} \Pr\{\boldsymbol{\Theta}_{\bar{\mathcal{A}}_n} = \boldsymbol{\theta}_{\bar{\mathcal{A}}_n}, \Theta_n = \theta_n \,|\mathbf{S}_{\bar{\mathcal{A}}_n} = \mathbf{s}_{\bar{\mathcal{A}}_n}, S_n = s_n\} \tag{20}$$

Since (19) has the same form as [3, (11)], $\Pr\{\Theta_n = 1 \,|\, S_n = 0\} = \zeta$ and $(\hat{f}_n(0), \hat{f}_n(1)) = (0, 1)$ are the solution to (19). Hence, the JI approach is locally optimal.

Furthermore, when spectrum opportunities evolve independently across channels, the measurements of different channels are independent. Hence, the JI sensor is equivalent to the II sensor.

References

[1] Y. Chen, Q. Zhao, and A. Swami, “Joint design and separation principle for opportunistic spectrum access,” in *Proc. of IEEE Asilomar Conference on Signals, Communications, and Networking*, Oct. 2006.

[2] Y. Chen, Q. Zhao, and A. Swami, “Joint PHY-MAC design for opportunistic spectrum access in the presence of sensing errors,” *To be submitted to IEEE Transactions on Signal Processing*, Jan. 2007.

[3] Y. Chen, Q. Zhao, and A. Swami, “Proof of the separation principle for opportunistic spectrum access,” tech. rep., University of California, Davis, Nov. 2006. http://www.ece.ucdavis.edu/q̃zhao/Report.html