

Impact of Imperfect Spectrum Sensing on Age of Information in Energy Harvesting Cognitive Radios

Shiyang Leng and Aylin Yener

Wireless Communications and Networking Laboratory (WCAN)
School of Electrical Engineering and Computer Science
The Pennsylvania State University, University Park, PA 16802.
sfl5154@psu.edu yener@engr.psu.edu

Abstract—The metric *age of information* (AoI) has recently been widely employed to quantify the freshness of the information delivered to the destination. This paper investigates long-term average AoI minimization of an energy harvesting secondary user (SU) in a cognitive radio network setting. Specifically, the paper focuses on the impact of imperfect spectrum sensing on AoI minimization in this setting. The SU makes a decision whether to sense the presence of a primary user, and if it determines the spectrum to be unoccupied, may send out a status update. Sensing and updating both cost energy and the sensing decision may be incorrect due to imperfect spectrum sensing. This setting is formulated as an infinite horizon partially observable Markov decision process (POMDP) to derive the optimal policy that minimizes the long-term average AoI of the SU. The existence of the optimal stationary sensing and update policy is proved and the threshold structure of the policy is shown. Numerical results are presented to demonstrate the SU's AoI performance.

I. INTRODUCTION

The concept of *age of information* (AoI) has been recently introduced to measure the freshness of information from the perspective of the destination [1], [2]. AoI is an especially suitable performance metric in time-sensitive wireless communications, for instance, data exchange in vehicular networks [1], where the status of physical processes have to be updated in a timely manner.

AoI is defined as the time elapsed since the generation of the latest successfully received information [1], [2]. In [2], AoI of a point-to-point communication system is characterized in a linear form, i.e., counting time units in these references, and minimized by considering the update inter-arrival time, waiting time in the queue, and transmission/service time over the channel. Peak age of information is introduced in [3], where only the peak age of an update is tracked, and packets are dropped or replaced to decrease age. Reference [4] introduces an AoI penalty function which generalizes linear and nonlinear models of age.

For energy harvesting communications, AoI has been first investigated in [5] by considering an energy harvesting source that performs status updates using the harvested energy. It is shown in [5] that subject to the energy causality constraint, waiting between updates improves AoI as compared to updating as fast as possible. References [6], [7] consider instantaneous update generation, transmission, and delivery,

and derive (asymptotically) optimal update policies in offline and online settings.

Cognitive radio networks (CRN) with energy harvesting have been studied previously with the throughput as the main metric [8]–[10]. In [8], [9], the spectrum sensing and/or accessing decisions are derived by solving partially observable Markov decision processes (POMDPs). Reference [10] proposes a fixed fraction power allocation policy to maximize long-term throughput. In [11], AoI minimization for a cognitive radio network (CRN) is considered. The AoI of the primary user is characterized and minimized from a queuing theoretic perspective. In [12], the AoI of an energy harvesting secondary user in a CRN is minimized by solving a finite-horizon sequential decision-making problem subject to the energy constraint. The optimal sensing and status update policy is derived in [12] considering *perfect* spectrum sensing at the secondary user.

Different from previous works, this paper studies the AoI of the energy harvesting secondary user (SU) with *imperfect spectrum sensing* in cognitive radio. The SU sends status updates to the destination by opportunistically sensing and accessing the spectrum. An update is successful if it is sent over an unoccupied spectrum and fails if a collision with the primary user (PU) occurs. Thus, the imperfect spectrum sensing of the SU impacts the status update process and the AoI evolution in the long run. In order to maintain the status updates received at the destination as fresh as possible, the SU has to make spectrum sensing and status update decisions optimally while satisfying the energy causality constraint. In particular, the SU may behave conservatively when the spectrum sensing accuracy is low, or energy is scarce, or the channel fading condition is poor even if energy is sufficient. Thus, an optimal sensing and update policy that minimizes the long-term average AoI of the SU is of interest. We formulate the sequential decision making problem as a partially observable Markov decision process (POMDP) due to the partial observation of the state of PU. Unlike [12], we consider an infinite horizon, and we prove that there exists an optimal stationary policy that solves the average cost POMDP problem. We further investigate the structure of the optimal policy and show that it is a threshold policy with respect to the

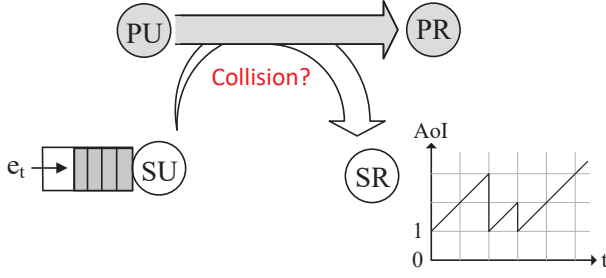


Fig. 1. The system model. The secondary user (SU) senses and accesses the primary user's spectrum opportunistically to send updates to the secondary receiver (SR).

energy state. Numerical results show the impact of imperfect sensing by comparing with the perfect sensing case.

II. SYSTEM MODEL

We consider a cognitive communications scenario where an energy harvesting SU sends status updates to its destination by accessing the PU's spectrum opportunistically over time slots $t = 0, 1, 2, \dots$, as shown in Fig. 1. Without loss of generality, the length of each slot is normalized. The PU has access rights to the spectrum. In each time slot, the PU is either in an active (A) state occupying the spectrum or stays inactive (I) denoted by $q_t \in \{A, I\}$, which forms a stationary two-state Markov chain. The transition probabilities are

$$p_{ii} \triangleq \mathbb{P}(q_{t+1} = I | q_t = I), \quad \forall t, \quad (1)$$

$$p_{ai} \triangleq \mathbb{P}(q_{t+1} = I | q_t = A), \quad \forall t, \quad (2)$$

which are obtained by long-term measurements and known to the SU.

