Optimal Sensing-Transmission Structure for Dynamic Spectrum Access

Senhua Huang, Xin Liu, and Zhi Ding
University of California, Davis, CA 95616, USA

***Abstract*—In cognitive wireless networks where secondary users (SUs) opportunistically access spectral white spaces of primary users (PUs), there exists an inherent tradeoff between sensing and transmission due to the competing goals of PU protection and SU access maximization. This paper studies means of sensing-transmission for SUs to better manage the competing goals by defining utility function to reward the SU for successful packet transmissions and to penalize it for colliding with PU. To maximize the SU utility, we present a threshold-based sensing-transmission structure that is optimal under a technical constraint. Both perfect sensing and imperfect sensing are considered, with or without SU acknowledgement of reception. This SU access scheme optimizes SU access efficiency while protecting PU performance. It sets a benchmark and provides insight for the design of sensing-transmission control in cognitive networks such as IEEE 802.22.**

I. Introduction

Cognitive radio has recently emerged as a potential technology to mitigate spectrum shortage by allowing secondary users (SUs) to opportunistically utilize spectral white spaces of primary users (PUs). To enable secondary cognitive spectral access, PU protection is vital since PUs would not be inclined to admit SUs without such assurance. To this end, the so-called "Listen-Before-Talk (LBT)" strategy has been extensively considered because it requires little infrastructure support and planning. It should be recognized, however, both SU sensing (listening) and PU collision caused by SU transmission incur cost. In this work, we study optimal sensing and transmission control of SUs in order to maximize SU spectrum usage while providing satisfactory level of PU protection.

Various types of periodic sensing have been considered. The state-of-the-art IEEE 802.22 standard considers periodic in-band measurement (i.e., sensing), using both fast sensing and fine sensing. There also exist prior works that assume a per-packet sensing where the SU senses the PU channel for a certain amount of time before each packet transmission [1], [2], [3], [4], [5]. While periodic sensing, especially per-packet sensing, simplifies the design of PHY/MAC protocol, the question remains as to what the optimal sensing-transmission control should be. This is particularly important when sensing takes a long time in many practical cases. For instance, in IEEE 802.22, the fine sensing time is 25ms for field-sync detection of a standard DTV system in US [6], and a typical sensing time is 10ms for the detection of cordless microphone signals with reasonable fidelity [7]. This sensing time is rather long and not negligible when compared to typical millisecond packet duration. On the other hand, transmission without sensing tends to cause more collisions with the PU, leading to higher penalty. Therefore, we consider both the benefit and penalty when designing sensing-transmission strategies of SUs.

In addition, it is not well understood the impact of feedback information from a secondary receiver (SU-Rx). While the secondary transmitter (SU-Tx) cannot detect primary signals during transmission, the SU-Rx is affected by the PU signal interference when attempting to receive the SU-Tx signal. It can send feedback information to the SU-Tx to indicate the successful packet reception using ACK when PU is inactive, or to indicate the packet failure using NACK (or the absence of an ACK) because of PU interference. Though such feedback may not be reliable, it can be explored by the SU-Tx for better decision making.

Motivated by these issues, we study the following questions in this paper:

- What is the optimal sensing-transmitting control, especially when the sensing time is large, as often in practice?
- What is the impact of (imperfect) feedback, which serves as an indicator of the PU activity during a secondary packet transmission?

To tackle these questions in this paper, we adopt a utility-maximization framework. We define the objective of the SU as to maximize its average channel access time, subject to collision penalty imposed by the PU. We formulate it as an unconstrained stochastic dynamic programming problem that maximizes the average utility/return over time.

Our contributions are as follows: we establish that the optimal spectrum access policy has a simple threshold-based structure, under a technical constraint. The result holds for both perfect and imperfect sensing, and is valid with or without SU-Rx acknowledgement. Note that we assume general PU idle time distributions, which differs from many prior works on dynamic spectrum allocation. Our optimal scheme provides a tight performance upper bound for LBT strategies. It can serve as a benchmark for the performance of the rule-of-thumb periodic sensing and shed light on the design of IEEE 802.22 sensing-access strategies.

II. Related Works

Compared to related works, this paper differs in one or more of the following aspects: 1) general sensing-transmission structure, including but not limited to periodic or per-packet sensing; 2) general unslotted PU idle time distribution; 3) emphasise on the interplay between PU protection and SU performance; and 4) considering perfect or imperfect sensing, with or without SU-Rx feedback. We elaborate as follows.

Channel probing/selection has been studied where the objective is to maximize SU performance by intelligently deciding the sequence of channel probing/sensing and optimal action at each channel, among a set of PU channels. In [8], the authors showed that the optimal strategy on channel probing, guessing, and transmission has a threshold-based structure. The authors assume temporal independence of channel states in a slotted structure and the state of each channel remains unchanged in each slot. In comparison, our objective is to decide the optimal sensing-transmission strategy in one PU channel and we assume general unslotted PU idle time distributions. In [9], the authors considered the MAC-layer sensing scheduling for general PU idle/busy time distributions, and proposed a scheme that optimizes the sensing period and probing sequence to maximize the discovery of spectrum opportunities in multiple channels. In comparison, our objective is to balance the PU protection requirement and the SU access, which is not considered in [9].

In [4], [5], the authors showed that by choosing a suitable sensing time, the SU can achieve an optimal tradeoff between the false-alarm probability and throughput. The works belong to the category of periodic sensing. In comparison, we study general sensing-transmission structure to maximize the SU access time while achieving desired PU protection based on the collision cost and the observation (sensing result and ACK/NACK).

Assuming slotted PU activities, in [1], [10], the authors derived the optimal operating point of the spectrum sensor, access decision, and channel selection with constrained PU collision probability. It was shown that the optimal spectrum access should utilize the whole observation history, and the joint optimization problem can be separated and simplified to an unconstrained problem. Furthermore, authors in [11] show that the results in [10] can be extended to unslotted PU activities. The papers assume that the PU state transition probability does not depend on time, and use a per-packet sensing/transmission structure. Our focus differs and we do not require the above mentioned assumptions.

In our preliminary work [12], we derived an optimal threshold-based policy for the exponential idle time distribution relying on its memoryless property. The approach cannot be generalized to other idle time distributions because the PU state transition depends on time. In addition, perfect sensing is also critical in [12], but not in this paper.

Centralized and distributed channel allocation and scheduling algorithms (e.g. [13], [14]) have been proposed to enable the spectrum sharing among users with cognitive radio capabilities in different network scenarios. In comparison, our objective and assumptions differ.

