Proof of the Separation Principle for Opportunistic Spectrum Access

Yunxia Chen, Qing Zhao, and Ananthram Swami

I. Problem Statement

A. Network Model

Consider a spectrum that consists of N channels (*e.g.*, separated frequency bands), each with bandwidth B_n ($n = 1, \cdots, N$). These N channels are licensed to a primary network whose users communicate according to a synchronous slot structure. Let $S_n \in \{0 \text{ (busy)}, 1 \text{ (idle)}\}$ denote the availability of channel n in a slot. The spectrum occupancy of the primary network $\mathbf{S} \stackrel{\Delta}{=} [S_1, \ldots, S_N]$ is modeled by a discrete-time Markov process with 2^N states. Let $\mathbb{S} \stackrel{\Delta}{=} \{0, 1\}^N$ denote the state space and $P_{\mathbf{s},\mathbf{s}'}$ the probability that the spectrum occupancy state transits from $\mathbf{s} \in \mathbb{S}$ to $\mathbf{s}' \in \mathbb{S}$.

We consider a secondary ad hoc network whose users independently and selfishly exploit instantaneous spectrum opportunities in these N channels. At the beginning of each slot, a secondary user with data to transmit chooses a channel $a \in \{1, \ldots, N\}$ to sense. Based on the sensing outcome Θ_a, the secondary user decides whether to access the sensed channel $\Phi_a \in \{0 \text{ (no access)}, 1 \text{ (access)}\}$ according to transmission probability $f_a(\theta) \stackrel{\Delta}{=} \Pr\{\Phi_a = 1 \,|\, \Theta_a = \theta\} \in [0, 1]$. At the end of the slot, the receiver acknowledges every successful transmission: $K_a \in \{0 \text{ (no ACK)}, 1 \text{ (ACK)}\}$. If the secondary user successfully accesses an idle channel, it will receive a reward measured by the number of transmitted information bits. A collision occurs if the secondary user accesses a busy channel.

B. Spectrum Sensor

The spectrum sensor of the secondary user performs a binary hypothesis test:

$$\begin{aligned} &\mathcal{H}_0: \text{ the sensed channel is idle}, \\ \text{vs.} \quad &\mathcal{H}_1: \text{ the sensed channel is busy}. \end{aligned} \tag{1}$$

The sensing outcome (*i.e.*, the hypothesis test result) is denoted by $\Theta_a \in \{0 \text{ (busy)}, 1 \text{ (idle)}\}$.

If the sensor mistakes $\mathcal{H}_0$ for $\mathcal{H}_1$ in a channel, a false alarm occurs, and a spectrum opportunity is overlooked by the sensor. On the other hand, if the sensor mistakes $\mathcal{H}_1$ for $\mathcal{H}_0$, we have a miss detection. The probabilities of false alarm and miss detection in channel a are denoted, respectively, by

$$\epsilon_a \triangleq \Pr\{\Theta_a = 0 \,|\, S_a = 1\}, \tag{2a}$$

$$\delta_a \triangleq \Pr\{\Theta_a = 1 \,|\, S_a = 0\}. \tag{2b}$$

The performance of the spectrum sensor is specified by the receiver operating characteristics (ROC) curve which gives the probability of detection $1 - \delta_a$ as a function of ϵ_a. We assume here that the ROC curve of spectrum sensors has already been obtained, and we focus on the tradeoff between false alarm and miss detection, *i.e.*, which point on the given ROC curve should spectrum sensor operate at.

C. Sensing and Access Policies

The sensing strategy decides, in each slot, which channel to sense, and the access strategy determines the transmission probabilities for different sensing outcomes. As shown in [1], [2], the optimal OSA design can be formulated as a partially observable Markov decision process (POMDP):

- State space: $\mathbb{S} = \{0, 1\}^N$.
- Action space: $\mathbb{A} = \mathbb{A}_s \times \mathbb{A}_\delta \times \mathbb{A}_c$, where $a \in \mathbb{A}_s = \{1, \ldots, N\}$ (the sensing action space), $\delta_a \in \mathbb{A}_\delta = [0, 1]$ (the sensor operating space), and $\{f_a(0), f_a(1)\} \in \mathbb{A}_c = [0, 1]^2$ (the transmission probability space.
- Observation space: $\mathbb{K} = \{0, 1\}$. Note that to maintain transceiver synchronization, the transmitter and the receiver must have the same observation in each slot. Since sensing outcomes may have errors, we use acknowledgement K_a as their common observation. See details in [2].
- Observation model: when action $(a, \delta, \{f(0), f(1)\})$ is taken, the probability of observing

$K_a = k \in \{0,1\}$ under the current spectrum occupancy state $\mathbf{S} = \mathbf{s}$ is given by

$$U_{\mathbf{s},1}(a,\delta,f(0),f(1)) = \sum_{\theta=0}^{1} \Pr\{\Theta_a = \theta \,|\, \mathbf{S} = \mathbf{s}\} 1_{[s_a=1]} f(\theta)$$
$$= 1_{[s_a=1]}[\epsilon f(0) + (1-\epsilon) f(1)], \tag{3a}$$
$$U_{\mathbf{s},0}(a,\delta,f(0),f(1)) = 1 - U_{\mathbf{s},1}(a,\delta,f(0),f(1)), \tag{3b}$$

where ϵ is the probability of false alarm corresponding to sensor operating point δ.