The SU is slot-synchronized with the PU. At the beginning of each slot, the SU decides to either sense the channel or not. If it stays idle, no action is needed. If it decides to sense, it takes a fixed fraction of one slot to sense the PU's spectrum and obtain the channel state information. The SU obtains an observation of the PU state, denoted by $\hat{q}_t \in \{A, I\}$, which could be erroneous due to imperfect spectrum sensing. Let p_f and p_d denote the probabilities of false alarm and detection:

$$p_f \triangleq \mathbb{P}(\hat{q}_t = A | q_t = I), \quad \forall t, \quad (3)$$

$$p_d \triangleq \mathbb{P}(\hat{q}_t = A | q_t = A), \quad \forall t. \quad (4)$$

The channel state information is obtained causally at each slot if sensing is conducted. We consider a block fading channel with channel gain $\mathbf{h}_t$, where $\mathbf{h}_t$ is a random variable with realization $\mathbf{h}_t = h_t$, and independently identically distributed (i.i.d.) over all t .

After sensing, the SU further decides whether to update status or not during the remainder of the slot. Let $x_t = (w_t, z_t)$ be the decision for slot t , where $w_t \in \{0 \text{ (idle)}, 1 \text{ (sense)}\}$ and $z_t \in \{0 \text{ (not update)}, 1 \text{ (update)}\}$ denote the sensing and update decisions, respectively. We consider that both spectrum sensing and status updating expend the energy harvested from ambient energy sources. The amount of the harvested energy e_t is a random variable with realization $\mathbf{e}_t = e_t$, and i.i.d. over

all t . Each sensing action consumes a fixed amount of energy σ . The energy consumed for an update, denoted by $u(h_t)$, includes a fixed cost for generating an update data packet and a varying cost for transmission, where $u(h_t)$ is nonincreasing with respect to the channel gain h_t [13]. For a battery of capacity $\bar{b}$, the battery state evolves as

$$b_{t+1} = \min\{b_t + e_t - w_t\sigma - z_t u(h_t), \bar{b}\}, \quad \forall t. \quad (5)$$

The energy causality constraint [10] is given by

$$w_t\sigma + z_t u(h_t) \leq b_t. \quad (6)$$

Note that we adopt a store-then-use model, by which the harvested energy is first stored in the battery and then used at the next slot onwards.

We use a linear model for AoI [1], [2], where AoI is defined as the time elapsed since the time instant when the most recently received update is generated, denoted by a_t at the beginning of slot t . Once the SU decides to update status, it generates and transmits a data packet. We consider the generate-at-will scheme [5]–[7], i.e., the data packet is generated when the update decision is made. The amount of update data is small enough that it is generated and transmitted instantaneously when spectrum sensing is completed and received by the end of the current slot. If the update is successfully received, the AoI decreases to 1; otherwise increases by 1. As shown in Fig. 1, the AoI increases between updates and, at the beginning of a slot, drops to 1 if an update is delivered during the last slot since it takes one slot to generate and transmit the update. We consider an error-free channel, i.e., the update can be received successfully if transmitted over an unoccupied spectrum and update failure occurs only if the SU collides with the PU. Thus, we have at the beginning of slot $t + 1$, $\forall t$,

$$a_{t+1} = \begin{cases} 1, & \text{if } x_t = (1, 1), \hat{q}_t = q_t = I, \\ a_t + 1, & \text{otherwise.} \end{cases} \quad (7)$$

Since SU can only partially observe the state of PU by opportunistically sensing and accessing the spectrum, a sufficient statistic of the spectrum availability is obtained. That is, the belief ρ_t , representing the conditional probability of PU being inactive, i.e., $q_t = I$, given SU's action and observation history. After taking action $x_t = (w_t, z_t)$, the belief is updated by $\rho_{t+1} = \Lambda(\rho_t)$ taking into account the imperfect spectrum sensing. Specifically, if the SU stays idle without sensing, the new belief is updated solely based on the underlying Markov chain of the PU state. That is, for $x_t = (0, 0)$,

$$\rho_{t+1} = \Lambda_0(\rho_t) = \rho_t p_{ii} + (1 - \rho_t) p_{ai}. \quad (8)$$

If the SU senses but does not update, the belief is updated based on the sensing result. When the PU is sensed to be active, the SU does not update, i.e., $x_t = (1, 0)$, $\hat{q}_t = A$, the new belief is

$$\rho_{t+1} = \Lambda_{1A}(\rho_t) = \theta_t p_{ii} + (1 - \theta_t) p_{ai}, \quad \text{where} \\ \theta_t \triangleq \mathbb{P}(q_t = I | \hat{q}_t = A) = \frac{\rho_t p_f}{\rho_t p_f + (1 - \rho_t) p_d}. \quad (9)$$

When the PU is sensed to be inactive and the SU decides not to update, i.e., $x_t = (1, 0)$, $\hat{q}_t = \text{I}$, the new belief is given by

$$\rho_{t+1} = \Lambda_{1I}(\rho_t) = \bar{\theta}_t p_{\text{ii}} + (1 - \bar{\theta}_t) p_{\text{ai}}, \text{ where} \\ \bar{\theta}_t \triangleq \mathbb{P}(q_t = \text{I} | \hat{q}_t = \text{I}) = \frac{\rho_t(1 - p_{\text{f}})}{\rho_t(1 - p_{\text{f}}) + (1 - \rho_t)(1 - p_{\text{d}})}. \quad (10)$$

If an update is transmitted, the SU receives a feedback signal from the destination that indicates update success or failure. If the PU is sensed to be inactive, i.e., $\hat{q}_t = \text{I}$, and the SU updates successfully, the sensing result correctly indicates the true state of PU, i.e., $q_t = \text{I}$. This gives

$$\rho_{t+1} = \Lambda_I(\rho_t) = p_{\text{ii}}. \quad (11)$$

If the PU is sensed to be inactive, i.e., $\hat{q}_t = \text{I}$, and the SU fails to update, a miss event has occurred in spectrum sensing and the true state is $q_t = \text{A}$. We have

$$\rho_{t+1} = \Lambda_A(\rho_t) = p_{\text{ai}}. \quad (12)$$

The SU aims to minimize the long-term average AoI in an infinite horizon by making optimal sensing and update decisions over time $t = 0, 1, 2, \dots$. The decisions are made adaptively based on SU's system states. Next, we formulate an infinite horizon POMDP to minimize the long-term average AoI under imperfect spectrum sensing.