III. System Model

Consider an SU transceiver pair opportunistically accesses a PU channel. The PU transmits its data at will and exhibits a BUSY/IDLE usage pattern. It is unaware of SU activities. The individual idle (X) and busy (Y) periods are independent. They follow distributions $f_X(\cdot)$ and $f_Y(\cdot)$ with means $\bar{x}$ and $\bar{y}$, respectively.

The SU takes two actions: sensing and transmission. It uses a spectrum sensor (e.g., based on energy or feature) to determine whether the PU is idle or busy at a given time. The non-ideal sensing is characterized by a false-alarm probability P_f and a detection probability P_d. False alarm (detection) represents the event of declaring BUSY when the PU state is IDLE (BUSY). Note that despite the best effort of the SU sensor, sensing of current PU state cannot assure that the PU will be idle in the future because PU traffic is not synchronized or slotted. Thus, if the PU reclaims the channel when the SU is transmitting, a collision happens.

Given Δ as the SU time unit, we assume

- A fixed sensing time $K_S\Delta$, $K_S \in \mathcal{N}$;
- A fixed SU packet length $K_T\Delta$, $K_T \in \mathcal{N}$.
- Relative short sensing/transmission time, i.e., $K_S\Delta, K_T\Delta \ll \bar{x}$, and $K_S\Delta, K_T\Delta \ll \bar{y}$.
- Each successful SU packet earns a reward $R \cdot K_T\Delta$.
- Collision penalty C per time unit Δ.
- $\Delta = 1$ for notation simplicity.

As shown later, for the SU to apply its spectrum access strategy to maximize its average utility/reward (benefit minus collision cost), the PU or the regulator can adjust the collision penalty C to control the aggressiveness of the SU access and the resulting collision rate observed by the PU. Therefore, the PU/SU spectrum sharing system can achieve a desired tradeoff between (SU) spectrum efficiency and PU protection by adjusting C.

We also assume that the SU knows the PU idle time distribution. The distribution function can be estimated through extensive measurement during the deployment of secondary network, through various methods as in [15], [16], [9]. We note that this assumption is no stronger than earlier works, e.g., assuming exponential distribution and the knowledge of its parameter.

In practice, upon receiving a packet from the SU-Tx, the SU-Rx may send an acknowledgement. The acknowledgement here serves two purposes. First, it validates the packet transmission of the SU. More importantly, it provides information to the SU-Tx on whether a collision with the PU may have occurred. In an ideal scenario for sensing collision detection, a NACK (ACK) from the SU-Rx accurately signifies that a collision with the PU has (not) happened during last transmission. However, in wireless communications, the SU-Rx may be able to decode the SU-Tx's packet even when the PT is transmitting. Moreover, a NACK may result from SU channel fading or interference from other users. Therefore, we view the acknowledgement as an inaccurate information source for collision detection, different from most existing works.

We define the following two probabilities:

$$\gamma_1 = Pr[\text{NACK}|\text{Collision with PU}]$$
$$\gamma_0 = Pr[\text{NACK}|\text{No Collision with PU}].$$

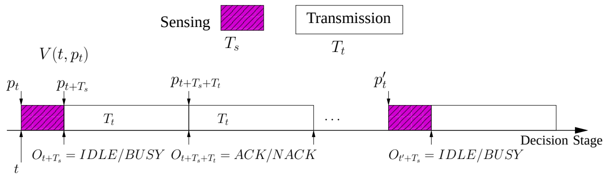

Fig. 1. Sequences of Spectrum Access

Since the interference from the busy primary transmitter can only worsen the packet error rate of SU-Rx, we have $\gamma_1 > \gamma_0$. When there is no capture effect at SU-Rx, i.e., $\gamma_1 = 1$, the SU-Tx can accurately infer the absence of PU during the previous transmission from the feedback ACK message from SU-Rx. Furthermore, if the PHY technique used by the SU link can mitigate the fading and noise fluctuation, then the SU-Rx can decode the packet reliably when there is no PU interference, i.e., $\gamma_0 = 0$. With $\gamma_0 = 0$ and $\gamma_1 = 1$, the SU-Tx can infer accurately that a collision has happened when it receives a NACK from SU-Rx, i.e., the ACK/NACK of the SU can faithfully reflect the event of collision with the PU. In general, we have $0 \leq \gamma_0 \leq \gamma_1 \leq 1$.

IV. Problem Formulation

To exploit the spectrum opportunities in the PU channel effectively while avoiding collision cost, the SU should adaptively schedule sensing and transmission. In other words, the SU can potentially transmit multiple packets or sense multiple times in a row, depending on the situation. Because the SU does not always have complete information on the PU state (IDLE/BUSY), this is a partially observable decision process.

A. PU State Transition

Since the PU's traffic pattern is independent of the SU activities, its state transition does not depend on the SU actions. However, the PU state transition depends on the time instance. Define t as the time elapsed since the PU's most recent state transition from BUSY to IDLE. Given that the PU is idle at time t, the probability that the PU will remain idle during the SU action (transmission or sensing) is

$$g_t^S = \frac{1 - F_X(t + K_S)}{1 - F_X(t)}, \qquad g_t^T = \frac{1 - F_X(t + K_T)}{1 - F_X(t)}, \quad (1)$$

where $F_X(\cdot)$ is the cumulative distribution function of the PU idle time, and superscripts S and T represent that the SU senses the channel and transmits a packet, respectively.

B. Information State

Denote the action space of the SU as $\mathcal{A} = \{a : 1(\text{Transmit}), 0(\text{Sense})\}$. After each action, the SU observes the outcome of its action, O_t. For the sensing action, $O_t^S \in \{I(\text{Sensing IDLE}), B(\text{Sensing BUSY})\}$; for transmission, $O_t^T \in \{A(\text{ACK}), N(\text{NACK})\}$.

Let p_t denote the conditional probability that the PU is IDLE at time t given $(p_0, a_0, \cdots, a_{t-1}, o_0, \cdots, o_{t-1})$, where a_t is the action of the SU at time t, and o_t is the observation after the action at time t. In other words, p_t is the information state at time t. As shown in Fig. 1, at the end of action, the SU will update its estimation on the PU idle probability p_{t+K_S} or p_{t+K_T} based on the observation it received. Specifically, by Bayes' rule, we have the following information dynamics: When $a_t = 0$,

$$p_{t+K_S}(O^S) = \begin{cases} \frac{p_t g_t^S (1-P_f)}{p_t g_t^S (1-P_f) + (1-p_t g_t^S)(1-P_d)}, & O^S = I \\ \frac{p_t g_t^S P_f}{p_t g_t^S P_f + (1-p_t g_t^S) P_d}, & O^S = B. \end{cases} \quad (2)$$

When $a_t = 1$,

$$p_{t+K_T}(O^T) = \begin{cases} \frac{p_t g_t^T (1-\gamma_0)}{p_t g_t^T (1-\gamma_0) + (1-p_t g_t^T)(1-\gamma_1)}, & O^T = A \\ \frac{p_t g_t^T \gamma_0}{p_t g_t^T \gamma_0 + (1-p_t g_t^T)\gamma_1}, & O^T = N. \end{cases} \quad (3)$$

Since we assume that the SU knows the beginning of the PU idle time, we have $p_0 = 1$.