- Transition probabilities: $\{P_{\mathbf{s},\mathbf{s}'}\}$ are determined by the primary traffic and thus independent of the secondary user's actions.
- Reward: $R^a_{K_a} = K_a B_a$. Note that we have assumed that the number of bits that can be delivered over a channel is proportional to the channel bandwidth.

Due to partial spectrum monitoring and sensing errors, secondary users cannot directly observe the current spectrum occupancy state. Their knowledge of the current spectrum occupancy state based on all past actions and observations can be summarized by a belief vector $\mathbf{\Lambda}(t) \triangleq \{\lambda_{\mathbf{s}}(t)\}_{\mathbf{s}\in\mathbb{S}} \in [0,1]^{|\mathbb{S}|}$ [3], where $\lambda_{\mathbf{s}}(t)$ is the conditional probability (given the decision and observation history) that the spectrum occupancy state is given by $\mathbf{S} = \mathbf{s}$ at the beginning of the slot t prior to the state transition. After action $(a, \delta_a, \{f(0), f(1)\}) \in \mathbb{A}$ is taken and observation $K_a = k$ is made, the belief vector can be updated by

$$\mathbf{\Lambda}(t+1) = \mathcal{T}(\mathbf{\Lambda}(t) \,|\, a, \delta, \{f(0), f(1)\}, k)$$
$$\lambda_{\mathbf{s}}(t+1) = \frac{\sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} U_{\mathbf{s},k}(a,\delta,f(0),f(1))}{\sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} U_{\mathbf{s},k}(a,\delta,f(0),f(1))}. \tag{4}$$

For ease of presentation, we define

$$p(k|a,\delta,f(0),f(1),\mathbf{\Lambda}(t)) \triangleq \sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} U_{\mathbf{s},k}(a,\delta,f(0),f(1)), \tag{5}$$

which is the probability of observing $K_a = k$ when the belief vector is given by $\mathbf{\Lambda}(t)$ and the action $(a, \delta, \{f(0), f(1)\})$ is taken. Note that $p(k|a,\delta,f(0),f(1),\mathbf{\Lambda}(t))$ is obtained by averaging the conditional observation probability $U_{\mathbf{s},k}(a,\delta,f(0),f(1))$ over the current spectrum occupancy state.

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

$$\pi_s = [\mu_1, \ldots, \mu_T], \qquad \text{where } \mu_t : [0,1]^{|\mathbb{S}|} \rightarrow \mathbb{A}_s, \tag{6}$$

where μ_t maps the belief vector $\mathbf{\Lambda}(t) \in [0,1]^{|\mathbb{S}|}$ at the beginning of slot t to a channel $a \in \mathbb{A}_s$ to be sensed in slot t. Similarly, a sensor operating policy π_δ that specifies the sensor operating point δ on the ROC curve and an access policy π_c that determines the transmission probabilities under different sensing outcomes are given, respectively, by sequences of functions:

$$\pi_\delta = [\omega_1, \ldots, \omega_T], \qquad \text{where } \omega_t : [0,1]^{|\mathbb{S}|} \rightarrow \mathbb{A}_\delta, \tag{7}$$

$$\pi_c = [\nu_1, \ldots, \nu_T], \qquad \text{where } \nu_t : [0,1]^{|\mathbb{S}|} \rightarrow \mathbb{A}_c. \tag{8}$$

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 below ζ, *i.e.*,

$$\begin{aligned} &\{\pi_\delta^*, \pi_s^*, \pi_c^*\} = \arg\max_{\pi_\delta, \pi_s, \pi_c} \sum_{t=1}^{T} \mathbb{E}_{\{\pi_\delta, \pi_s, \pi_c\}} \left[R(t) | \, \mathbf{\Lambda}(1) \right] \\ &\text{s.t. } P_a(t) \triangleq \Pr\{\Phi_a = 1 \,|\, S_a = 0\} \leq \zeta, \quad \forall t, \end{aligned} \tag{9}$$

where $\mathbb{E}_{\{\pi_\delta, \pi_s, \pi_c\}}$ is the expectation given sensor operating policy π_δ and sensing and access policies $\{\pi_s, \pi_c\}$, $P_a(t)$ is the probability of collision perceived by the primary network in chosen channel a in slot t, and $\mathbf{\Lambda}(1)$ is the initial belief vector which can be the stationary distribution of spectrum occupancy. Note that in the trivial case where $\Pr\{S_a = 0\} = 0$, the optimal sensing and access decisions are straightforward: choose this channel to access; the sensor operating policy is irrelevant. In this paper, we consider the non-trivial case where $\Pr\{S_n = 0\} \neq 0$ in any slot and thus the constraint in (9) is well-defined.