III. PROBLEM FORMULATION

The POMDP is formulated as follows.

Actions: The action taken in each slot is $x_t = (w_t, z_t) \in \mathcal{X} \triangleq \{(0, 0), (1, 0), (1, 1) : b_t \geq w_t \sigma + z_t u(h_t)\}$, where $w_t \in \Gamma_w \triangleq \{0, 1 : b_t \geq w_t \sigma\}$ and $z_t \in \Gamma_z \triangleq \{0, 1 : b_t \geq \sigma + z_t u(h_t)\}$.

States: The completely observable state consists of AoI state, battery state, energy harvesting state, and channel state, denoted by $s_t \triangleq (a_t, b_t, e_t, h_t)$. In particular, the age a_t is upper bounded by $\hat{a}$ such that $a_t \in \mathcal{A} \triangleq \{1, 2, \dots, \hat{a}\}$. AoI approaching $\hat{a}$ indicates that the information received at the destination is expired so that there is no need for counting. Without loss of generality, we consider finite sample spaces for the harvested energy and the channel fading level, so that $e_t \in \mathcal{E}$ and $h_t \in \mathcal{H}$, $\forall t$, for finite sets $\mathcal{E}$ and $\mathcal{H}$. As a result, the battery state b_t takes finite values between 0 and $\bar{b}$. Denote the finite set by $\mathcal{B}$. Note that the state space, i.e., $\mathcal{S} \triangleq \mathcal{A} \times \mathcal{B} \times \mathcal{E} \times \mathcal{H}$, is thus finite. The partially observable state is the spectrum occupancy, which is represented by the belief ρ_t . The belief space $\mathcal{I}$ is countably infinite as $t \rightarrow \infty$. The complete system state is denoted by $(s_t, \rho_t) \in \mathcal{S} \times \mathcal{I}$.

Transition probabilities: Taking action x_t , the transition probability is denoted by $p_{x_t}(s_{t+1}|s_t)$ for transitioning from state $s_t = (a_t, b_t, e_t, h_t)$ to state s_{t+1} . Since energy harvesting and channel fading are i.i.d. each, and independent from one another, we have

$$p_{x_t}(s_{t+1}|s_t) = \mathbb{P}(a_{t+1}|a_t, x_t) \mathbb{P}(b_{t+1}|b_t, e_t, h_t, x_t) \cdot p_{\mathcal{E}}(e_{t+1}) p_{\mathcal{H}}(h_{t+1}), \quad (13)$$

where $p_{\mathcal{E}}(\cdot)$ and $p_{\mathcal{H}}(\cdot)$ are the distribution of the harvested energy and the channel level, respectively, that are apriori known at the SU, and $\mathbb{P}(b_{t+1}|b_t, e_t, h_t, x_t) = 1$ if $b_{t+1} = \min\{b_t + e_t - w_t \sigma - z_t u(h_t), \bar{b}\}$. Imperfect sensing results involve in the transition probability of the age. Specifically,

$$\mathbb{P}(a_{t+1}|a_t, x_t) = \begin{cases} 1, & \text{if } z_t = 0, a_{t+1} = a_t + 1, \\ \bar{\theta}_t, & \text{if } z_t = 1, a_{t+1} = 1, \\ 1 - \bar{\theta}_t, & \text{if } z_t = 1, a_{t+1} = a_t + 1, \\ 0, & \text{otherwise,} \end{cases} \quad (14)$$

where $\bar{\theta}_t$ is given in (10).

Cost: The immediate cost $C(s_t)$ of state s_t is defined as the accumulated AoI for slot t , i.e., the area under the age curve of slot t ,

$$C(s_t) = \frac{1}{2} + a_t, \quad \forall t. \quad (15)$$

Policy: Denote the policy $\pi = \{\mu_0, \mu_1, \dots\}$, where μ_t is a deterministic decision rule that maps a system state $(s_t, \rho_t) \in \mathcal{S} \times \mathcal{I}$ into an action $x_t \in \mathcal{X}$, i.e., $x_t = \mu_t(s_t, \rho_t)$. For an infinite horizon, we focus on the set of deterministic stationary policies Π^s , where $\pi = \{\mu_0, \mu_1, \dots\} \in \Pi^s$ such that $\mu_{t_1} = \mu_{t_2}$ when $(s_{t_1}, \rho_{t_1}) = (s_{t_2}, \rho_{t_2})$ for any t_1, t_2 . Thus, we omit the time index in the sequel.

Given initial state (s_0, ρ_0) , the long-term average AoI under policy π is given by

$$J^\pi(s_0, \rho_0) = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=0}^{T-1} C(s_t) | s_0, \rho_0 \right]. \quad (16)$$

The goal is to find an optimal stationary sensing and update policy that solves the long-term average AoI minimization problem

$$\min_{\pi \in \Pi^s} J^\pi(s_0, \rho_0). \quad (17)$$

IV. OPTIMAL POLICY

The infinite-horizon POMDP has a countably infinite set of beliefs $\mathcal{I}$ leading to a countably infinite set of system states. Based on [14, Theorem 4.2], we prove that a solution exists for the POMDP with average cost formulated in (17).