C. Immediate Reward

Given the current idle probability p_t, the probability that the PU will remain idle during a SU packet duration K_T at time t is $p_t g_t^T$. Therefore, the probability that the SU successfully transmits a packet (considering capture effect) is $p_t g_t^T (1 - \gamma_0) + (1 - p_t g_t^T)(1 - \gamma_1)$. Hence, the immediate expected reward/utility the SU obtains at time t, $r_t(p_t, a_t)$, with information state p_t and action a_t, is

$$\begin{aligned} & r_t(p_t, 1) = [p_t g_t^T (1 - \gamma_0) R + (1 - p_t g_t^T)((1 - \gamma_1) R - C)] K_T \\ & r_t(p_t, 0) = 0. \end{aligned} \quad (4)$$

D. Optimality Criterion

Let X_l/Y_l be the length of the lth idle/busy period of the PU. Since $X_l's$ and $Y_l's$ are independent and identically distributed (i.i.d.) random variables, and $X_l's$ are independent with $Y_l's$, the SU spectrum access in a long time duration can be treated as repeated trials of access policy in a PU idle-busy cycle. An access policy of the SU is a sequence of functions $\pi = [d_0, d_1, \cdots, d_t, \cdots]$, where d_t defines a mapping from the SU belief space $\mathcal{P} = \{p_t\} = [0, 1]$ to the action space $\mathcal{A}$ at time t (recall that t is the time elapsed since the beginning of the latest idle period). Here, we consider the case where the SU adopts the same spectrum access policy π in each idle/busy period. Our objective is to find a sensing-transmission strategy that maximizes the expected average utility per time unit, i.e.,

$$\max_{\pi} \qquad \lim_{L \to \infty} \frac{(\sum_{l=1}^{L} \sum_{t=0}^{N_l} r_t(p_t, a_t))/L}{(\sum_{l=1}^{L} (X_l + Y_l))/L},$$

where, N_l is the number of actions taken in the lth idle-busy period. The total utility the SU obtain during each idle-busy cycle is identically and independently distributed. Therefore, by the law of large numbers, we have:

$$\lim_{L \to \infty} \frac{(\sum_{l=1}^{L} \sum_{t=0}^{N_l} r_t(p_t, a_t))/L}{(\sum_{l=1}^{L} (X_l + Y_l))/L} \overset{a.s.}{=} \frac{E[\sum_{t=0}^{N_l} r_t(p_t, a_t)]}{\bar{x} + \bar{y}}.$$

Since $\bar{x}+\bar{y}$ is fixed, we translate the average-utility optimization to a problem that aims to maximize the total expected utility in an idle-busy period. The optimal spectrum access policy is the policy achieving the maximum utility function

$$V(0,p)=\sup_{\pi} V_{\pi}(0,p), \quad \text{and}$$
$$V_{\pi}(0,p)=E_{\pi}[\sum_{t=0}^{N_l} r_t(a_t,p_t)|p_0=p],$$

where $V_{\pi}(0,p)$ is the utility achieved by policy π.

Note here we assume that the SU can detect the beginning of the idle period; i.e., the value of t is known perfectly. This assumption is reasonable. Since the PU busy time is long compared with the SU sensing time, to avoid collision penalty, the SU performs multiple sensing on the primary channel until it finds the PU idle. With a reasonable detection performance, the SU can detect the beginning of the PU idle period. The impact of inaccuracy on this detection is studied through simulations in Section VII-E.

E. Optimality Equation

Define $V(t,p)$ as the maximum expected utility the SU can obtain at time t with information state p. We have

$$V(t,p)=\max_{\{0,1\}}\Big\{L(t,p),M(t,p)\Big\}, \tag{5}$$

where $L(t,p)$ and $M(t,p)$ are the expected utility the SU can obtain by sensing the channel (**L**istening) and trans**M**itting a packet, respectively. Specifically, we have

$$\begin{aligned} L(t,p) &= \sum_{i\in\{I,B\}} Pr[O^S_{t+K_S}=i]V(t+K_S,p_{t+K_S}(i)) \\ M(t,p) &= [pg^T_t(1-\gamma_0)R+(1-pg^T_t)((1-\gamma_1)R-C)]K_T \\ &\quad + \sum_{j\in\{A,N\}} Pr[O^T_{t+K_T}=j]V(t+K_T,p_{t+K_T}(j)). \end{aligned} \tag{6}$$

V. Perfect Sensing without ACK/NACK

In this section, for easy illustration, we consider a simple scenario where the spectrum sensor can detect the PU state accurately when it senses the channel, i.e., $P_f=0$, and $P_d=1$. We also assume that there is no acknowledgement from SU-Rx to SU-Tx, and a SU transmission is successful if and only if it does not collide with the PU, i.e., $\gamma_0=0$, and $\gamma_1=1$. In this case, the SU updates information state as:

$$\begin{aligned} p_{t+K_S}(O^S) &= \begin{cases} 1, & O^S=I, \\ 0, & O^S=B; \end{cases} \\ p_{t+K_T} &= p_t g^T_t. \end{aligned} \tag{7}$$

Therefore, $V(t,p)$ can be simplified as:

$$\begin{aligned} V(t,p)=\max_{\{0,1\}}\Big\{&pg^S_tV(t+K_S,1)+(1-pg^S_t)V(t+K_S,0), \\ &[pg^T_t(R+C)-C]K_T+V(t+K_T,pg^T_t)\Big\}. \end{aligned}$$

Note we assume that the sensing result is IDLE only if the PU channel is idle during the whole sensing duration.