II. SEPARATION PRINCIPLE

To solve (9), we derive the value function $V_t(\mathbf{\Lambda}(t))$ defined as the maximum expected remaining reward that can be obtained from slot t:

$$V_t(\mathbf{\Lambda}(t)) = \max_{a\in\{1\ldots,N\}} \max_{\delta\in[0,1]} \max_{\{f(0),f(1)\}\in[0,1]^2} \sum_{k=0}^{1} p(k|a,\delta,f(0),f(1),\mathbf{\Lambda}(t)) \times [kB_a + V_{t+1}(\mathcal{T}(\mathbf{\Lambda}(t)\,|\,a,\delta,\{f(0),f(1)\},k))], \quad 1\le t<T, \tag{10a}$$

$$V_T(\mathbf{\Lambda}(T)) = \max_{a\in\{1,\ldots,N\}} \max_{\delta\in[0,1]} \max_{\{f(0),f(1)\}\in[0,1]^2} p(1|a,\delta,f(0),f(1),\mathbf{\Lambda}(t))B_n, \tag{10b}$$

$$\text{s.t.} \quad P_a(t) = \sum_{\theta=0}^{1} \Pr\{\Theta_a=\theta\,|\,S_a=0\}\Pr\{\Phi_a=1\,|\,\Theta_a=\theta\} = (1-\delta)f(0)+\delta f(1)\le\zeta, \tag{10c}$$

where $p(k|a,\delta,f(0),f(1),\mathbf{\Lambda}(t))$ is given in (5).

Theorem 1: **The Separation Principle**

The joint OSA design can be carried out in two steps without losing optimality. First, choose sensor operating policy π_δ^* and spectrum access policy π_c^* to maximize the instant throughput (i.e., the expected immediate reward) under the design constraint. Hence, given any belief vector $\mathbf{\Lambda}(t)$ at the beginning of any slot t, the optimal sensor operating point δ_a^* and the optimal transmission probabilities $\{f_a^*(0), f_a^*(1)\}$ of chosen channel a are given by

$$\{\delta_a^*, f_a^*(0), f_a^*(1)\} = \arg \max_{\substack{\delta_a\in[0,1]\\ \{f_a(0),f_a(1)\}\in[0,1]^2}} \mathbb{E}_{K_a}[R_{K_a}^a(t)\,|\,\mathbf{\Lambda}(t)] = \arg \max_{\substack{\delta_a\in[0,1]\\ \{f_a(0),f_a(1)\}\in[0,1]^2}} \epsilon_a f_a(0) + (1-\epsilon_a)f_a(1) \tag{11a}$$

$$\textit{s.t.} \quad (1-\delta_a)f_a(0)+\delta_a f_a(1)\le\zeta. \tag{11b}$$

Second, choose spectrum sensing policy to maximize the overall throughput (i.e., total expected reward). That is, the optimal sensing policy π_s^* can be obtained by solving an unconstrained POMDP:

$$\pi_s^* = \arg\max_{\pi_s} \mathbb{E}_{\{\pi_s,\pi_\delta^*,\pi_c^*\}} \left[\sum_{t=1}^{T} R_{K_a}^a(t) \middle| \mathbf{\Lambda}(1) \right], \tag{12}$$

where π_δ^* and π_c^* are obtained from (11a).

Proof: The second part of the separation principle follows directly from the first part since the design constraint has been ensured by (11) for any sensing action a. The proof of first part is built upon the following three Lemmas.

Lemma 1: The value function given in (10) *is convex in the belief state. Specifically, at any time t, the value functions $V_t(\mathbf{\Lambda}_1)$ and $V_t(\mathbf{\Lambda}_2)$ of any two belief vectors $\mathbf{\Lambda}_1 \in [0,1]^{|\mathbb{S}|}$ and $\mathbf{\Lambda}_2 \in [0,1]^{|\mathbb{S}|}$ satisfy*

$$V_t(\tau\mathbf{\Lambda}_1 + (1-\tau)\mathbf{\Lambda}_2) \leq \tau V_t(\mathbf{\Lambda}_1) + (1-\tau)V_t(\mathbf{\Lambda}_2), \quad \forall 0 \leq \tau \leq 1. \tag{13}$$

Proof: We use mathematical induction to prove Lemma 1. From (10b), we can see that the value function $V_T(\mathbf{\Lambda})$ at slot $t = T$ is linear and hence convex in the belief vector $\mathbf{\Lambda}$. Now suppose that $V_t(\mathbf{\Lambda})$ is convex for every $t > t_0$. Let $Q_{t_0}(\mathbf{\Lambda} \,|\, a, \delta, f(0), f(1))$ denote the maximum expected remaining reward if action $(a, \delta, f(0), f(1)) \in \mathbb{A}$ is taken in slot t_0:

$$\begin{aligned} &Q_{t_0}(\mathbf{\Lambda} \,|\, a, \delta, f(0), f(1)) \\ &\quad = \sum_{k=0}^{1} p(k|a, \delta, f(0), f(1), \mathbf{\Lambda})[kB_a + V_{t_0+1}(\mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta, \{f(0), f(1)\}, k))]. \end{aligned} \tag{14}$$

If we can show that $Q_{t_0}(\mathbf{\Lambda} \,|\, a, \delta, f(0), f(1))$ is convex in belief vector $\mathbf{\Lambda}$ for any action, then it will proves the convexity of the value function $V_{t_0}(\mathbf{\Lambda}) = \max_{a,\delta,\{f(0),f(1)\}} Q_{t_0}(\mathbf{\Lambda} \,|\, a, \delta, f(0), f(1))$ since the maximum of a set of convex functions is convex.