Theorem 1: There exists $(J^*, G(s, \rho))$ that satisfies the Bellman equation

$$J^* + G(s, \rho) = \min_{w \in \Gamma_w} Q^w(s, \rho), \quad \forall (s, \rho) \in \mathcal{S} \times \mathcal{I}, \quad (18)$$

where J^* is the optimal average cost which is a constant for all $(s, \rho) \in \mathcal{S} \times \mathcal{I}$, $G(s, \rho)$ is the relative value function defined in (22), and $Q^w(s, \rho)$ is the Q-function for taking sensing action w , which is given in (24). The optimal policy π^* exists and is obtained by

$$w^*(s, \rho) \in \arg \min_{w \in \Gamma_w} Q^w(s, \rho), \quad (19)$$

$$z^*(s, \rho) \in \arg \min_{z \in \Gamma_z} Q^{1z}(s, \rho), \quad (20)$$

where $Q^{1z}(s, \rho)$ is the Q-function for taking action $(1, z)$ as given in (24). Furthermore, for $\beta \in (0, 1)$, we have

$(1 - \beta)V_\beta(s, \rho) \xrightarrow{\beta \uparrow 1} J^*$, where $V_\beta(s, \rho)$ is the value function of the corresponding discounted cost problem with objective

$$J_\beta^\pi(s_0, \rho_0) = \lim_{T \rightarrow \infty} \mathbb{E} \left[\sum_{t=0}^{T-1} \beta^t C(s_t) | s_0, \rho_0 \right], \quad (21)$$

and $V_\beta(s, \rho) \triangleq \min_{\pi \in \Pi^\pi} J_\beta^\pi(s, \rho)$.

The proof is provided in Appendix A. The relative value function is defined as

$$G(s, \rho) \triangleq \tilde{V}(s, \rho) - \tilde{V}(s^0, \rho^0), \quad (22)$$

where $(s^0, \rho^0) \in \mathcal{S} \times \mathcal{I}$ is a reference system state, and $\tilde{V}(s, \rho)$ is computed as

$$\tilde{V}(s, \rho) = \min_{w \in \Gamma_w} Q^w(s, \rho), \quad (23)$$

$$Q^0(s, \rho) = C(s) + \sum_{s'} p_{00}(s'|s) G(s', \Lambda_0(\rho)), \quad (24a)$$

$$Q^1(s, \rho) = (1 - \eta(\rho)) Q^{1A}(s, \rho) + \eta(\rho) \min_{z \in \Gamma_z} Q^{1z}(s, \rho), \quad (24b)$$

$$Q^{1A}(s, \rho) = C(s) + \sum_{s'} p_{10}(s'|s) G(s', \Lambda_{1A}(\rho)), \quad (24c)$$

$$Q^{10}(s, \rho) = C(s) + \sum_{s'} p_{10}(s'|s) G(s', \Lambda_{1I}(\rho)), \quad (24d)$$

$$Q^{11}(s, \rho) = C(s) + \sum_{s'} p_{11}(s', a' = 1 | s) G(s', \Lambda_I(\rho)) \\ + \sum_{s'} p_{11}(s', a' = a + 1 | s) G(s', \Lambda_A(\rho)), \quad (24e)$$

where $\eta(\rho) = \rho(1 - p_f) + (1 - \rho)(1 - p_d)$ is the probability of observing PU inactive.

Next, we investigate the structure of the optimal policy. We first focus on the discounted cost problem (21) and prove the properties of its value function $V_\beta(s, \rho)$ and the optimal policy π_β , then relate to the average cost problem (17) by Theorem 1. The value function of the discounted problem (21) satisfies the following Bellman equation [14, Theorem 2.1]

$$V_\beta(s, \rho) = \min_{w \in \Gamma_w} Q_\beta^w(s, \rho), \quad (25)$$

$$Q_\beta^0(s, \rho) = C(s) + \beta \sum_{s'} p_{00}(s'|s) V_\beta(s', \Lambda_0(\rho)), \quad (26a)$$

$$Q_\beta^1(s, \rho) = (1 - \eta(\rho)) Q_\beta^{1A}(s, \rho) + \eta(\rho) \min_{z \in \Gamma_z} Q_\beta^{1z}(s, \rho), \quad (26b)$$

$$Q_\beta^{1A}(s, \rho) = C(s) + \beta \sum_{s'} p_{10}(s'|s) V_\beta(s', \Lambda_{1A}(\rho)), \quad (26c)$$

$$Q_\beta^{10}(s, \rho) = C(s) + \beta \sum_{s'} p_{10}(s'|s) V_\beta(s', \Lambda_{1I}(\rho)), \quad (26d)$$

$$Q_\beta^{11}(s, \rho) = C(s) + \beta \sum_{s'} p_{11}(s', a' = 1 | s) V_\beta(s', \Lambda_I(\rho)) \\ + \beta \sum_{s'} p_{11}(s', a' = a + 1 | s) V_\beta(s', \Lambda_A(\rho)). \quad (26e)$$

The discounted problem can be solved by value iteration [15]. We prove the monotonicity of the value function $V_\beta(s, \rho)$ as follows.