To avoid collision penalty, the SU should keep sensing the channel if it knows the channel is busy (i.e., $p_t=0$) until the PU becomes idle again, at which time the next idle-busy cycle begins. Hence,

$$\begin{aligned} &V(t,0)=0, \forall t\in[0,X_l+Y_l] \\ &V(t,p)=\max_{\{0,1\}}\Big\{pg^S_tV(t+K_S,1), \\ &\qquad [pg^T_t(R+C)-C]K_T+V(t+K_T,pg^T_t)\Big\}. \end{aligned} \tag{8}$$

Lemma 1. *$V(t,p)$ is a convex function of p for a given t.*

Proof: For $0\le\lambda\le 1$, $0\le p_1,p_2\le 1$, we show that $V(t,\lambda p_1+(1-\lambda)p_2)\le \lambda V(t,p_1)+(1-\lambda)V(t,p_2)$, following a similar argument as in [17]. Suppose that the initial state p is determined by flipping a biased-coin with probability λ that a head appears. If a head appears, we choose the initial state as $p=p_1$, else, $p=p_2$. If we know the outcome of the coin flipping, the most reward we can get is $\lambda V(t,p_1)+(1-\lambda)V(t,p_2)$. On the other hand, the most reward we can get if we do not know the outcome of the coin flipping is $V(t,\lambda p_1+(1-\lambda)p_2)$. Since we can always do no worse with more information, the result holds. ■

Note that this proof does not require any assumption on the sensing performance and ACK/NACK mechanism, and thus applies to all cases in the paper.

Next, we define a technical condition under which we can derive an optimal threshold-based policy. Define

$$T^*=\min\{\tau: g^T_t<\frac{C-(1-\gamma_1)R}{(\gamma_1-\gamma_0)R+C}, \forall t>\tau\}. \tag{9}$$

When $\gamma_0=0$, $\gamma_1=1$, this condition is simplified to

$$T^*=\min\{\tau: g^T_t<C/(R+C), \forall t>\tau\}. \tag{10}$$

The physical meaning of T^* is that the SU should not transmit after T^*. To elaborate, recall that g^T_t is the conditional probability that the PU will not return to the channel during the SU's transmission of length K_T, given that the PU is idle at time t. Relying on (9) and the definition of immediate reward $r_t(p_t,a_t)$ as in (4), we have $\forall t>T^*$, $r_t(1,1)<0$, i.e,. all transmissions after T^* will result in a negative immediate return even when the channel is known to be idle at time t. Therefore, $V(t,1)=0, \forall t>T^*$.

Two cases exist: $T^*<\infty$ and $T^*=\infty$. In the first case, we present an optimal threshold-based policy in this paper. In the latter case, we need approximations to calculate the optimal policy. The physical meaning of $T^*<\infty$ is as follows: for t large enough, the probability of primary return during a SU transmission is high enough, and thus the potential cost outweighs the potential gain, if C is chosen appropriately. Many distributions satisfy the condition $T^*<\infty$. Examples of such distribution include uniform distribution, Gaussian distribution, Rayleigh distribution, a class of Weibull distributions with the shape parameter $\beta>1$, and a class of beta distributions with parameter $\alpha\ge 1$.

The condition, $T^*<\infty$, does not hold for all distributions. For example, for Weibull distribution with $\beta<1$, T^* may be infinite because g_t^T is an increasing function of t. Such distributions are not as common. In this case, we need to set T^* large enough so that $P(X > T^*)$ is small enough and we approximate $V(t,1) = 0$ for $t > T^*$.

Exponential distribution is a special case, where g_t^T is a constant due to its memoryless property. By exploiting it, an optimal scheme was derived in [12], where an optimal-stopping technique is used that does not involve T^*. The technique does not apply to the general case where the PU state transition may depend on time t.

In the rest of the paper, we assume $T^* < \infty$. We note that the PU busy time distribution is insignificant. The results in this paper apply to all busy time distributions.

We first introduce the following lemma.

Lemma 2. *$V(t,p)$ increases in p for a given t.*

Proof: We use backward induction to prove this Lemma. By convexity, we have $V(t,p) \le pV(t,1)$, and thus $V(t,p) = 0\ \forall p \in [0,1]$ when $t \ge T^*$. Therefore, $\forall t \ge T^*$, $V(t,p)$ is an increasing function of p. Suppose that for $t \ge T^* - k$, $V(t,p)$ increases in p. Then, for $t = T^* - k - 1$, we have the first term on the RHS in (8), i.e., $pg_t^S V(t+K_S,1)$, increases in p linearly with constant slope $g_t^S V(t+K_S,1) \ge 0$; and the second term, i.e., $[pg_t^T(R+C) - C]K_T + V(t+K_T, pg_t^T)$, also increases in p by the induction hypothesis. Therefore, $V(T^* - k - 1, p)$ is an increasing function of p. ∎

With the previous two lemmas, we are ready to show the structure of the optimal spectrum access strategy.

Theorem 1. The following policy is optimal

$$a^*(t) = \begin{cases} 1 \text{ (Transmit)}, & \text{if } p_t > p_t^* \\ 0 \text{ (Sense)}, & \text{otherwise}, \end{cases} \tag{11}$$

where $0 \le p_t^* \le 1$ is a threshold, which depends on time t.

Proof: The first term on the optimality equation (8) is a linear increasing function of p with a non-negative constant slope $g_t^S V(t+K_S,1)$ for any given t. The second term on the optimality equation (8) is a convex and increasing function of p because $V(t+K_T,p)$ is a convex function of p. In addition, $L(t,0) = 0 > -C = M(t,0)$. Therefor, there exists at most one intersection between the two functions, as illustrated in Fig. 2. This intersection, when exists, is the threshold p_t^*. When the intersection does not exist, we set the threshold $p_t^* = 1$. Because $M(t,p) \le L(t,p)$ for $p \le p_t^*$ and $M(t,p) \ge L(t,p)$ for $p > p_t^*$, Eq. (11) defines an optimal policy. ∎

Next, we present a property for a special class of idle time distributions where $\frac{f_X(x)}{1-F_X(x)}$ is a decreasing function of x. In this case, we have

Lemma 3. *$V(t,p)$ is a decreasing function of t.*

Proof: We prove the above result by induction. Because $\frac{f_X(x)}{1-F_X(x)}$ decreases in x, from (1), we have both g_t^S and g_t^T decrease in t. For $t \ge T^*$, we have $V(t,p) = 0$, and thus the result holds trivially. Suppose that the result is true for