To show that $Q_{t_0}(\mathbf{\Lambda} \,|\, a, \delta, f(0), f(1))$ is convex in belief vector $\mathbf{\Lambda}$, we consider any two belief vectors $\mathbf{\Lambda}_1 = \{\lambda_{1,\mathbf{s}}\}_{\mathbf{s}\in\mathbb{S}} \in [0,1]^{|\mathbb{S}|}$, $\mathbf{\Lambda}_2 = \{\lambda_{2,\mathbf{s}}\}_{\mathbf{s}\in\mathbb{S}} \in [0,1]^{|\mathbb{S}|}$ and any $\tau \in [0,1]$. Applying (4) yields

$$\begin{aligned} &\overline{\mathbf{\Lambda}} \triangleq \mathcal{T}(\tau\mathbf{\Lambda}_1 + (1-\tau)\mathbf{\Lambda}_2 \,|\, a, \delta, \{f(0), f(1)\}, k) \\ &\overline{\lambda}_{\mathbf{s}} = \frac{\sum_{\mathbf{s}'\in\mathbb{S}}[\tau\lambda_{1,\mathbf{s}'} + (1-\tau)\lambda_{2,\mathbf{s}'}]P_{\mathbf{s}',\mathbf{s}}U_{\mathbf{s},k}(a, \delta, f(0), f(1))}{p(k \,|\, a, \delta, f(0), f(1), \tau\mathbf{\Lambda}_1 + (1-\tau)\mathbf{\Lambda}_2)} \\ &\quad = \tau' \frac{\sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{1,\mathbf{s}'}P_{\mathbf{s}',\mathbf{s}}U_{\mathbf{s},k}(a, \delta, f(0), f(1))}{p(k \,|\, a, \delta, f(0), f(1), \mathbf{\Lambda}_1)} + (1-\tau')\frac{\sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{2,\mathbf{s}'}P_{\mathbf{s}',\mathbf{s}}U_{\mathbf{s},k}(a, \delta, f(0), f(1))}{p(k \,|\, a, \delta, f(0), f(1), \mathbf{\Lambda}_2)}, \end{aligned} \tag{15}$$

where

$$\tau' = \frac{\tau p(k \,|\, a, \delta, f(0), f(1), \mathbf{\Lambda}_1)}{p(k \,|\, a, \delta, f(0), f(1), \tau\mathbf{\Lambda}_1 + (1-\tau)\mathbf{\Lambda}_2)} \in [0,1]. \tag{16}$$

Note that by definition (5)

$$\begin{aligned} &p(k \,|\, a, \delta, f(0), f(1), \tau\mathbf{\Lambda}_1 + (1-\tau)\mathbf{\Lambda}_2) \\ &\quad = \tau p(k \,|\, a, \delta, f(0), f(1), \mathbf{\Lambda}_1) + (1-\tau)p(k \,|\, a, \delta, f(0), f(1), \mathbf{\Lambda}_2). \end{aligned} \tag{17}$$

Hence, we can write

$$\begin{aligned}&\mathcal{T}(\tau\mathbf{\Lambda}_1+(1-\tau)\mathbf{\Lambda}_2\,|\,a,\delta,\{f(0),f(1)\},k)\\&\quad=\tau'\mathcal{T}(\mathbf{\Lambda}_1\,|\,a,\delta,\{f(0),f(1)\},k)+(1-\tau')\mathcal{T}(\mathbf{\Lambda}_2\,|\,a,\delta,\{f(0),f(1)\},k)\end{aligned} \tag{18}$$

Applying (17) and using (18) with the convexity of $V_{t_0+1}(\cdot)$ (see Lemma 1), we can obtain that

$$\begin{aligned}&Q_{t_0}(\tau\mathbf{\Lambda}_1+(1-\tau)\mathbf{\Lambda}_2\,|\,a,\delta,f(0),f(1))\\&=\sum_{k=0}^{1}p(k\,|\,a,\delta,f(0),f(1),\tau\mathbf{\Lambda}_1+(1-\tau)\mathbf{\Lambda}_2)\\&\quad\times[kB_a+V_{t_0+1}(\mathcal{T}(\tau\mathbf{\Lambda}_1+(1-\tau)\mathbf{\Lambda}_2\,|\,a,\delta,\{f(0),f(1)\},k))]\\&\le\sum_{k=0}^{1}[\tau p(k\,|\,a,\delta,f(0),f(1),\mathbf{\Lambda}_1)+(1-\tau)p(k\,|\,a,\delta,f(0),f(1),\mathbf{\Lambda}_2)]kB_a\\&\quad+\sum_{k=0}^{1}p(k\,|\,a,\delta,f(0),f(1),\tau\mathbf{\Lambda}_1+(1-\tau)\mathbf{\Lambda}_2)[\tau'V_{t_0+1}(\mathcal{T}(\mathbf{\Lambda}_1\,|\,a,\delta,\{f(0),f(1)\},k))\\&\quad+(1-\tau')V_{t_0+1}(\mathcal{T}(\mathbf{\Lambda}_1\,|\,a,\delta,\{f(0),f(1)\},k))]\\&=\tau\sum_{k=0}^{1}p(k\,|\,a,\delta,f(0),f(1),\mathbf{\Lambda}_1)[kB_a+V_{t_0+1}(\mathcal{T}(\mathbf{\Lambda}_1\,|\,a,\delta,\{f(0),f(1)\},k))]\\&\quad+(1-\tau)\sum_{k=0}^{1}p(k\,|\,a,\delta,f(0),f(1),\mathbf{\Lambda}_2)[kB_a+V_{t_0+1}(\mathcal{T}(\mathbf{\Lambda}_2\,|\,a,\delta,\{f(0),f(1)\},k))]\\&=\tau Q_{t_0}(\mathbf{\Lambda}_1\,|\,a,\delta,f(0),f(1))+(1-\tau)Q_{t_0}(\mathbf{\Lambda}_2\,|\,a,\delta,f(0),f(1)).\end{aligned} \tag{19}$$