Algorithm 1 Relative value iteration algorithm

- 1: For all $(s, \rho) \in \mathcal{S} \times \tilde{\mathcal{I}}$, initialize $\tilde{V}_0(s, \rho) = 0$, choose (s^0, ρ^0) , set $G_0(s, \rho) = \tilde{V}_0(s, \rho) - \tilde{V}_0(s^0, \rho^0)$ and $k = 0$.
 - 2: **repeat**
 - 3: **for** $(s, \rho) \in \mathcal{S} \times \tilde{\mathcal{I}}$ **do**
 - 4: compute $\tilde{V}_{k+1}(s, \rho)$ by (23) and (24) using $G_k(s, \rho)$.
 - 5: Let $G_{k+1}(s, \rho) = \tilde{V}_{k+1}(s, \rho) - \tilde{V}_{k+1}(s^0, \rho^0)$.
 - 6: **end for**
 - 7: **until** $G_k(s, \rho) \rightarrow G(s, \rho)$ for all (s, ρ) , otherwise increase k by 1.
 - 8: $J^* = \tilde{V}(s^0, \rho^0)$, optimal policy is obtained by (19).
-

Proposition 1: (i) $V_\beta(s, \rho)$ is nondecreasing with respect to the AoI state a . (ii) $V_\beta(s, \rho)$ is nonincreasing with respect to battery state b , energy harvesting state e , and channel state h . (iii) $V_\beta(s, \rho)$ is nonincreasing with respect to belief ρ if the transition probabilities of the state of PU given in (1), (2) and the probabilities of false alarm and detection events given in (3), (4) satisfy $\frac{p_{ii}}{p_{ai}} \geq \frac{p_d}{p_f} > 1$.

The proof is provided in Appendix B. The threshold structures of the optimal sensing and update policy of average cost problem (17) are proved in the following theorems.

Theorem 2: For the optimal sensing policy, the SU senses, i.e., $w^*(a, b, e, h, \rho) = 1$, if $e \geq \sigma + \bar{b} - b$.

The proof of Theorem 2 is provided in Appendix C.

Theorem 3: For the optimal update policy, (i) if $e \geq \sigma + \bar{b} - b$, and $z^*(a, b, e, h, \rho) = 1$, then for any $b' \geq b$, $z^*(a, b', e, h, \rho) = 1$; (ii) if $e \geq \sigma + u(h) + \bar{b} - b$, the SU updates, i.e., $z^*(a, b, e, h, \rho) = 1$.

The proof of Theorem 3 is provided in Appendix D. Theorem 2 and 3 imply that if the harvested energy is larger than certain thresholds, sensing and update are implemented.

To solve (18) for $(J^*, G(s, \rho))$ and the optimal policy, we apply the relative value iteration algorithm [15]. The algorithm is summarized in Algorithm 1. Since the set of belief states $\mathcal{I}$ is countably infinite, we approximate it by a finite set $\tilde{\mathcal{I}}$ for a given initial belief ρ_0 . From the initial belief ρ_0 , the belief can be updated to 5 new states in each slot as specified in (8)-(12), and evolves as a belief tree. As t grows, the newly updated beliefs deviate little from the previous ones. Thus, we choose t sufficiently large such that $\tilde{\mathcal{I}}$ can be obtained by including p_{ii} , p_{ai} , ρ_0 and all the updated beliefs till t .

V. NUMERICAL RESULTS

In this section, we present numerical results of AoI for imperfect spectrum sensing. The PU has state transition probabilities $p_{ii} = 0.8$ and $p_{ai} = 0.2$. The probability of detecting an active PU is $p_d = 0.8$. The energy consumption for sensing is $\sigma = 1$. The energy harvesting process is i.i.d. Bernoulli with probability p_e for harvesting $e = 3$ and probability $1 - p_e$ for $e = 0$. For the channel fading level, we set $\mathcal{H} = \{h_1, h_2\}$ with p_h for h_1 , where each level corresponds to an energy cost on update. We set $u(h_1) = 2$ and $u(h_2) = 4$. We compare the proposed optimal policy with a myopic policy.

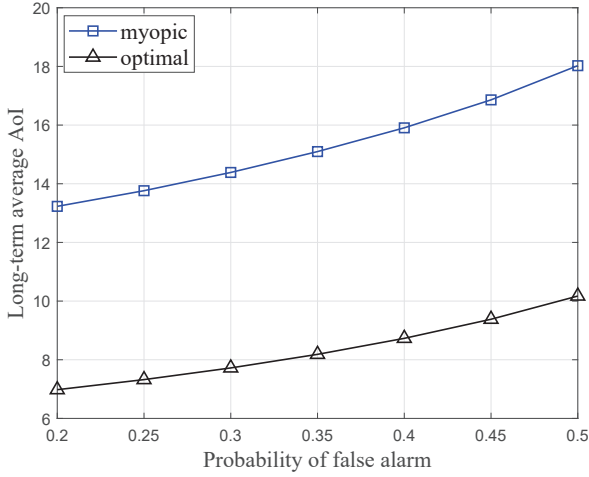


Fig. 2. Average AoI versus the probability of false alarm p_f for $\bar{b} = 5$, $p_e = 0.3$, $p_h = 0.5$, $\hat{a} = 100$, $T = 10000$.

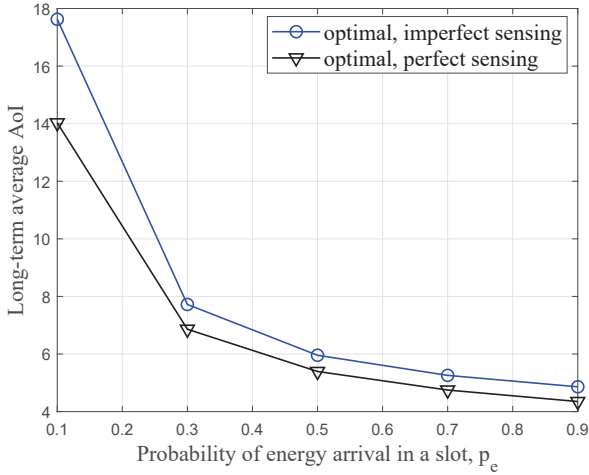


Fig. 3. Average AoI versus the probability of energy arrival p_e for $\bar{b} = 5$, $p_f = 0.3$, $p_h = 0.5$, $\hat{a} = 100$, $T = 10000$.