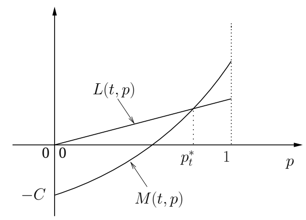

Fig. 2. Illustration of $L(t,p)$ and $M(t,p)$

$t \ge T^* - k$. For $t = T^* - k - 1$, by the induction hypothesis, we have: $V(T^* - k - 1 + K_S, 1) \ge V(T^* - k + K_S, 1)$. Because g_t^S decreases in t, we have $L(T^*-k-1,p) \ge L(T^*-k,p)$. Furthermore, $g_{T^*-k-1}^T \ge g_{T^*-k}^T$, and thus $pg_{T^*-k-1}^T \ge pg_{T^*-k}^T$. According to Lemma 1, $V(t,p)$ increases in p, we have $V(T^*-k-1+K_T, pg_{T^*-k-1}^T) \ge V(T^*-k-1+K_T, pg_{T^*-k}^T)$. Again, by the induction hypothesis, we have $V(T^*-k-1+K_T, pg_{T^*-k}^T) \ge V(T^*-k+K_T, pg_{T^*-k}^T)$. As a result, $M(T^*-k-1,p) \ge M(T^*-k,p)$. Therefore, we have

$$\begin{aligned} V(T^*-k-1,p) &= \max\Big\{L(T^*-k-1,p), M(T^*-k-1,p)\Big\} \\ &\ge \max\{L(T^*-k,p), M(T^*-k,p)\Big\} \\ &= V(T^*-k,p), \end{aligned}$$

which completes the proof. ∎

For these distributions, $M(t,p)$ increases in p. Then, for decreasing g_t^T, we can find the value of T^* by solving

$$g_{T^*}^T = \frac{C - (1-\gamma_1)R}{(\gamma_1 - \gamma_0)R + C}.$$

For general distributions that satisfy the technical constraint $T^* < \infty$, we can first find a t large enough so that condition (9) is satisfied. Then, we can use backward induction find the optimal threshold p_t^* and the maximum utility function $V(t,p)$.

In Fig. 3, we show the optimal threshold for cases where $C = 10$, 20, and $K_S = 1, 5, 30$. The PU idle time is uniformly distributed in $[0, 1000]$, and the packet length is set as $K_T = 5$. The per time unit reward is set to be $R = 1$. We have the following observations: The threshold is time-dependent, and there is no obvious structure. The larger the K_S, the smaller the threshold; i.e., the SU senses less often when sensing time is larger. The larger the C, the larger the threshold; i.e., the SU senses more often to avoid collision penalty. We also note that T^* is smaller for larger C, as shown in the figure.

VI. General Cases

In this section, we consider more general scenarios, i.e., with SU-Rx acknowledgement, and with imperfect sensing.

A. Perfect Sensing with (Imperfect) ACK/NACK

Assuming sensing is perfect, we consider the situation where the SU-Rx sends an ACK to the SU-Tx for a successful transmission. The probability of receiving ACK/NACK is

$$\begin{aligned} &Pr(O_{t+K_T}^T = A) = pg_t^T(1-\gamma_0) + (1 - pg_t^T)(1-\gamma_1) \\ &Pr(O_{t+K_T}^T = N) = 1 - Pr(O_{t+K_T}^T = A). \end{aligned} \tag{12}$$

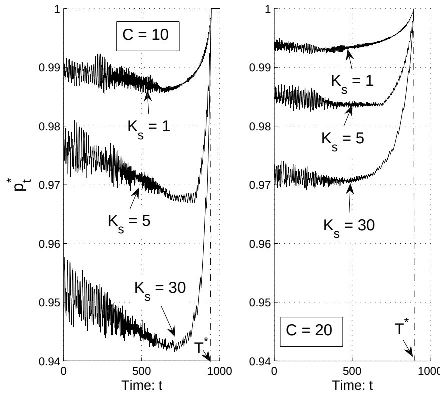

Fig. 3. Threshold for SU Access Decision

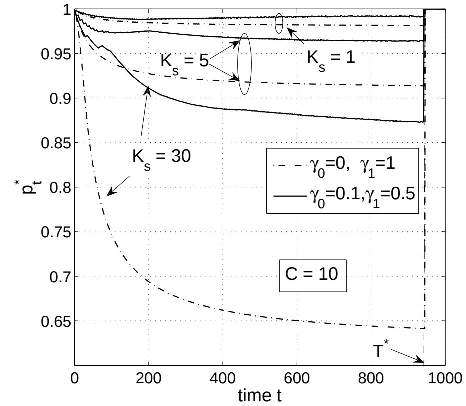

Fig. 4. Threshold for SU Access Decision with ACK/NACK

The SU-Tx updates its belief upon the reception of ACK/NACK, $p_{t+K_T}(O^T_{t+K_T})$, based on (3). The value function at time t given the current information state p is

$$V(t,p) = \max_{\{0,1\}}\Big\{pg_t^S V(t+K_S,1), M(t,p)\Big\}, \tag{13}$$

where $M(t,p)$ is expressed as in (6). Next we show a threshold-based structure, similar to that in previous section, is optimal.

Lemma 4. *With SU-Rx acknowledgement, the optimal utility function $V(t,p)$ increases in p for a given t.*

Proof: We prove the result by induction on t. By convexity, we have $V(t,p) \leq pV(t,1) = 0, \forall t \geq T^*$. Then, the statement holds for $t \geq T^*$ trivially. Suppose it holds for $t \geq T^* - k$. First note that the first term on the RHS of (13) increases linearly with p with a constant slope $g_t^S V(t+K_S,1)$. Thus if we prove the second term on the RHS of (13), i.e., $M(t,p)$, also increases in p, we then prove the Lemma. Take the first-order derivative of $M(t,p)$ with regard to p for $t = T^* - k - 1$ as follows:

$$\begin{aligned}\frac{dM(t,p)}{dp} =& g_t^T((\gamma_1-\gamma_0)R + C)K_T \\ &+ g_t^T(\gamma_1-\gamma_0)V(t+K_T, p_{t+K_T}(A)) \\ &- g_t^T(\gamma_1-\gamma_0)V(t+K_T, p_{t+K_T}(N)) \\ &+ \frac{g_t^T(1-\gamma_0)(1-\gamma_1)V'(t+K_T,p_{t+K_T}(A))}{pg_t^T(1-\gamma_0)+(1-pg_t^T)(1-\gamma_1)} \\ &+ \frac{g_t^T\gamma_0\gamma_1 V'(t+K_T,p_{t+K_T}(N))}{pg_t^T\gamma_0+(1-pg_t^T)\gamma_1}\end{aligned}$$

Because $\gamma_1 > \gamma_0$, we have:

$$p_{t+K_T}(A) > p_{t+K_T}(N),$$

and thus by the induction hypothesis, we have:

$$V(t+K_T, p_{t+K_T}(A)) \geq V(t+K_T, p_{t+K_T}(N)).$$

Thus, the sum of the second and the third terms of $\frac{dM(t,p)}{dp}$ is non-negative. The last two terms of $\frac{dM(t,p)}{dp}$ are also non-negative due to the induction hypothesis that $V(t,p)$ increases in p for all $t \geq T^* - k$. Therefore, $M(t,p)$ increases in p for $t = T^* - k - 1$; and we complete the proof. ■