Hence, $Q_{t_0}(\mathbf{\Lambda}\,|\,a,\delta,f(0),f(1))$ is convex in belief vector $\mathbf{\Lambda}$. Lemma 1 follows. □

Lemma 2: If observation $K_a=1$ is made in slot t, the maximum expected future reward $V_{t+1}(\mathcal{T}(\mathbf{\Lambda}\,|\,a,\delta_a,\{f_a(0),f_a(1)\},1))$ is independent of the sensor operating point δ_a and the transmission probabilities $\{f_a(0),f_a(1)\}$ employed in the current slot. That is, for any δ_a, $\{f_a(0),f_a(1)\}$ and δ'_a, $\{f'_a(0),f_a(1)'\}$

$$V_{t+1}(\mathcal{T}(\mathbf{\Lambda}\,|\,a,\delta_a,\{f_a(0),f_a(1)\},1))=V_{t+1}(\mathcal{T}(\mathbf{\Lambda}\,|\,a,\delta'_a,\{f'_a(0),f'_a(1)\},1)). \tag{20}$$

Proof: Applying the observation probability $U_{\mathbf{s},1}(a,\delta_a,f_a(0),f_a(1))$ given in (3) to (4), we obtain the updated belief vector $\mathbf{\Lambda}_1(t+1)=\{\lambda_{1,\mathbf{s}}\}_{\mathbf{s}\in\mathbb{S}}\triangleq\mathcal{T}(\mathbf{\Lambda}\,|\,a,\delta_a,\{f_a(0),f_a(1)\},1)$ as

$$\lambda_{1,\mathbf{s}}(t+1)=\frac{\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}}{\sum_{\mathbf{s}'\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}} \tag{21}$$

which is independent of the sensor operating point δ_a and the transmission probabilities $\{f_a(0),f_a(1)\}$.

□

Lemma 3: If observation $K_a = 0$ is made in slot t, the maximum expected future reward $V_{t+1}(\mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta_a, \{f_a(0), f_a(1)\}, 0))$ satisfies the following inequality for any operating points $\delta_a, \delta'_a \in \mathbb{A}_\delta$, any transmission probabilities $\{f_a(0), f_a(1)\}, \{f'_a(0), f_a(1)'\} \in \mathbb{A}_c$:

$$\begin{aligned} V_{t+1}(\mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta_a, \{f_a(0), f_a(1)\}, 0)) \leq \tau V_{t+1}(\mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta_a, \{f_a(0), f_a(1)\}, 1)) \\ + (1-\tau) V_{t+1}(\mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta'_a, \{f'_a(0), f'_a(1)\}, 0)), \end{aligned} \tag{22}$$

where τ is given by

$$\tau = \frac{p(1 \,|\, a, \delta'_a, f'_a(0), f'_a(1), \mathbf{\Lambda}) - p(1 \,|\, a, \delta_a, f_a(0), f_a(1), \mathbf{\Lambda})}{p(0 \,|\, a, \delta_a, f_a(0), f_a(1), \mathbf{\Lambda})}. \tag{23}$$

Proof: Consider any sensing action $a \in \mathbb{A}_s$ in slot t. Given any operating points $\delta_a, \delta'_a \in \mathbb{A}_\delta$ and any transmission probabilities $\{f_a(0), f_a(1)\}, \{f'_a(0), f_a(1)'\} \in \mathbb{A}_c$, if we can demonstrate that the updated belief vectors satisfy

$$\begin{aligned} \mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta_a, \{f_a(0), f_a(1)\}, 0) = \tau \mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta_a, \{f_a(0), f_a(1)\}, 1) \\ + (1-\tau) \mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta'_a, \{f'_a(0), f'_a(1)\}, 0), \end{aligned} \tag{24}$$

where τ is given in (23), then since the value function is convex in belief vector (see Lemma 1), this will prove Lemma 3.