In particular, in a myopic policy, the SU senses the primary spectrum whenever it has enough energy for sensing. If the primary spectrum is sensed to be unoccupied, an update takes place if the residual energy is sufficient for an update.

Fig. 2 presents the average AoI versus the probability of false alarm. As the probability of false alarm increases, the average AoI becomes larger since the SU observes the spectrum to be occupied and decides not to update due to this sensing error. The optimal policy outperforms the myopic policy significantly since the myopic policy does not take into account the statistics of the spectrum occupancy. Fig. 3 shows the AoI versus the probability of energy arrival. The impact of imperfect spectrum sensing is illustrated by comparing with the perfect spectrum sensing scenario. As expected, the long-term average AoI is larger in imperfect sensing scenario due to the false alarm and miss detection events.

VI. CONCLUSION

We have considered an energy harvesting cognitive secondary user with the objective of AoI minimization. An

infinite horizon POMDP is formulated to investigate the impact of imperfect spectrum sensing on AoI minimization in the long run. The optimal sensing and update decisions that minimize the long-term average AoI are derived and solved by dynamic programming. The threshold structure of the optimal policy is proved. Numerical results highlight the impact of spectrum sensing parameters and demonstrate that the optimal policy significantly outperforms the myopic policy.

APPENDIX A PROOF OF THEOREM 1

According to [14, Theorem 4.2], it suffices to show that the following two conditions are satisfied: (i) $\Lambda_i^{-1}(\rho)$, $\forall i \in \{0, 1A, 1S, A, S\}$ is a countable set; (ii) there is a constant $L \geq 0$ such that $|V_\beta(s, \rho) - V_\beta(s', \rho')| \leq L$, $\forall 0 < \beta < 1$, $\forall (s, \rho), (s', \rho') \in \mathcal{S} \times \mathcal{I}$.

For (i), the condition holds if $\Lambda_i^{-1}(\rho)$ is an injective map. Since $\frac{p_{ii}}{p_{ai}} \geq \frac{p_d}{p_f} > 1$, the matrices $\begin{pmatrix} p_{ii} & 1-p_{ii} \\ p_{ai} & 1-p_{ai} \end{pmatrix}$ and $\begin{pmatrix} p_d & 1-p_d \\ p_f & 1-p_f \end{pmatrix}$ are nonsingular. Thus, $\Lambda_i^{-1}(\rho)$ is an injective map based on [14, Lemma 4.2]. For (ii), consider a system state $(\bar{s}, \bar{\rho}) = (\hat{a}, 0, 0, h_{\min}, 0)$, where $\hat{a}$ is the upper bound of AoI and $h_{\min} \in \mathcal{H}$ is the worst channel level. Due to the monotonicity of $V_\beta(s, \rho)$ as proved below, $0 \leq V_\beta(s, \rho) \leq V_\beta(\bar{s}, \bar{\rho})$ for any $(s, \rho) \in \mathcal{S} \times \mathcal{I}$. Then, it suffices to show that $V_\beta(\bar{s}, \bar{\rho})$ is no larger than a constant L . For imperfect sensing, $\Lambda_0(0) = p_{ai}$, then by (26a), $Q_\beta^0(\bar{s}, \bar{\rho}) = C(\hat{a}) + \beta \sum_{s'} p_{00}(s'|s) V_\beta(s', p_{ai})$. Thus, $V_\beta(\bar{s}, \bar{\rho}) \leq Q_\beta^0(\bar{s}, \bar{\rho}) \leq C(\hat{a}) + \beta V_\beta(\bar{s}, \bar{\rho})$, which results in $V_\beta(\bar{s}, \bar{\rho}) \leq C(\hat{a})/(1 - \beta) \leq (\frac{1}{2} + \hat{a})/(1 - \beta)$. Then, $L = (\frac{1}{2} + \hat{a})/(1 - \beta)$.

APPENDIX B PROOF OF PROPOSITION 1

For clarity of exposition, we omit the notation for irrelevant state components in the sequel. Subscript k denotes the iteration index in value iteration.

(i) *Nondecreasing in a* : We show $V_\beta(a') \geq V_\beta(a)$ for $a' \geq a$ by induction over value iteration. For $k = 0$, $V_{\beta,0}(a') = V_{\beta,0}(a) = 0$. Suppose $V_{\beta,k}(a') \geq V_{\beta,k}(a)$ for some k . From (26a), $Q_{\beta,k}^0(a') \geq Q_{\beta,k}^0(a)$ holds as $C(a') \geq C(a)$ and $V_{\beta,k}(a') \geq V_{\beta,k}(a)$. Similarly, we have $Q_{\beta,k}^{1A}(a') \geq Q_{\beta,k}^{1A}(a)$, $Q_{\beta,k}^{10}(a') \geq Q_{\beta,k}^{10}(a)$ and $Q_{\beta,k}^{11}(a') \geq Q_{\beta,k}^{11}(a)$. Then, $Q_{\beta,k}^1(a') \geq Q_{\beta,k}^1(a)$. By value iteration for (25), $V_{\beta,k+1}(a') \geq V_{\beta,k+1}(a)$. Thus, $V_{\beta,k}(a') \geq V_{\beta,k}(a)$ for all k , and $V_{\beta,k}(a') \rightarrow V_\beta(a)$ as k goes large. Therefore, we conclude $V_\beta(a') \geq V_\beta(a)$ for $a' \geq a$.