Now, we are ready to present an SU optimal sensing/transmission strategy in the presence of (imperfect) ACK/NACK:

Theorem 2. Assuming perfect sensing and acknowledgement on collision detection, an optimal policy is:

$$a^*(t) = \begin{cases} 1 \text{ (Transmit)}, & \text{if } p_t > p_t^* \\ 0 \text{ (Sense)}, & \text{otherwise} \end{cases} \tag{14}$$

Proof: The key is to prove that $M(t,p)$ is convex and increasing in p, which is not obvious because of the negative p term involved in (6). See Appendix VIII-A for details. ■

The calculation of p_t^* is similar to that in the previous section. In Fig. 4, we show the optimal threshold for $K_S = 1, 5, 30$. The PU idle time is uniformly distributed over $[0, 1000]$, and the SU transmission length is set to be $K_T = 5$. The per time unit reward is set to be $R = 1$. We observe that the more certain the feedback information is about the PU state (the larger the γ_1 for a fixed γ_0), the smaller the threshold. The threshold is also lower than the case without ACK. This is because that more accurate information about the PU state allows more aggressive SU transmission. Similar to the case without ACK, $p_t^* = 1$ for $t > T^*$.

B. Imperfect Sensing

Now we consider a scenario where the spectrum sensor has imperfect detection performance (P_f, P_d) with/without SU-Rx acknowledgement. First, we have the following result for $V(t,p)$.

Lemma 5. *The optimal utility function $V(t,p)$ with imperfect sensing increases in p for a given t.*

The proof is similar to that in the previous section, and thus omitted here for brevity. Next, we present the following property of $L(t,p)$.

Lemma 6. *$L(t,p)$ is a convex increasing function of p.*

For convenience, recall

$$L(t,p) = \sum_{i\in\{I,B\}} Pr(O^S_{t+K_S} = i)V(t+K_S, p_{t+K_S}(i)).$$

The key is to prove that $L(t,p)$ has non-negative first and second order derivatives with regard to p. Notice that $L(t,p)$ with imperfect sensing is similar to $M(t,p)$ with SU-Rx acknowledgement except that the first-order item ($r_t(p_t,1)$) in $M(t,p)$ is not 0. The Lemma can be proved following a similar argument as in the appendix, and thus omitted here.

Note that we have already shown that the total expected utility of transmitting a packet $M(t,p)$ is a convex and increasing function of p for cases with/without SU-Rx acknowledgement. Together with the above lemma, we next show an optimal structure for the utility-based spectrum access problem. Define $p^*_t = \min\{p : L(t,p) \le M(t,p)\}$, and $p^{**}_t = \max\{p : L(t,p) \le M(t,p)\}$. If the set $\{p : L(t,p) \le M(t,p)\}$ is empty, set both p^*_t and p^{**}_t as 1.

Theorem 3. When the detection performance of sensing is imperfect, the following policy maximizes utility:
If $p^*_t = p^{**}_t$,

$$a^*(t) = \begin{cases} 1 \text{ (Transmit)}, & \text{if } p_t > p^*_t \\ 0 \text{ (Sense)}, & \text{if } p_t \le p^*_t. \end{cases} \quad (15)$$

If $p^*_t < p^{**}_t$,

$$a^*(t) = \begin{cases} 1 \text{ (Transmit)}, & \text{if } p^*_t < p_t < p^{**}_t \\ 0 \text{ (Sense)}, & \text{if } p_t \le p^*_t \text{ or } p_t \ge p^{**}_t. \end{cases} \quad (16)$$

Proof: The key is to prove that both $L(t,p)$ and $M(t,p)$ are convex and increasing in p, which we have already done. Since $L(t,0) = 0 \ge M(t,0)$, we have the above result. ■

The optimal strategy indicated by Theorem 3 is still threshold-based, but it could have two thresholds p^*_t and p^{**}_t.

When $p^{**}_t = p^*_t$, only one threshold exists. For example, we show the single decision threshold in Fig. 5 for different values of K_S. The same settings for the packet length and the PU idle time distribution are used. We also observe a lower threshold for larger sensing time.

VII. Simulation Results

In all simulations, the PU idle time is uniformly distributed in $[0, 1000]$, and the SU packet length is $K_T = 5$. The per time unit reward is $R = 1$, and the mean of PU busy time is set as 500. Therefore, the available spectrum opportunity is 50%. The performance shown in this section is normalized with respect to the sum of average PU idle and busy time. To investigate the impact of small, comparable, and large sensing time on the performance, we set $K_S = 1, 5, 30$ in our simulations.

Our performance metrics are SU average utility, SU throughput, and PU collision rate. SU throughput refers to SU successful transmission time normalized by the PU IDLE-BUSY cycle. SU throughput is upper bounded by the available channel opportunity 50% when there is no capture effect at SU-Rx. PU collision rate quantifies the interruption of the SU transmission on the PU, defined as:

$$P^c_p = \frac{\text{Total Number of PU Packet Collisions}}{\text{Total Number of PU Packets}}, \quad (17)$$

where we assume that the PU packet (time unit) length is Δ without loss of generality. When there is acknowledgement from SU-Rx to SU-Tx, every ACK is treated as a sign for a successful transmission (including capture effect).

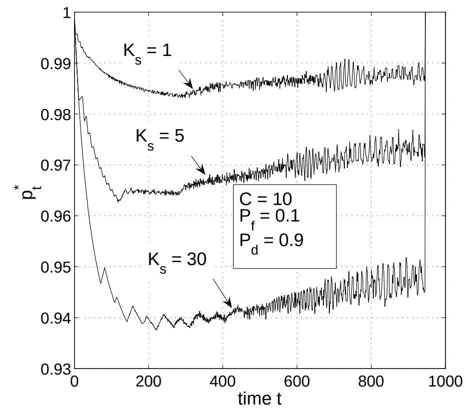

Fig. 5. Threshold for SU Access with Imperfect Sensing

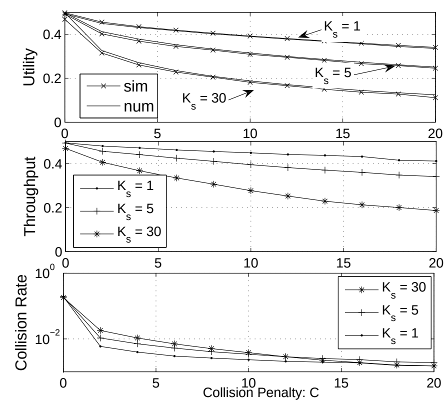

Fig. 6. Performance with Perfect Sensing, Without ACK/NACK

A. Perfect Sensing without ACK/NACK

The performance of the proposed scheme is shown in Fig. 6. We observe that the numerical results obtained via backward induction match with the simulation results. When there is no collision penalty, i.e., $C = 0$, the SU can fully utilize the spectrum opportunities by constantly transmitting packets. The throughput decreases with the collision cost C, and so does the PU collision rate. Therefore, the PU can adjust C to achieve its desired protection (i.e., collision rate), and thus effectively control the aggressiveness of the SU access. The proposed scheme achieves the optimal tradeoff between SU throughput and PU protection for a given C.