For ease of presentation, let $\mathbf{\Lambda}_0 = \mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta_a, \{f_a(0), f_a(1)\}, 0)$, $\mathbf{\Lambda}_1 = \mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta_a, \{f_a(0), f_a(1)\}, 1)$, and $\mathbf{\Lambda}'_0 = \mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta'_a, \{f'_a(0), f'_a(1)\}, 0)$. Applying (4) and noting that $\sum_{\mathbf{s}\in\mathbb{S}} \sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} = 1$, we obtain the difference between the elements of the updated belief vectors $\mathbf{\Lambda}_0$ and $\mathbf{\Lambda}'_0$:

$$\begin{aligned} &\lambda_{0,\mathbf{s}} - \lambda'_{0,\mathbf{s}} \\ &\quad= \frac{\sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} [1 - 1_{[s_a=1]} g]}{\sum_{\mathbf{s}\in\mathbb{S}} \sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} [1 - 1_{[s_a=1]} g]} - \frac{\sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} [1 - 1_{[s_a=1]} g']}{\sum_{\mathbf{s}\in\mathbb{S}} \sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} [1 - 1_{[s_a=1]} g']} \\ &\quad= \frac{\{\sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} - g \sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} 1_{[s_a=1]}\}\{1 - g' \sum_{\mathbf{s}\in\mathbb{S}} \sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} 1_{[s_a=1]}\}}{\{1 - g \sum_{\mathbf{s}\in\mathbb{S}} \sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} 1_{[s_a=1]}\}\{1 - g' \sum_{\mathbf{s}\in\mathbb{S}} \sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} 1_{[s_a=1]}\}} \\ &\qquad- \frac{\{\sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} - g' \sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} 1_{[s_a=1]}]\}\{1 - g \sum_{\mathbf{s}\in\mathbb{S}} \sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} 1_{[s_a=1]}\}}{\{1 - g \sum_{\mathbf{s}\in\mathbb{S}} \sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} 1_{[s_a=1]}\}\{1 - g' \sum_{\mathbf{s}\in\mathbb{S}} \sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} 1_{[s_a=1]}\}} \\ &\quad= \frac{(g' - g)\{\sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} 1_{[s_a=1]} - [\sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}}][\sum_{\mathbf{s}\in\mathbb{S}} \sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} 1_{[s_a=1]}]\}}{\{1 - g \sum_{\mathbf{s}\in\mathbb{S}} \sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} 1_{[s_a=1]}\}\{1 - g' \sum_{\mathbf{s}\in\mathbb{S}} \sum_{\mathbf{s}'\in\mathbb{S}} \lambda_{\mathbf{s}'}(t) P_{\mathbf{s}',\mathbf{s}} 1_{[s_a=1]}\}} \end{aligned} \tag{25}$$

where $g = \epsilon_a f_a(0) + (1-\epsilon_a) f_a(1)$ and $g' = \epsilon'_a f'_a(0) + (1-\epsilon'_a) f'_a(1)$ are constants, ϵ_a and ϵ'_a are associated with δ_a and δ'_a, respectively. Similarly, we obtain the difference between the elements of the updated belief vectors $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_0'$:

$$\begin{aligned}\lambda_{1,\mathbf{s}} - \lambda_{0,\mathbf{s}}' &\\ &= \frac{\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}}{\sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}} - \frac{\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}[1-1_{[s_a=1]}g']}{\sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}[1-1_{[s_a=1]}g']}\\ &= \frac{\{\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}\}\{1-g'\sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}\}}{\{\sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}\}\{1-g'\sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}\}}\\ &\quad - \frac{\{\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}} - g'\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}\}\{\sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}\}}{\{\sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}\}\{1-g'\sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}\}}\\ &= \frac{\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]} - [\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}][\sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}]}{\{\sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}\}\{1-g'\sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}\}}.\end{aligned} \tag{26}$$

Dividing $\lambda_{0,\mathbf{s}} - \lambda_{0,\mathbf{s}}'$ by $\lambda_{1,\mathbf{s}} - \lambda_{0,\mathbf{s}}'$ and applying (25) and (26), we obtain that

$$\begin{aligned}\frac{\lambda_{0,\mathbf{s}} - \lambda_{0,\mathbf{s}}'}{\lambda_{1,\mathbf{s}} - \lambda_{0,\mathbf{s}}'} &= \frac{(g'-g)\sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}}{1-g\sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}}\\ &= \frac{\sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}g' - \sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}g}{\sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}(t)P_{\mathbf{s}',\mathbf{s}}[1-1_{[s_a=1]}g]}\\ &= \frac{p(1\,|\,a,\delta_a',f_a'(0),f_a'(1),\mathbf{\Lambda}) - p(1\,|\,a,\delta_a,f_a(0),f_a(1),\mathbf{\Lambda})}{p(0\,|\,a,\delta_a,f_a(0),f_a(1),\mathbf{\Lambda})} = \tau \text{ in (23).}\end{aligned} \tag{27}$$

Since τ is independent of s, we can write

$$\frac{\mathbf{\Lambda}_0 - \mathbf{\Lambda}_0'}{\mathbf{\Lambda}_1 - \mathbf{\Lambda}_0'} = \tau. \tag{28}$$

which proves (24). This completes the proof of Lemma 3.