(ii) *Nonincreasing in b , e , and h* : The induction in (i) applies here to show that the value function is nonincreasing in b . For energy harvesting state, a larger e' results in a larger battery state in the next slot, which implies a lower value function. Similarly, a better channel state h' leads to a smaller update cost $u(h')$, due to the fact that $u(\cdot)$ is nonincreasing in h' , thus, results in a lower value function.

(iii) *Nonincreasing in ρ* : For $k = 0$, $V_{\beta,0}(\rho') = V_{\beta,0}(\rho) = 0$. Suppose $V_{\beta,k}(\rho') \leq V_{\beta,k}(\rho)$ for $\rho' \leq \rho$. From (8), $\Lambda_0(\rho)$ is nondecreasing in ρ as $p_{ii} > p_{ai}$. Then, $V_{\beta,k}(\Lambda_0(\rho')) \leq V_{\beta,k}(\Lambda_0(\rho))$ by assumption. This implies $Q_{\beta,k}^0(\rho') \leq Q_{\beta,k}^0(\rho)$

according to (26a). Similarly, it can be easily verified from (9) and (10) that $\Lambda_{1A}(\rho)$ and $\Lambda_{1I}(\rho)$ are nondecreasing in ρ , as well as $\Lambda_{1I}(\rho) \geq \Lambda_{1A}(\rho)$ due to $\frac{p_{ii}}{p_{ai}} \geq \frac{p_d}{p_f} > 1 > \frac{1-p_{ii}}{1-p_{ai}}$. Then from (26c)-(26e), $Q_{\beta,k}^{1A}(\rho') \leq Q_{\beta,k}^{1A}(\rho)$, $Q_{\beta,k}^{10}(\rho') \leq Q_{\beta,k}^{10}(\rho)$, and $Q_{\beta,k}^{11}(\rho') \leq Q_{\beta,k}^{11}(\rho)$. Then, for $Q_{\beta,k}^1(\rho')$ given in (26b),

$$Q_{\beta,k}^1(\rho') \leq (1 - \rho')Q_{\beta,k}^{1A}(\rho) + \rho' \min_{z_t \in \Gamma_z} \{Q_{\beta,k}^{10}(\rho), Q_{\beta,k}^{11}(\rho)\} \\ = Q_{\beta,k}^{1A}(\rho) + \rho' \Delta Q_{\beta,k}(\rho), \quad (27)$$

$$\Delta Q_{\beta,k}(\rho) \triangleq \min_{z_t \in \Gamma_z} \{Q_{\beta,k}^{10}(\rho), Q_{\beta,k}^{11}(\rho)\} - Q_{\beta,k}^{1A}(\rho) \\ \leq Q_{\beta,k}^{10}(\rho) - Q_{\beta,k}^{1A}(\rho) \leq 0 \quad (28)$$

The nonpositivity is by (26d) and (26c), where $V_{\beta,k}(\Lambda_I(\rho)) \leq V_{\beta,k}(\Lambda_A(\rho))$ holds by assumption for $\Lambda_I(\rho) = p_{ii} > \Lambda_A(\rho) = p_{ai}$. Therefore, $Q_{\beta,k}^1(\rho') \leq Q_{\beta,k}^{1A}(\rho) + \rho' \Delta Q_{\beta,k}(\rho) \leq Q_{\beta,k}^{1A}(\rho) + \rho \Delta Q_{\beta,k}(\rho) = Q_{\beta,k}^1(\rho)$. By value iteration for (25), we have $V_{\beta,k+1}(\rho') \leq V_{\beta,k+1}(\rho)$.

APPENDIX C PROOF OF THEOREM 2

Here, we show $Q_{\beta}^1(b, \rho) \leq Q_{\beta}^0(b, \rho)$. Then, by Theorem 1, $Q^1(b, \rho) \leq Q^0(b, \rho)$ holds by letting $\beta \rightarrow 1$, thus the statement in Theorem 2 follows. Since $e \geq \sigma + \bar{b} - b$, the new battery state becomes $b' = \min\{b + e - \sigma, \bar{b}\} = \bar{b}$ if sensing, and $b'' = \min\{b + e, \bar{b}\} = \bar{b}$ if not sensing.

$$Q_{\beta}^1(b, \rho) \leq (1 - \eta(\rho))Q_{\beta}^{1A}(b, \rho) + \eta(\rho)Q_{\beta}^{10}(b, \rho) \\ \stackrel{(1)}{\leq} C + \beta \sum_{s'} p_{10}(s'|s) V_{\beta}(\bar{b}, (1 - \eta(\rho))\Lambda_{1A}(\rho) + \eta(\rho)\Lambda_{1I}(\rho)) \\ \stackrel{(2)}{=} C + \beta \sum_{s'} p_{10}(s'|s) V_{\beta}(\bar{b}, \Lambda_0(\rho)) \\ = C + \beta \sum_{s''} p_{00}(s''|s) V_{\beta}(b'', \Lambda_0(\rho)) = Q_{\beta}^0(b, \rho), \quad (29)$$

where C is short for the immediate cost $C(s)$, (1) is by the concavity of $V_{\beta}(s, \rho)$ with respect to ρ , proved in [14, Theorem 2.1] and references therein, and (2) is from the belief update equations in (8)-(10).