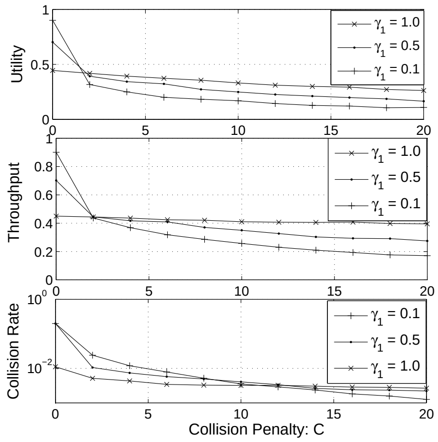

Fig. 7. Performance of Threshold-based Policy with ACK/NACK

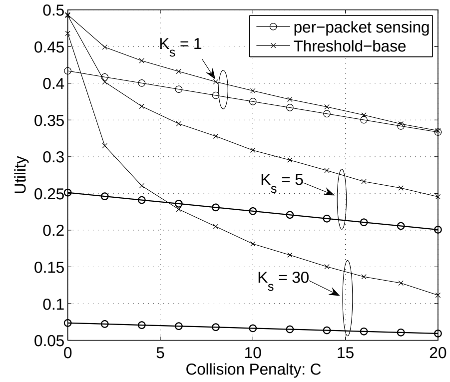

Fig. 8. Utility Comparison of Periodic LBT and Threshold-based Policy

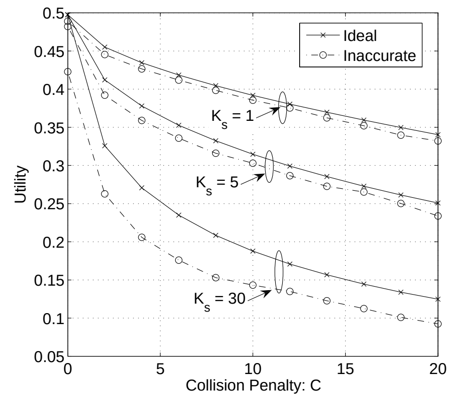

Fig. 9. Average Reward with inaccurate timing synchronization

B. Perfect Sensing with ACK/NACK

With SU-Rx acknowledgement, we consider imperfect SU/PU collision detection where $\gamma_0 = 0.1$, $\gamma_1 = 0.1$, 0.5, and 1. Recall that γ_0 is the packet error rate due to fading or noise, and $\gamma_1 < 1$ represents capture effect at the SU-Rx. The larger the value of γ_1, the larger the posterior probability of PU being idle when $O^T = A$, and thus the more certain the knowledge of SU on the PU state. Simulation result for $K_S = 30$ is shown Fig. 7. We can observe that when collision cost is small (or zero), the SU obtains higher average utility with higher capture effect. However, when C is large, the SU obtains higher average reward with better collision detection. In other words, there exists a tradeoff between the capture effect at SU-Rx and information on PU state. When C is small, smaller γ_1 is desired; when C is large, larger γ_1 is. We also observe that a larger collision cost leads to smaller PU collision rate and smaller SU throughput. Again, the collision cost can effectively balance SU throughput and PU protection. For smaller K_S, the trends are similar, but the performance difference at different values of γ_1 is smaller.

C. Imperfect Sensing

To test the impact of imperfect detection, we set $P_f = 0.1$ and the detection probability P_d at various values. We observe that the SU utility decreases as P_d decreases, a result of higher collision penalty. We also observe increasing the collision penalty reduces the SU throughput and PU collision rate. The simulation result is omitted due to space limit.

D. Comparison with Periodic LBT

We consider periodic LBT where the SU senses the channel for K_S unit time and transmits for K_T unit time if it is idle. The SU keeps sensing if the sensing result is busy. We set $K_T = 5$ and $K_S = 1, 5, 30$. The average utility of periodic LBT and that of our proposed scheme are shown in Fig. 8. For periodic LBT, the SU utility decreases linearly with respect to C because both the SU throughput and the PU collision rate are determined by K_S and K_T, but not C. In comparison, we observe utility gain as large as several-folds for the proposed scheme, especially for long sensing time. Similar observations hold for cases with acknowledgement and imperfect sensing.

We also compare the throughput performance of the periodic LBT and the proposed policy, where we first run simulations using the periodic LBT, and then find the appropriate collision penalty C for the threshold-based strategy to achieve the same PU collision rate. The throughput improvement achieved by the threshold-based policy varies between $2\% - 40\%$. In general, the gain is higher when the sensing time is moderate, and SU-Rx acknowledgement is present.

E. Error in Detecting the Beginning of Idle Period

To study the impact of the inaccuracy in detecting the beginning of idle period, we introduce a random delay, uniformly distributed in $[1, 3K_S]$, to the time that the SU detects the PU transition from BUSY to IDLE. This error, upto 90, is relatively large compared to the average idle length 500 for $K_S = 30$. The impact of imperfect timing is shown in Fig. 9. In the simulation, we assume the sensing is accurate, and there is no ACK/NACK. We observe small performance degradation compared with the ideal detection of PU transition time. The performance loss is larger with longer sensing time due to larger timing errors.

VIII. Conclusions

In this paper, we study general sensing-transmission structures for SUs in a cognitive radio network. The SU receives a reward for successful transmission, and a penalty for colliding with the PU. By maximizing its average utility, the SU achieves a balance between its average channel access time and the collision penalty imposed by the PU. The structure is flexible—the collision cost can be used to control the aggressiveness of the SU access and to limit the interruption on PU transmission. Because the scheme resides in the MAC layer, it simplifies the design of SU networks.

The optimality of our scheme requires that SU has the knowledge of PU idle time distribution, that it can detect the start of PU idle time, and that there exists a finite "stopping time". While these assumptions may not always hold in practice, our results provide a tight performance upper bound for general LBT strategies. They can serve as a performance benchmark for the rule-of-thumb periodic sensing, and shed light on the design of IEEE 802.22 sensing/access strategies.