□□□

With the above Lemmas, we now prove the separation principle. Given any belief vector $\mathbf{\Lambda}$ at the beginning of slot t and any sensing action $a \in \mathbb{A}_s$ taken in slot t. First, note that the expected immediate reward $\mathbb{E}_K[R_{K_a}^a\,|\,\mathbf{\Lambda}]$ is given by

$$\mathbb{E}_{K_a}[R_{K_a}^a\,|\,\mathbf{\Lambda}] = B_a p(1\,|\,a,\delta,f(0),f(1)) = [\epsilon f(0) + (1-\epsilon)f(1)]B_a\sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}, \tag{29}$$

where ϵ is the probability of false alarm associated with δ and $B_a\sum_{\mathbf{s}\in\mathbb{S}}\sum_{\mathbf{s}'\in\mathbb{S}}\lambda_{\mathbf{s}'}P_{\mathbf{s}',\mathbf{s}}1_{[s_a=1]}$ is a constant for a given belief vector $\mathbf{\Lambda}$ and sensing action a. Second, we note that the sensor operating point δ and the transmission probabilities $(f(0), f(1))$ only affect the expected remaining reward $Q_t(\mathbf{\Lambda}\,|\,a,\delta,f(0),f(1))$ (14) through the term $\epsilon f(0) + (1-\epsilon)f(1)$ (see (14) together with (5), (4), and (3)). Therefore, if we can show that $Q_t(\mathbf{\Lambda}\,|\,a,\delta,f(0),f(1))$ increases with $\epsilon f(0) + (1-\epsilon)f(1)$, then this will prove the separation principle.

Let $\delta, \delta' \in \mathbb{A}_\delta$ be two possible sensor operating points and $\{f(0), f(1)\}, \{f'(0), f'(1)\} \in \mathbb{A}_c$ be two possible sets of transmission probabilities such that $\epsilon' f'(0) + (1-\epsilon')f'(1) \geq \epsilon f(0) + (1-\epsilon)f(1)$. Comparing the resulting expected remaining rewards $Q_t(\mathbf{\Lambda} \,|\, a, \delta', f'(0), f'(1))$ and $Q_t(\mathbf{\Lambda} \,|\, a, \delta, f(0), f(1))$ in slot t, we obtain that

$$\begin{aligned} &Q_t(\mathbf{\Lambda} \,|\, a, \delta', f'(0), f'(1)) - Q_t(\mathbf{\Lambda} \,|\, a, \delta, f(0), f(1)) \\ &\quad= \sum_{k=0}^{1} p(k|a, \delta', f'(0), f'(1), \mathbf{\Lambda})[kB_a + V_{t+1}(\mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta', \{f'(0), f'(1)\}, k))] \\ &\quad\quad - \sum_{k=0}^{1} p(k|a, \delta, f(0), f(1), \mathbf{\Lambda})[kB_a + V_{t+1}(\mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta, \{f(0), f(1)\}, k))] \\ &\quad= [p(1|a, \delta', f'(0), f'(1), \mathbf{\Lambda}) - p(1|a, \delta, f(0), f(1), \mathbf{\Lambda})]B_a \\ &\quad\quad + p(0|a, \delta', f'(0), f'(1), \mathbf{\Lambda})V_{t+1}(\mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta', \{f'(0), f'(1)\}, 0)) \\ &\quad\quad - p(0|a, \delta, f(0), f(1), \mathbf{\Lambda})V_{t+1}(\mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta', \{f(0), f(1)\}, 0)) \\ &\quad\quad + p(1|a, \delta', f'(0), f'(1), \mathbf{\Lambda})V_{t+1}(\mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta', \{f'(0), f'(1)\}, 1)) \\ &\quad\quad - p(1|a, \delta, f(0), f(1), \mathbf{\Lambda})V_{t+1}(\mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta, \{f(0), f(1)\}, 1)). \end{aligned} \tag{30}$$

Since we have assumed that $\epsilon' f'(0)+(1-\epsilon')f'(1) \geq \epsilon f(0)+(1-\epsilon)f(1)$, we have $p(1|a, \delta', f'(0), f'(1), \mathbf{\Lambda}) \geq p(1|a, \delta, f(0), f(1), \mathbf{\Lambda})$ by (3) and (5). Applying Lemma 2, we obtain from (30) that

$$\begin{aligned} &Q_t(\mathbf{\Lambda} \,|\, a, \delta', f'(0), f'(1)) - Q_t(\mathbf{\Lambda} \,|\, a, \delta, f(0), f(1)) \\ &\quad\geq p(0|a, \delta', f'(0), f'(1), \mathbf{\Lambda})V_{t+1}(\mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta', \{f'(0), f'(1)\}, 0)) \\ &\quad\quad - p(0|a, \delta, f(0), f(1), \mathbf{\Lambda})V_{t+1}(\mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta', \{f(0), f(1)\}, 0)) \\ &\quad\quad + [p(1|a, \delta', f'(0), f'(1), \mathbf{\Lambda}) - p(1|a, \delta, f(0), f(1), \mathbf{\Lambda})]V_{t+1}(\mathcal{T}(\mathbf{\Lambda} \,|\, a, \delta, \{f(0), f(1)\}, 1)). \end{aligned} \tag{31}$$

Applying Lemma 3 to (31) and noting that $p(1|a, \delta, f(0), f(1), \mathbf{\Lambda}) + p(0|a, \delta, f(0), f(1), \mathbf{\Lambda}) = 1$, we obtain

$$Q_t(\mathbf{\Lambda}(t) \,|\, a, \delta', f'(0), f'(1)) - Q_t(\mathbf{\Lambda}(t) \,|\, a, \delta, f(0), f(1)) \geq 0 \tag{32}$$

Hence, the expected remaining reward $Q_t(\mathbf{\Lambda}(t) \,|\, a, \delta, f(0), f(1))$ increases with $\epsilon f(0) + (1-\epsilon)f(1)$. This completes the proof.