APPENDIX D PROOF OF THEOREM 3

(i) Similarly, we need to show that $Q_{\beta}^{11}(b', \rho) \leq Q_{\beta}^{10}(b', \rho)$. The new battery state is $\tilde{b} = \min\{b - \sigma + e, \bar{b}\} = \bar{b}$. Let $b'_{11} = \min\{\bar{b}, b' - \sigma - u(h) + e\}$ and $b_{11} = \min\{\bar{b}, b - \sigma - u(h) + e\}$. Then, $b'_{11} \geq b_{11}$. By (26e),

$$Q_{\beta}^{11}(b', \rho) = C + \beta \sum_{\tilde{s}'} p_{11}(\tilde{s}', \tilde{a}' = 1|s) V_{\beta}(b'_{11}, \Lambda_I(\rho)) \\ + \beta \sum_{\tilde{s}'} p_{11}(\tilde{s}', \tilde{a}' = a + 1|s) V_{\beta}(b'_{11}, \Lambda_A(\rho)) \\ \stackrel{(1)}{\leq} C + \beta \sum_{\tilde{s}} p_{11}(\tilde{s}, \tilde{a} = 1|s) V_{\beta}(b_{11}, \Lambda_I(\rho)) \\ + \beta \sum_{\tilde{s}} p_{11}(\tilde{s}, \tilde{a} = a + 1|s) V_{\beta}(b_{11}, \Lambda_A(\rho)) \\ \stackrel{(2)}{\leq} C + \beta \sum_{\tilde{s}} p_{10}(\tilde{s}|s) V_{\beta}(\tilde{b}, \Lambda_I(\rho)) \\ = Q_{\beta}^{10}(b', \rho), \quad (30)$$

where (1) is by the monotonicity of value function with respect to the battery state, and (2) is due to $z^*(a, b, e, h, \rho) = 1$ implying $Q_{\beta}^{11}(b, \rho) \leq Q_{\beta}^{10}(b, \rho)$.

(ii) By Theorem 2, sensing action is taken. Thus, we only need to show $Q_{\beta}^{10}(a, b, e, h, \rho) \geq Q_{\beta}^{11}(a, b, e, h, \rho)$. The new battery state is $b' = \bar{b}$ if update is transmitted. By (26e),

$$Q_{\beta}^{11}(a, b, e, h, \rho) \\ = C + \beta \sum_{e', h'} p_{\mathcal{E}}(e') p_{\mathcal{H}}(h') \left[\bar{\theta} V_{\beta}(1, \bar{b}, e', h', \Lambda_I(\rho)) \right. \\ \left. + (1 - \bar{\theta}) V_{\beta}(a + 1, \bar{b}, e', h', \Lambda_A(\rho)) \right] \quad (31)$$

$$\leq C + \beta \sum_{e', h'} p_{\mathcal{E}}(e') p_{\mathcal{H}}(h') \left[\bar{\theta} V_{\beta}(a + 1, \bar{b}, e', h', \Lambda_I(\rho)) \right. \\ \left. + (1 - \bar{\theta}) V_{\beta}(a + 1, \bar{b}, e', h', \Lambda_A(\rho)) \right] \quad (32)$$

$$\leq C + \beta \sum_{e', h'} p_{\mathcal{E}}(e') p_{\mathcal{H}}(h') V_{\beta}(a + 1, \bar{b}, e', h', \Lambda_{1I}(\rho)) \quad (33) \\ = Q_{\beta}^{10}(a, b, e, h, \rho), \quad (34)$$

where (32) follows the nondecreasing of value function in AoI state, (33) is due to the concavity of value function, and belief update equations (10)-(12).

REFERENCES

- [1] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, "Minimizing age of information in vehicular networks," in *Proc. 8th Annual IEEE Commun. Society Conf. on Sensor, Mesh and Ad Hoc Commun. and Networks (SECON)*, 2011, pp. 350–358.
- [2] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc. IEEE INFOCOM*, 2012, pp. 2731–2735.
- [3] M. Costa, M. Codreanu, and A. Ephremides, "On the age of information in status update systems with packet management," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 1897–1910, 2016.
- [4] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," *IEEE Trans. Inf. Theory*, vol. 63, no. 11, pp. 7492–7508, 2017.
- [5] R. D. Yates, "Lazy is timely: Status updates by an energy harvesting source," in *Proc. IEEE Intern. Sympos. on Inf. Theory*, 2015, pp. 3008–3012.
- [6] B. T. Bacinoglu, E. T. Ceran, and E. Uysal-Biyikoglu, "Age of information under energy replenishment constraints," in *Proc. IEEE Inf. Theory Workshop*, 2015, pp. 25–31.
- [7] A. Arafa, J. Yang, S. Ulukus, and H. V. Poor, "Age-minimal transmission for energy harvesting sensors with finite batteries: Online policies," *submitted to IEEE Trans. Information Theory*, arXiv:1806.07271, 2018.
- [8] A. Sultan, "Sensing and transmit energy optimization for an energy harvesting cognitive radio," *IEEE Commun. Lett.*, vol. 1, no. 5, pp. 500–503, 2012.
- [9] S. Park and D. Hong, "Optimal spectrum access for energy harvesting cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 6166–6179, 2013.
- [10] B. Varan and A. Yener, "Online transmission policies for cognitive radio networks with energy harvesting secondary users," in *Proc. IEEE Wireless Commun. and Networking Conf.*, 2017.
- [11] A. Kosta, N. Pappas, A. Ephremides, and V. Angelakis, "Age of information and throughput in a shared access network with heterogeneous traffic," *arXiv preprint arXiv:1806.08776*, 2018.
- [12] S. Leng and A. Yener, "Minimizing age of information for an energy harvesting cognitive radio," in *Proc. IEEE Wireless Commun. and Networking Conf.*, 2019.
- [13] Y. Chen, Q. Zhao, and A. Swami, "Distributed spectrum sensing and access in cognitive radio networks with energy constraint," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 783–797, 2009.
- [14] E. Fernández-Gaucherand, A. Arapostathis, and S. I. Marcus, "On the average cost optimality equation and the structure of optimal policies for partially observable markov decision processes," *Annals of Operations Research*, vol. 29, no. 1, pp. 439–469, 1991.
- [15] D. P. Bertsekas, *Dynamic programming and optimal control, Vol. I & II*. Athena scientific, 2005.