Our current access structure focuses on one SU (pair) accessing one PU channel. It is future work to study the cases with multiple PU bands and multiple SUs, which is a non-trivial task. Another possible extension is to study the impact of SU power control.

References

[1] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *IEEE Journal on Selected Areas in Communications (JSAC): Special Issue on Adaptive, Spectrum Agile and Cognitive Wireles Networks*, vol. 25, no. 3, pp. 589–600, 2007.

[2] L. Lai, H. E. Gamal, H. Jiang, and H. V. Poor. (2007) Cognitive medium access: Exploration, exploitation and competition. [Online]. Available: http://arxiv.org/abs/0710.1385

[3] Q. Zhao, S. Geirhofer, L. Tong, and B. M. Sadler, "Optimal dynamic spectrum access via periodic channel sensing," in *Proc. Wireless Communications and Networking Conference (WCNC)*, 2007.

[4] P. Wang, L. Xiao, S. Zhou, and J. Wang, "Optimization of detection time for channel efficiency in cognitive radio systems," in *Proc. Wireless Communications and Networking Conference (WCNC)*, 2007, pp. 111–115.

[5] Y.-C. Liang, Y. Zeng, E. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 4, pp. 1326–1337, April 2008.

[6] C. Cordeiro, K. Challapali, D. Birru, and S. Shankar, "IEEE 802.22: An introduction to the first wireless standard based on cognitive radios," *Journal of Communications*, vol. 1, no. 1, pp. 38–47, April 2006.

[7] H.-S. Chen, W. Gao, and D. Daut, "Spectrum sensing for wireless microphone signals," *5th IEEE Annual Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks Workshops, 2008. SECON Workshops '08.*, pp. 1–5, June 2008.

[8] N. B. Chang and M. Liu, "Optimal channel probing and transmission scheduling for opportunistic spectrum access," in *Proc. of the 13th annual ACM international conference on Mobile computing and networking*. New York, NY, USA: ACM, 2007, pp. 27–38.

[9] H. Kim and K. Shin, "Efficient discovery of spectrum opportunities with MAC-layer sensing in cognitive radio networks," *IEEE Transactions on Mobile Computing*, vol. 7, no. 5, pp. 533–545, May 2008.

[10] Y. Chen, Q. Zhao, and A. Swami, "Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 2053–2071, 2008.

[11] Q. Zhao and K. Liu, "Detecting, tracking, and exploiting spectrum opportunities in unslotted primary systems," *in Proc. of IEEE Radio and Wireless Symposium (RWS)*, 2008.

[12] S. Huang, X. Liu, and Z. Ding, "On optimal sensing and transmission strategies for dynamic spectrum access," *Poster Session, Third IEEE International Symposium on Dynamic Spectrum Access Networks, DySPAN*, October 2008.

[13] Y. Yuan, P. Bahl, R. Chandra, T. Moscibroda, and Y. Wu, "Allocating dynamic time-spectrum blocks in cognitive radio networks," in *Proc. of the 8th ACM international symposium on Mobile ad hoc networking and computing*. New York, NY, USA: ACM, 2007, pp. 130–139.

[14] J. Zhao, H. Zheng, and G.-H. Yang, "Distributed coordination in dynamic spectrum allocation networks," *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 259–268, 8-11 Nov. 2005.

[15] A. Azzalini, "A note on the estimation of a distribution function and quantiles by a kernel method," *Biometrika*, vol. 68, no. 1, pp. 326–328, April 1981.

[16] S. Wang, "Nonparametric estimation of distribution functions," *Metrika*, vol. 38, no. 1, pp. 259–267, December 1991.

[17] S. M. Ross, *Introduction to Stochastic Dynamic Programming*. Orlando, FL, USA: Academic Press, Inc., 1983, ch. 3, pp. 58–59.

Appendix

A. Proof of Theorem 2

Proof: If we prove that the expected utility of transmission with ACK/NACK $M(t,p)$ is convex and increasing in p, we can reach the result using the similar argument as in Theorem 1. We have already proved that $\frac{dM(t,p)}{dp} \geq 0$. Next, we show that $\frac{d^2M(t,p)}{dp^2} \geq 0$.

$$
\begin{aligned}
&\frac{d^2M(t,p)}{dp^2}\\
=&\frac{g_t^T(\gamma_1-\gamma_0)g_t^T(1-\gamma_0)(1-\gamma_1)}{(pg_t^T(1-\gamma_0)+(1-pg_t^T)(1-\gamma_1))^2}V'(t+K_T,p_{t+K_T}(A))\\
&-\frac{g_t^T(\gamma_1-\gamma_0)g_t^T\gamma_0\gamma_1}{(pg_t^T\gamma_0+(1-pg_t^T)\gamma_1)^2}V'(t+K_T,p_{t+K_T}(N))]\\
&+\frac{g_t^T(\gamma_0-\gamma_1)g_t^T(1-\gamma_0)(1-\gamma_1)V'(t+K_T,p_{t+K_T}(A))}{(pg_t^T(1-\gamma_0)+(1-pg_t^T)(1-\gamma_1))^2}\\
&+(\frac{g_t^T(1-\gamma_0)(1-\gamma_1)V''(t+K_T,p_{t+K_T}(A))}{pg_t^T(1-\gamma_0)+(1-pg_t^T)(1-\gamma_1)})^2\\
&+\frac{g_t^T(\gamma_1-\gamma_0)g_t^T\gamma_0\gamma_1V'(t+K_T,p_{t+K_T}(N))}{(pg_t^T\gamma_0+(1-pg_t^T)\gamma_1)^2}\\
&+(\frac{g_t^T\gamma_0\gamma_1}{pg_t^T\gamma_0+(1-pg_t^T)\gamma_1})^2V''(t+K_T,p_{t+K_T}(N))\\
=&(\frac{g_t^T(1-\gamma_0)(1-\gamma_1)}{pg_t^T(1-\gamma_0)+(1-pg_t^T)(1-\gamma_1)})^2V''(t+K_T,p_{t+K_T}(A))\\
&+(\frac{g_t^T\gamma_0\gamma_1}{pg_t^T\gamma_0+(1-pg_t^T)\gamma_1})^2V''(t+K_T,p_{t+K_T}(N))\\
\geq&0
\end{aligned}
$$

Here, we use the convexity of $V(t,p)$ in p. When $p=0$, $L(t,0)=0$; while $M(t,0)=(1-\gamma_1)R-C<0$. Since $L(t,p)$ is a linearly increasing function of p, and $M(t,p)$ is a convex increasing function of p, therefore, at most one intersection ion exists between $L(t,p)$ and $M(t,p)$, which is the threshold for decision time slot p_t^*. ∎