□□□

Theorem 2: **Closed-Form Transmission Probabilities** *For any chosen channel* a *and belief vector* $\mathbf{\Lambda}(t)$*, the optimal transmission probabilities* $(f_a^*(0|\delta_a), f_a^*(1|\delta_a))$ *at a fixed sensor operating point* δ_a *are given by*

$$(f_a^*(0|\delta_a), f_a^*(1|\delta_a)) = \begin{cases} (\frac{\zeta-\delta_a}{1-\delta_a}, 1), & \delta_a < \zeta, \\ (0, 1), & \delta_a = \zeta, \\ (0, \frac{\zeta}{\delta_a}), & \delta_a > \zeta. \end{cases} \tag{33}$$

Proof: With the aid of the separation principle, it suffices to show that the transmission probabilities given in (33) maximize the objective function $\epsilon_a f_a(0) + (1-\epsilon_a) f_a(1)$ under the constraint $(1-\delta_a) f_a(0) + \delta_a f_a(1) \leq \zeta$ (see (11)).

Since the transmission probabilities are within [0,1], the constraint in (11) can be written as

$$0 \leq f_a(0) \leq \frac{\zeta - \delta_a f_a(1)}{1-\delta_a}. \tag{34}$$

Applying the upper bound in (34) to the objective function (11a), we obtain that

$$\epsilon_a f_a(0) + (1-\epsilon_a) f_a(1) \leq f_a(1)\left[1 - \frac{\epsilon_a}{1-\delta_a}\right] + \frac{\epsilon_a \zeta}{1-\delta_a}. \tag{35}$$

Since the probability of detection is no less than the probability of false alarms,*i.e.*, $1-\delta_a \geq \epsilon_a$. The upper bound given in (35) increases with $f_a(1)$. Since $f_a(0), f_a(1) \in [0,1]$, the optimal transmission probability $f_a^*(1|\delta_a)$ is given by

$$f_a^*(1|\delta_a) = \min\left\{1, \frac{\zeta}{\delta_a}\right\}. \tag{36}$$

When $\delta_a \leq \zeta$, $f_a^*(1|\delta_a) = 1$ and correspondingly $f_a^*(0|\delta_a) = \frac{\zeta-\delta_a}{1-\delta_a}$. When $\delta_a \geq \zeta$, $f_a^*(1|\delta_a) = \frac{\zeta}{\delta_a}$ and correspondingly $f_a^*(0|\delta_a) = 0$, which completes the proof of Theorem 2.

□□□

Theorem 3: **Deterministic Optimal Policies** *The optimal sensor operating point is* $\delta_a^* = \zeta$*. Correspondingly, the optimal transmission probabilities are given by* $\{f_a^*(0), f_a^*(1)\} = \{0, 1\}$*. Hence, the access decision is deterministic: trust sensing outcome* $\Phi_a = \Theta_a$*.*

Proof: Note that the optimal transmission probabilities obtained in Theorem 2 have ensured the design constrained for each sensor operating point. In the light of the separation principle, it suffices to show that the sensor operating point $\delta_a = \zeta$ maximizes the objective function in (11a).

Substituting the optimal transmission probabilities $(f_a^*(0|\delta_a), f_a^*(1|\delta_a))$ obtained in Theorem 2 to the objective function (11a), we obtain that

$$\epsilon_a f_a(0) + (1-\epsilon_a) f_a(1) = \begin{cases} 1 - \frac{\epsilon_a}{1-\delta_a}(1-\zeta), & \delta_a \leq \zeta, \\ \frac{1-\epsilon_a}{\delta_a}\zeta, & \delta_a \geq \zeta. \end{cases} \tag{37}$$

Since the ROC curve is concave [4], both $\frac{\epsilon_a}{1-\delta_a}$ and $\frac{1-\epsilon_a}{\delta_a}$ increase with ϵ_a and hence decrease with δ_a. From (37), we can see that $\epsilon_a f_a(0) + (1-\epsilon_a) f_a(1)$ increases with δ_a when $\delta_a \leq \zeta$, but decreases when $\delta_a \geq \zeta$. Hence, the maximum $\epsilon_a f_a(0) + (1-\epsilon_a) f_a(1)$ is achieved when $\delta_a^* = \zeta$. Correspondingly, the optimal transmission probabilities $\{f_a^*(0), f_a^*(1)\}$ are given by $\{0, 1\}$.

□□□

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, Systems, and Computers*, Oct. 2006.

[2] Y. Chen, Q. Zhao, and A. Swami, "Joint PHY-MAC design for opportunistic spectrum access: spectrum sensor and sensing/access strategies," *To be submitted to IEEE Transactions on Signal Processing*, Nov. 2006.

[3] R. Smallwood and E. Sondik, "The optimal control of partially observable Markov processes over a finite horizon," *Operations Research*, pp. 1071–1088, 1971.

[4] H. L. V. Trees, "Detection, Estimation, and Modulation Theory, Part I," Wiley-Interscience, Sept., 2001.